# Epitope designing on avian influenza disease

Suchitra kumari

Department of Bioinformatics, NTHRYS Biotech Labs,
Hyderabad, India

**Abstract**— Avian influenza is a disease caused by influenza virus. Avian influenza is a very variable and highly contagious virus that is widespread in birds, particularly in wild waterfowl and shorebird. In this present work, we predicted "FKRTNGSSV" epitope by using TMHMM, BCpred, ProPred, Propred I, MHCpred, VexiJen and Pepitope tools.There are total 17 alleles of MHC which binds with designed epitope. This epitope can be used for new vaccine production, discovery and development of diagnostic and therapeutic antibodies after wet lab validation.

**Index Terms**— Avian influenza,Epitope design, BCpred, TMHMM, ProPred, MHCpred, Pepitope.

## 1 INTRODUCTION

Avian influenza is a highly contagious viral disease .Influenza A viruses infects many different animals including ducks, chickens, pigs, whales, horses, and seals. Typically infection and transmission among one animal species sometimes can cross over and cause illness in another species [1]. Avian influenza is also known as bird flu. The Avian influenza A subtype H5N1 is a extremely pathogenic (HPAI) strain of the virus that has been confirmed in poultry populations across Asia, Russia and some southern European countries [1].

.

In this study, we designed epitope for avian influenza treatment. We used several tools in this work along with Epitope mapping. Epitope mapping is used for identification and characterization of the minimum molecular structures that are recognized by the Immune System, mainly T and B cells. Most epitopes recognized by B cells are three-dimensional surface features of a protein antigen; these features fit exactly and thus bind to antibodies [2]. The new vaccine production, discovery and development of diagnostic are the main objective of epitope mapping.

## 2 METHODOLOGY

### 2.1 Target sequence Selection

First, we selected avian influenza protein target sequence for this study. On the basis of non- homologous human protein and essential protein for pathogens, target sequence has selected. The selected target sequence showing essential protein or not was predicted by Database of Essential Genes (DEG) database. According to DEG, the functions encoded by essential genes are considered as the foundation of life and these are required for their growth, reproduction and their survival [3]. Non-homologous protein sequences were checked by using BLASTp.

• *Suchitra kumari, Trainee at Nthrys Biotech Labs Hyderabad, India. Email: suchitra.bioinfo@nthrys.org*

### 2.2 Transmembrane region and allergenicity prediction

After target selection, prediction of transmembrane was done by TMHMM tool [4]. In TMHMM, protein sequence which lies in the transmembrane region shows antigenicity. EVALLER tool [5] is used for allergenicity prediction.

### 2.3 Prediction of B cell epitope by BCpred

BCpred server [6] was used for B cell prediction.

### 2.4 Epitope prediction by ProPred, ProPred I, MHCpred

Propred, Propred I, MHCpred were used after B cell prediction. The ProPred I is an on-line service for identifying the MHC Class I binding regions in antigens. It implements matrices for 47 MHC Class-I alleles, proteasomal and immunoproteasomal models [7]. The aim of ProPred server is to predict MHC Class-II binding regions in an antigen sequence. There are total 51 MHC Class-II alleles present in ProPred server [8]. MHCpred is used to predict the binding affinity of major histocompatibility complex (MHC) class I and II molecules [9].

### 2.5 Antigenicity check by vaxiJen

Antigenicity prediction was performed through VaxiJen server. VaxiJen is based on the physicochemical properties of proteins without recourse to sequence alignment [10].

### 2.6 Epitope mapping for 3D structure

The structure of target sequence retrieved from PDB for epitope mapping study.In this step, Pepitope server was used. Pepitope is mainly based on two algorithms namely pepsurf and mapitope. PepSurf aligns each peptide to a graph which represents the surface of the input 3D structure while mapitope first identifies pairs of residues that are significantly over represented in the panel of peptides, compared to their expected frequencies [11].

## 3 Results

We found that the polymerase PB2 as a target protein sequence. The exomembrane sequence was checked by using TMHMM. The whole target protein i.e. Polymerase PB2 (759 amino acids) was outside the membrane [Fig.1] which favors epitope designing. The target protein shows no allergenicity, confirmed by EVALLER tool. There are total 13 B cell epitopes were predicted by using BCpred. All 13 B cell epitopes shows the good score value. Considered those T cell epitopes as a part of B cell epitopes and that lies in the transmembrane region. There are 55 epitopes nanomer sequence comes after performing ProPred analysis. We found that total 16 epitopes comes under PropredI. Out of 16 epitopes, only one epitope "FKRTNGSSV" is showing antigenicity by vaxiJen [Table.2].

"FKRTNGSSV" is the promising vaccine peptide as it binds to most MHC alleles with good confidence value and shows antigenicity also. "FKRTNGSSV" epitope binds with 3 alleles DRB1-0101, DRB1-0701 and DRB1- 0703 from ProPred analysis. From ProPred I, our designed epitope binds with 14 allele- Hla –A1, HLA-A3, HLA-A*3302, HLA-B40, HLA-B*5301, HLA-B*5401, HLA-B*51, HLA-B*5801, HLA-B61, HLA-B8, MHC-Db. Since, total 17 alleles are able to bind with proposed epitope. From MHCpred, the binding affinity of DRB0101 is less than 500nm i.e. 69.82 nm to the epitope. It shows the high binding affinity with FKRTNGSSV.

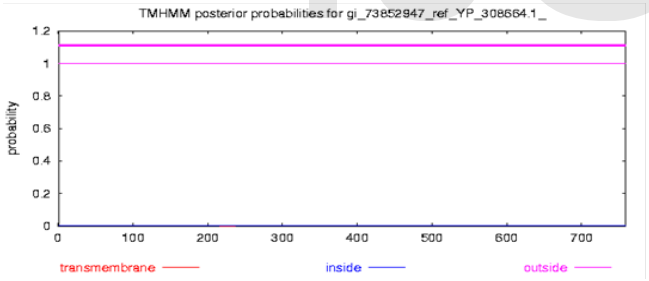


Fig1:TMHMM result

TABLE .1. BCpred result

| EPITOPE | SCORE |
|---|---|
| TSESQLTITKEKKEELQDCK | 0.996 |
| TFKRTNGSSVKKEEEVLTGN | 0.995 |
| TQGTCWEQMYTPGGEVRNDD | 0.989 |
| PERNEQGQTLWSKTNDAGSD | 0.979 |
| VDHMAIIKKYTSGRQEKNPA | 0.975 |
| HGTFGPVHFRNQVKIRRRVD | 0.958 |
| AKEAQDVIMEVVFPNEVGAR | 0.957 |
| WWNRNGPTTSTVHYPKVYKT | 0.95 |
| NGPESVLVNTYQWIIRNWET | 0.942 |
| MRILVRGNSPVFNYNKATKR | 0.941 |
| PFAAAPPEPSRMQFSSLTVN | 0.927 |
| LSPEEVSETQGTEKLTITYS | 0.893 |
| GKDAGALTEDPDEGTAGVES | 0.816 |

TABLE.2. VaxiJen result

| EPITOPE | SCORE | ANTIGEN/NON-ANTIGEN |
|---|---|---|
| LVRGNSPVF | -0.4553 | probable non-antigen |
| RILVRGNSP | -0.0498 | probable non-antigen |
| ILVRGNSPV | -0.2489 | probable non-antigen |
| MRILVRGNS | 0.1716 | probable non-antigen |
| KRTNGSSV | 0.2429 | probable non-antigen |
| **FKRTNGSSV** | **0.5563** | **Probable antigen** |
| TFKRTNGS | 0.2061 | probable non-antigen |
| VDHMAIIKK | 0.3771 | probable non-antigen |
| DHMAIIKKY | 0.2993 | probable non-antigen |
| MAIIKKYTS | -0.3134 | probable non-antigen |
| HMAIIKKYT | 0.2293 | probable non-antigen |
| LVRGNSPVF | -0.4553 | probable non-antigen |
| RILVRGNSP | -0.0498 | probable non-antigen |
| ILVRGNSPV | -0.2489 | probable non-antigen |
| RILVRGNSP | -0.0498 | probable non-antigen |
| MRILVRGNS | 0.1716 | probable non-antigen |

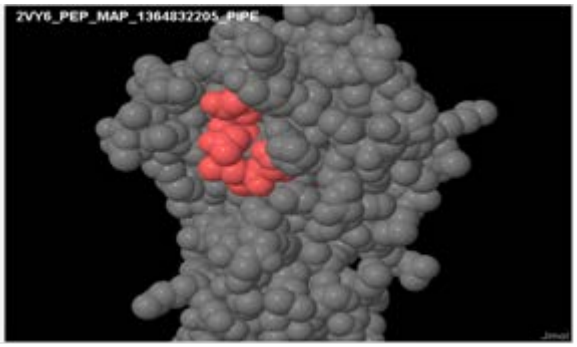


Fig 2: Surfaced exposed epitope "FKRTNGSSV" epitope (red color). The target protein chain is represented as a space-filling model colored in grey.

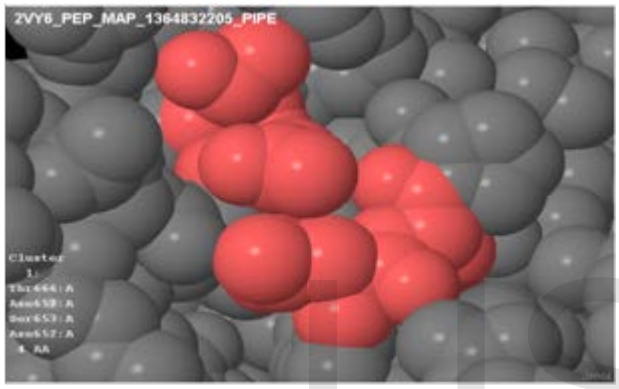


Fig 3: List of residues that the cluster contains and the peptides that are aligned to cluster 1.

It is found that the designed epitope is exposed on the outer surface (red colour) [Fig.2]. PDB ID: 2VY6 was used for epitope mapping study. By using combined algorithms i.e. pepsurf and mapitope, we found four center residues [Fig.3]. Fom Pepsurf algorithm, we got score 7.85 with eight residues: LEU648; ASN652; SER653; ASN657; TYR658; LYS663; ARG664; THR666 from best cluster result. This 7.85 score is very less which indicates the novel epitope.

This research will provide new insight for accelerating immunotechnology for development of vaccines.

## 4 Conclusion

In summary, in this study, we have predicted epitope "FKRTNGSSV". This epitope can be checked for their predicted activity for avian influenza treatment. It can also be used for new vaccine production, discovery and development of diagnostic and therapeutic antibodies after wet lab validation.

## Acknowledgement

Express my sincere gratitude to Mr. Balaji S Rao, Director, Nthrys Biotech Labs, Hyderabad. Special thanks to Krishna Kant Gupta, Ankit Gupta, Narayana Murthy Lalam.

## References

[1] Hse.gov.uk/biosafety/diseases/avianflu.htm

[2] Creative-biostructure.com/others3.htm

[3] Zhang R, Ou HY, Zhang CT. (2004), "DEG: a database of essential genes", Nucleic Acids Res., 1; 32(Database issue):D271-2.

[4] Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305(3): 567–580. doi: 10.1006/jmbi.2000.4315.

[5] Barrio A M, Atmadja D S, Nister A, Gustafsson M G, Hammerling U, et al . (2007) EVALLER: a web server for in silico assessment of potential protein allergenicity; Nucleic Acids Res. 2007 July; 35(Web Server issue): W694–W700;. doi: 10.1093/nar/gkm370

[6] El-Manzalawy Y, Dobbs D, Honavar V (2008) Predicting linear B-cell epitopes using string kernels. J Mol Recognit. 21(4): 243–255.

[7] Singh H, Raghava GP (2003) ProPred1: prediction of promiscuous MHC Class-I binding sites. Bioinformatics 19(8): 1009–1014. doi:10.1093/bioinformatics/btg108.

[8] Singh H, Raghava GP (2001) ProPred: prediction of HLA-DR binding sites. Bioinformatics 17(12): 1236–7. doi: 10.1093/bioinformatics/17.12.1236.

[9] Guan P, Doytchinova IA, Zygouri C, Flower DR (2003) MHCPred: A server for quantitative prediction of peptide-MHC binding, Nucleic Acids Res. 31(13): 3621–3624.

[10] Doytchinova I A and Flower D R. (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8:4.

[11] Mayrose I, Penn O, Erez E, Rubinstein ND, Shlomi T, et al. (2007) Pepitope: epitope mapping from affinity-selected peptides. 3244–3246. doi: 10.1093/bioinformatics/btm493.