Vedic Research International
Bioinformatics & Proteomics
eISSN 2330-0272
JOURNAL HOME PAGE AT WWW.VEDICJOURNALS.COM
RESEARCH ARTICLE
DOI: http://dx.doi.org/10.14259/bp.v2i1.86

Comparative Epitope Stability Analysis of Influenza A Virus Antigenic Proteins

KRISHNA CHAITANYA AMAJALA*, R MEENA, DSVGK KALADHAR

Department of Bioinformatics & Biochemistry, GITAM Institute of Science, GITAM University, Visakhapatnam-530045, Andhra Pradesh, India

Article Info: Received: *December 5th, 2013*; **Accepted:** *December 17th, 2013*

ABSTRACT

Influenza A Viruses are the most virulent human pathogens among the three influenza types (Influenza A, B, C) and cause the most severe disease. The influenza A virus are subdivided into different serotypes based on the antibody response to these viruses and they resulted in many number of human pandemic deaths. The known flu pandemics are from 1918 Spanish flu, which is the serious pandemic in recent history and the most recent one was the 2009 flu pandemic. In each of the pandemics, different subtypes of H & N of Influenza A virus strain are involved. The process of antigenic shift in between species as well subtype combinations are leading to the novel strains which are highly virulent. As the subtypes are based on two important proteins viz. Hemagglutinin and Neuraminidase (H&N), our focus is to identify the amino acid changes/mutations in the epitopes of H1-H16 and N1-N9 strains which are evolved through time. The epitopic regions are predicted using Immune Epitope DataBase, and the predicted epitopic peptide is modeled using the CPH modeling server and each of the modeled epitope(s) are subjected to structural stability analysis using Hyperchem program. Epitope modeling and energy value comparison studies of Influenza A Virus epitopes showed that recent (2010) strains - H5 and N6 are possessing higher energies indicating their instability and thereby are more prone to mutations and may become highly pathogenic (virulent) in the near future than the earlier H1 and N1 strains.

Keywords: *Epitope, Hemagglutinin, Neuraminidase, IEDB*

INTRODUCTION

Influenza viruses are classified into three types (species), A, B and C, on the basis of identity of the major internal protein antigens such as nucleoprotein (NP) and matrix proteins (M1). Influenza A viruses are the most virulent of the three and cause severe and fatal acute respiratory disease, which is epidemic and sometimes worldwide pandemic [1, 2]. Influenza A viruses have been isolated from humans and various animals, including pigs, horses, sea mammals such as seals and whales, wild waterfowl, including ducks, and poultry [3]. Swine influenza was first proposed to be a disease related to human influenza during the 1918 flu pandemic [4]. The key proteins involved in causing the flu are hemagglutinin (H) and/or neuraminidase (N) [5]. As flu viruses are more prone to mutations, there were gradual changes introduced in the proteins of H & N leading 16 numbers of H subtypes (H1-16) and 9 numbers of N subtypes (N1-9). Among those subtypes the H1N1 form of swine flu, one of the descendants of the strain which caused the 1918 flu pandemic [6, 7]. Avian influenza viruses in aquatic birds serve as the natural reservoir for all known subtypes of influenza A virus and probably are the ultimate source of human pandemic influenza strains [8]. The objective of the current study is to identify the mutations in the predicted epitopic regions of hemagglutin (H1-H16) & Neuraminidase (N1-N9) subtypes of Influenza A virus and to analyze the epitopic structural stability through Epitope modeling studies. Hence to perform epitope analysis, we have employed Immunoinformatic resources which

*Corresponding Author
Krishna Chaitanya Amajala
Department of Bioinformatics & Biochemistry, GITAM Institute of Science, GITAM University, Visakhapatnam-530045, Andhra Pradesh, India.
Phone: +91 9966771134; Email: chy2ak@gmail.com

ISSN 2330-0272 (Online); CODEN: VRIBCN, VRI Bioinformatics & Proteomics, VRI Press, Vedic Research Inc. USA

include searchable databases of epitopes and immune-related molecules, as well analysis tools for T cell and B cell epitope prediction [9].

Methods

Retrieving the Protein Sequences of Influenza A Virus (H1-H16) & (N1-N9) using Swiss-Prot Database:
Swiss-Prot is a high-quality, manually annotated, non-redundant protein sequence database. It combines information extracted from scientific literature and biocurator-evaluated computational analysis [10]. The protein sequences of both hemagglutinin (H1-H16) and neuraminidase (N1-N9) subtypes of first isolated in the year of nineties to the recent (upto 2010) are comparatively searched in swiss-prot database and found that only H1, H3, H5 and N1, N2, N6 subtypes are evolved through various mutations continuously till 2010 year. The protein sequences of first isolated and recent isolated subtypes of hemagglutinin (H1, H3, H5) and neuraminidase (N1, N2, N6) are retrieved for further analysis.

T-cell Epitope Prediction Using IEDB Database:
The Immune Epitope Database and Analysis Resource (IEDB) contains data related to antibody and T cell epitopes for humans, non-human primates, rodents, and other animal species [9]. The Protein sequences retrieved from Swiss-Prot Database were submitted to IEDB which gives the Epitope sequences (9 aminoacids length each) with Percentile Scores. The Epitope Sequences with Low Percentile Scores were selected since they are Good Binders and have High Affinity for MHC Class I molecules.

Tertiary structure prediction using CPH Model Server:
CPH models-3.0 is a web-server predicting protein 3D-structure by use of single template homology modeling. The server employs a hybrid of the scoring functions of CPH models-2.0 and a novel remote homology-modeling algorithm [11]. The Epitope Sequences obtained using IEDB were submitted to CPH Model which gives the PDB model for the submitted Epitope sequences.

Comparative Single Point Energy Computation Studies of Influenza A Virus subtypes (H1, H3, H6) and (N1, N2, N6) Epitopes using HyperChem:
HyperChem is a powerful computational software package that is capable of examining potential energy surfaces of molecules. HyperChem can perform an energy minimization or geometry optimization of a molecule using a variety of computational methods [7, 12]. The PDB models of the Epitope sequences downloaded from CPH Model were uploaded in HyperChem for the single point energy calculations each epitope sequence.

Construction of Phylogenetic tree using MEGA:
Molecular Evolutionary Genetic Analysis is an integrated tool for conducting automatic and manual sequence alignment, inferring phylogenetic trees, mining web-based databases, estimating rates of molecular evolution, inferring ancestral sequences, and testing evolutionary hypotheses [13]. All the Protein sequences of Hemagglutinin (H1-H16) and Neuraminidase (N1-N9) previously retrieved from Swiss-Prot Database were subjected to MEGA which constructs Phylogenetic Tree.

Results and Discussion

Prediction of Epitope regions in Hemagglutinin & Neuraminidase subtypes: The protein sequences of both Hemagglutinin and Neuraminidase subtype(s) are subjected to T-cell epitope prediction analysis using IEDB. The sequences information are given in the **Table 1 & 2**. The epitope regions are predicted on the basis of IC50 values for respective alleles. Low IC50 values are good binders. The epitopes which are having low IC50 values are chosen for epitope mutation analysis. From results of the IEDB analysis which are tabulated in the **Table 3 & 4**, it can be inferred that the change in sequence patterns of the epitopic regions of all the subtypes of the first isolate and recent isolate strains are differing, which suggests a residue specific antigenic shift. The change in the residues / epitope sequence mutations of the first isolate and recent (2010) isolate subtypes are indicated in red-colored bold. The percentile indicates the prediction score of the epitope sequence/region. Many different epitopic regions are predicted

Table 1: Hemagglutinin protein sequences retrieved from Swiss-prot database
Recent (2010) strains were found only for H1, H3, and H5 subtypes

Strain Subtype	Year	Accession Number	Organism	Length (AA)
H1	1918	Q9WFX3	Influenza A virus (strain A/Brevig Mission/1/1918 H1N1)	566
	2010	E0V1G0	Influenza A virus (A/Egypt/N00277/2010(H1N1))	
H3	1963	P03442	Influenza A virus (strain A/Duck/Ukraine/1/1963 H3N8)	566
	2010	E1TXU6	Influenza A virus (A/Managua/4009.02/2010(H3N2))	
H5	1959	P09345	Influenza A virus (strain A/Chicken/Scotland/1959 H5N1)	564
	2010	E1B535	Influenza A virus (A/duck/Egypt/106d/2010(H5N1))	567

using IEDB database, among those the lowest percentile based epitope region is considered as it is the good binder to the MHC molecules for antigen presentation. The percentile scores for the predicted epitopes in Haemaglutinin (recent-2010) H1, H3 & H5 subtypes are 0.15, 0.30 & 0.10 and for Neuraminidase (recent 2010) N1, N2 & N6 are 0.20, 43.95 & 0.60 respectively.

Comparative Energy Profile analysis of Epitope & Phylogeny of H & N Subtypes:

Using Hyperchem, the Single Point Energy computation for H & N subtypes are performed. The energy values are tabulated in the **Table 3 & 4.** The highest energy among the first isolate/ older strains and recent strains is exhibited by H5 subtype i.e., 556.99 kcal/mol, followed by H1 subtype. Although there is an increase in the energy of 2010-H3 subtype but comparatively the energy profile of H5 and H1 recent subtypes/strains are greater indicating the instability of the mutation which may leads to further rapid frame shift mutation in the sequence. From the obtained percentile scores, it can be further inferred that the recent H5 strain epitopic region is a good binder to the MHC molecules comparatively with the other H1 & H3 strains. The Phylogenetic Analysis **(Figure 1)** further supports the inference as the H5 and H1 subtypes/strains are clustered together than the H3 subtype.

In the case of Neuraminidase subtypes, the subtype N2 recent (2010) isolate is exhibiting the highest energy i.e., 145.93 kcal/ mol among the other subtypes but the predicted epitope percentile score is 43.95, which is comparatively very higher than N1 & N6 subtypes and hence it can be inferred that it is a poor binder to the MHC molecule for the antigenic presentation. Hence, by considering the lowest percentile score

Table 2: Neuraminidase protein sequences retrieved from Swiss-prot database
Recent (2010) strains were found only for N1, N1, and N6 subtypes

Strain Subtype	Year	Accession Number	Organism	Length (AA)
N1	1933	P03470	Influenza A virus (strain A/Wilson-Smith/1933 H1N1)	453
	2010	E0UYR7	Influenza A virus (A/Athens/INS390/2010(H1N1))	469
N2	1957	Q1K9Q1	Influenza A virus (strain A/Japan/305/1957 H2N2)	469
	2010	E1TXS7	Influenza A virus (A/Managua/4763.03/2010(H3N2))	
N6	1956	Q6XV27	Influenza A virus (strain A/Duck/England/1/1956 H11N6)	470
	2010	D7UT70	Influenza A virus (A/duck/Vietnam/OIE-2334/2010(H9N6))	

Table 3: IEDB results of Peptide Binding to MHC Class I Molecules and Identified Mutations in Epitopes of Haemagglutin (H1, H3, H5 subtypes) and Single point Energy obtained from HyperChem

Strain Subtype	Year	Accession Number	Position	Predicted Epitope Region	Percentile	Single Point Energy Kcal/mol
H1	1918	Q9WFX3	1:358-366	WTGMIDGWY	0.15	553.89
	2010	E0V1G0	1:358-366	WTGMVDGWY	**0.15**	551.93
H3	1963	P03442	1:106-114	RSNAFSNCY	0.30	56.56
	2010	E1TXU6	1:106-114	RSKAYSNCY	**1.15**	505.56
H5	1959	P09345	1:15-23	KSDQICIGY	0.10	556.99
	2010	E1B535	1:15-23	KSDQICIGY	**0.10**	**556.99**

of 0.60 for N6 subtype and the single point energy value of 144.06 kcal/mol it can be deduced that the N6-2010 subtype is unstable form in nature and can led to further mutations in the future at the same epitopic region. The comparative changes in the N2 subtype epitopic residues in both the 1957 & 2010 isolates are very clear indication for the instability in the year of 1957 which is involved of five amino acid residue mutations, which resulted in higher energy profile of single point calculation. The Phylogenetic Analysis **(Figure 2)** further supports the inference as the N6 and N2 subtypes/strains are clustered closer than the N1 subtype.

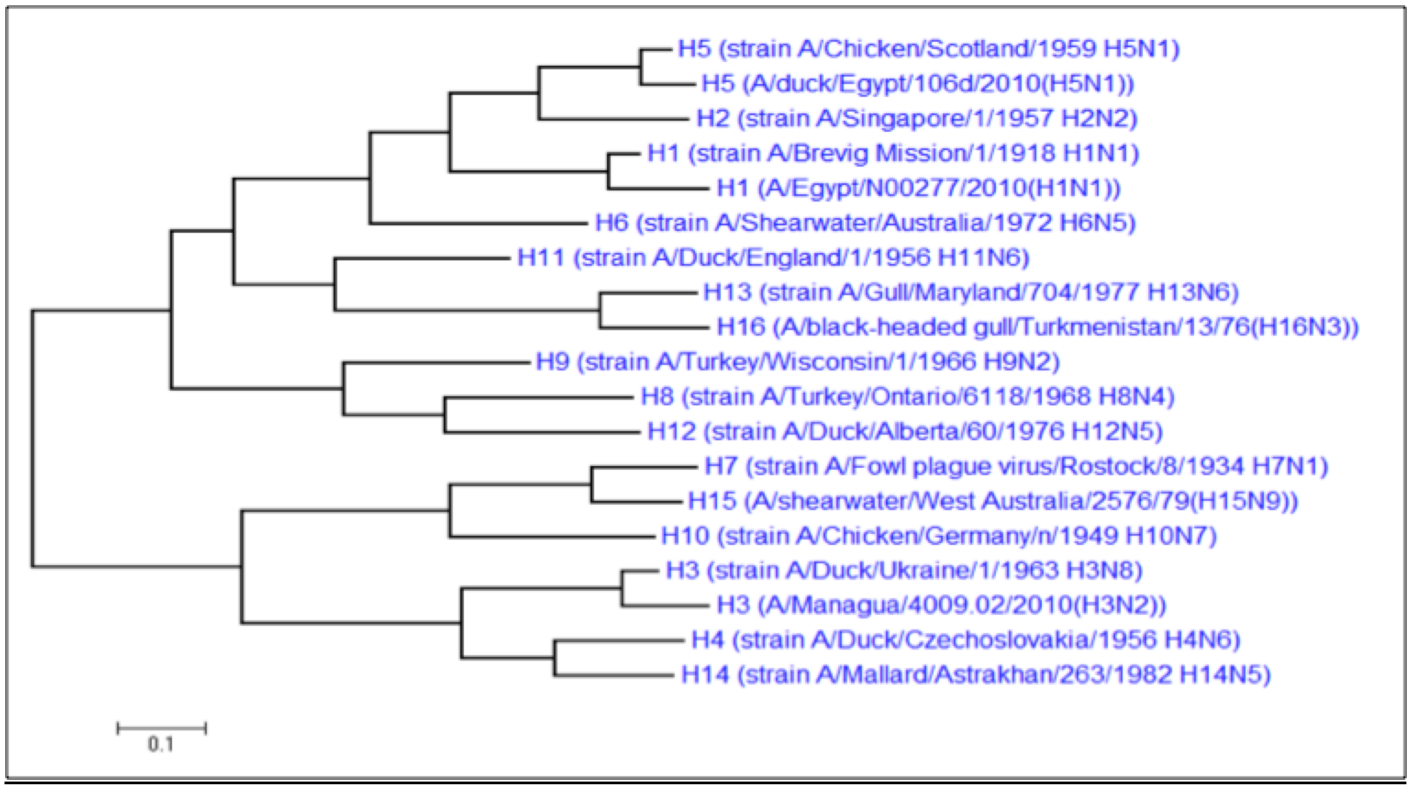
Conclusions

Due to the mutations occurring in the Epitope Positions of Old and Recent 2010 strains of Hemagglutinin and Neuraminidase proteins, there are changes in the Energy Levels as well. From the Energy Comparisons, it is apparent that among the Recent 2010 strains, H5 and N6 are possessing higher energies indicating instability of the sequence/structure conformations, by which we can infer that the mutations possessing pathogenicity (Virulence) in the future evolving strains is comparatively higher than H1 and N1 subtypes/

Table 4: IEDB results of Peptide Binding to MHC Class I Molecules and Identified Mutations in Epitopes of Neuraminidase (N1, N2, N6 subtypes) and Single point Energy obtained from HyperChem

Strain Subtype	Year	Accession Number	Position	Sequence	Percentile	Single Point Energy Kcal/mol
N1	1933	P03470	1:288-296	VSFDQNLDY	0.20	119.08
	2010	E0UYR7	1:304-312	VSFNQNLEY	**0.20**	93.85
N2	1957	Q1K9Q1	1:302-310	VIDINMEDY	0.25	583.33
	2010	E1TXS7	1:302-310	IVDINIKDH	**43.95**	145.93
N6	1956	Q6XV27	1:384-392	ETDIQSGPI	1.15	120.29
	2010	D7UT70	1:384-392	ETDTQSGPV	**0.60**	**144.06**

Figure 1: Phylogeny of Influenza A Virus – Hemagglutinin (H1-H16) using Maximium Likelihood Method



strains respectively. The currently available vaccination for highly variable Influenza virus is of seasonal type consisting of antigens representing trivalent vaccine or quadrivalent vaccine influenza virus strains with the combinations such as: one influenza type A subtype H1N1 virus strain, another influenza type A subtype H3N2 virus strain, and either one or two influenza type B virus strains. As there is no vaccine still under development for the other subtypes of virus strains such as H5 & N6, our studies can be helpful in understanding the mutational rate of the Haemagglutinin & Neuraminidase proteins and the strains variability leading to highly pathogenic future outbreak of the Influenza pandemics.

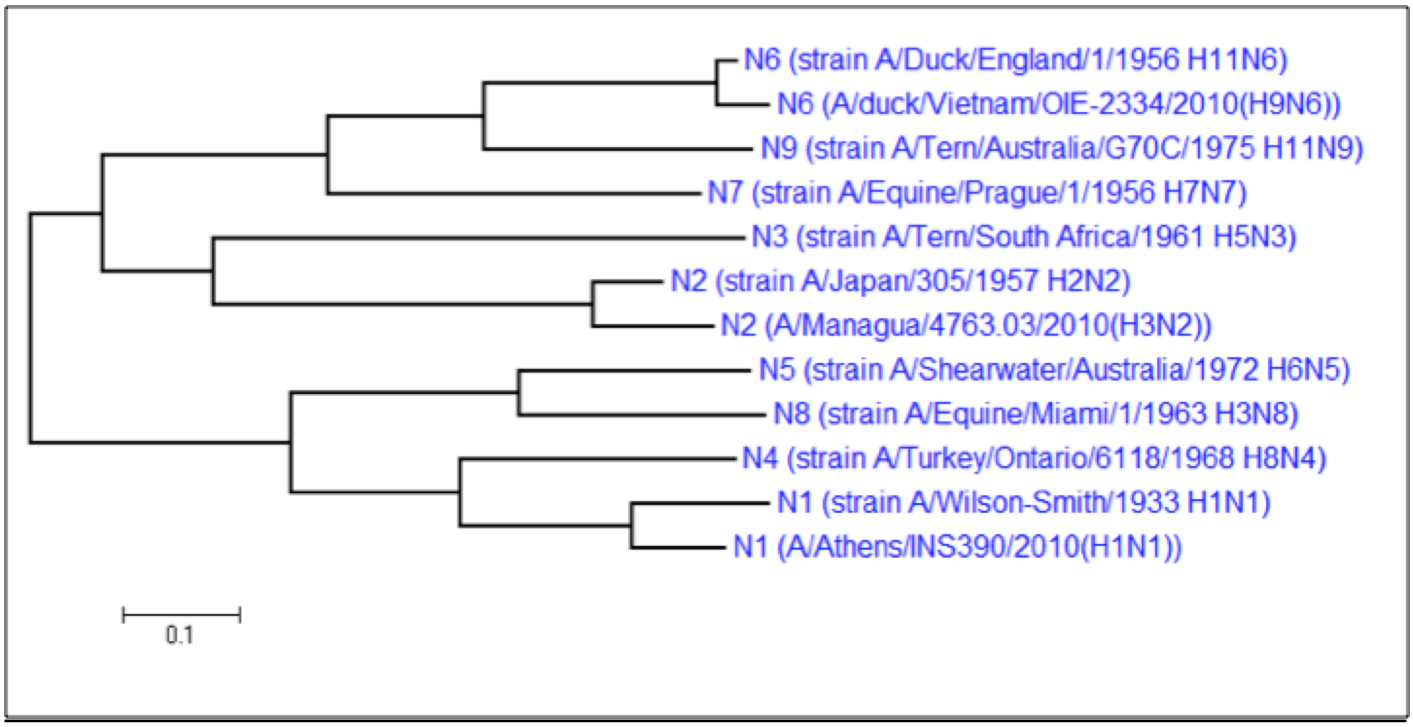
Acknowledgement

We would like to thank GITAM University for providing Bioinformatics lab facility to carry out the research work.

References

1. Yasuo SUZUKI: **Sialobiology of Influenza: Molecular Mechanism of Host Range Variation of Influenza Viruses.** Biol Pharm Bull. 2005 Mar;28(3):399-408.
2. Gong J, Xu W, Zhang: **Structure and functions of influenza virus neuraminidase.** J.Curr Med Chem. 2007;14(1): 113-22.
3. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y: **Evolution and ecology of influenza A viruses.** Microbiol Rev. 1992 Mar;56(1):152-79.
4. Jason AG, Maya FK, Yohan K, Carla O, Kerrie V, Nima S, Randi V, Julia P, Richard HS, Alessandro S, Bjoern P: **Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population,** PNAS, 106(48), 2009, 20365-20370.
5. Easton DM, Nijnik A, Mayer ML, Hancock RE: **Potential of immunomodulatory host defense peptides as novel anti-infectives**. Trends Biotechnol. 27(10), 2009, 582-590.
6. Khanna M, Gupta N: **Influenza A (H1N1) pandemic: preparedness and clinical management**. Indian J Exp Biol, 47(11), 2009, 929-32.
7. Amajala Krishna Chaitanya, VSV Ravi Chandu, DSVGK Kaladhar, I Bhaskar Reddy.: **Epitope Modeling, Comparative Mutation & Structural Stability Analysis of Influenza A Virus H1N1**, Drug Invention Today, 2012, 4(3), 351-354.
8. Taubenberger JK, Morens DM: **The pathology of influenza virus infections.** Annu Rev Pathol. 2008;3:499-522.
9. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, Bourne PE, Bui HH, Buus S, Frankild S, Greenbaum J, Lund O, Lundegaard C, Nielsen M, Ponomarenko J, Sette A, Zhu Z, Peters B: **Immune epitope database analysis resource (IEDB-AR)**. Nucleic Acids Res, 36, 2008, W513-8.
10. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I. et al.: **The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003**. Nucleic Acids Res. 2003;8(1): 365–370. doi: 10.1093/nar/gkg095.
11. Nielsen M., Lundegaard C., Lund O., Petersen TN: **CPHmodels-3.0 - Remote homology modeling using structure guided sequence profiles.** Nucleic Acids Research, 2010, Vol. 38, doi:10.1093/nar/gkq535.
12. **HyperChem(TM) Professional 7.51**, Hypercube, Inc., 1115

Figure 2: Phylogeny of Influenza A Virus – Neuraminidase (N1-N9) using Maximium Likelihood Method



NW 4th Street, Gainesville, Florida 32601, USA.

13. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S: **MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods.** Molecular Biology and Evolution 2011, 28: 2731-2739.

Visit us @ www.vedicjournals.com : DOI: http://dx.doi.org/10.14259/bp.v2i1.86

Authors Column