STUDY ON EPITOPE MAPPING OF HEMAGGLUTININ PROTEIN IN H5N1 INFLUENZA VIRUS

A Thesis submitted in partial fulfillment of the requirements for the award of the degree of

MASTER OF SCIENCE IN MICROBIOLOGY

Submitted By
Anterpreet Kaur Chahal
Registration No. 301105002

Under the Supervision of
DR. MANOJ BARANWAL

DEPARTMENT OF BIOTECHNOLOGY& ENVIRONMENTAL SCIENCES
THAPAR UNIVERSITY, PATIALA
JULY, 2013

CERTIFICATE

This is to certify that the thesis entitled **"STUDY ON EPITOPE MAPPING OF HEMAGGLUTININ PROTEIN IN H5N1 INFLUENZA VIRUS"** being submitted by Anterpreet Kaur Chahal, Registration No. 301105002 in partial fulfillment of the requirements for the award of degree of Master of Science in Microbiolgy, Department of Biotechnology and Environmental Sciences, Thapar University, Patiala, is a bonafide work carried out under my supervision and guidance. The thesis has not been submitted for award of any other degree or certificate in this or any other university.

Dr. Manoj Baranwal
Supervisor
DBTES
THAPAR UNIVERSITY
Patiala

Dr. M.S Reddy
Head of Department
DBTES
THAPAR UNIVERSITY
Patiala

Dr. S.K Mohapatra
Dean , Academic Affairs
THAPAR UNIVERSITY
Patiala

CANDIDATE'S DECLARATION

I hereby declare that the work which is being presented in the dissertation entitled "**STUDY ON EPITOPE MAPPING OF HEMAGGLUTININ PROTEIN IN H5N1 INFLUENZA VIRUS**" in partial fulfillment of the requirements for the award of Master of Science in Microbiology, Department of Biotechnology and Environmental Sceinces, Thapar University, Patiala is an authentic record of my own work during a period of six months from January 2013 to June 2013, under the supervision of Dr. Manoj Baranwal, Assistant Professor, Department of Biotechnology and Environmental Sceinces, Thapar University, Patiala. The Report has not been submitted for the award of any other degree or certificate in this or any other university.

Anterpreet Kaur Chahal

Place: PATIALA
Date: 15/7/2013

Anterpreet Kaur Chahal
Roll No. 301105002

This is to certify that the above statement given by the candidate is correct and true to the best of our knowledge.

Dr. Manoj Baranwal
Supervisor
DBTES
THAPAR UNIVERSITY
Patiala

Dr. M.S. Reddy
Head Of Department
DBTES
THAPAR UNIVERSITY
Patiala

ACKNOWLEDGEMENT

I take immense pleasure in thanking **Dr. Manoj Baranwal, Assistant Professor,** Department of Biotechnology and Environmental Sciences, Thapar University, Patiala for giving me an opportunity to accomplish my training under his supervision and for his able guidance and useful suggestions, which helped me in completing the project work, in time.

I wish to express my deep sense of gratitude to Dr. **M.S. Reddy, Head of Department,** Department of Biotechnology and Environmental Sciences, Thapar University, Patiala for his immense concern throughout the project work.

Needless to mention **Ms. Neha Lohia, Research scholar** who had been a source of inspiration and for her timely guidance in the conduct of my project work.

I take pleasure to acknowledge the help given by **Avni Vij**, who was always there when I needed her help.

Words fail me to express my thanks to my parents & my brother for their sacrifice encouragement and heart full blessings that continue to enlighten my life.

Above all I thank almighty God for blessing me with strength and wisdom to complete this project successfully.

Date 15/7/2013

Anterpreet Kaur Chahal

Anterpreet Kaur Chahal

ABSTRACT

Influenza viruses circulating in animals pose threats to human health. Humans can be infected with influenza viruses from animal sources, such as pigs and birds. H5N1 also called bird flu is one of recent strain of influenza which causes death in birds as well as in human. Although human to human transmission of this disease is not clearly reported. The primary risk factor for human infection appears to be direct or indirect exposure to infected live or dead animals or contaminated environments. Current influenza virus vaccines protect mostly against one particular strain. Hence great challenge in the field of influenza virus research is to design a universal vaccine. In our study we have considered the conserved peptide of hemagglutinin protein (HA) of H5N1 Virus to predict the epitope. Nine conserved peptide sequences from 163 HA protein sequences were obtained. Thirty and eleven epitopes were predicted for MHC Class I T and B cell respectively based on immuninformatics approach. **Seven and two** immunogenic peptides were generated by considering the overlapping epitopes MHC Class I T and B cell respectively**.** These immunogenic peptides can be validated for their immunogenic response in the *in-vitro system* which can be an interesting candidate for vaccine design against the H5N1 influenza virus vaccine.

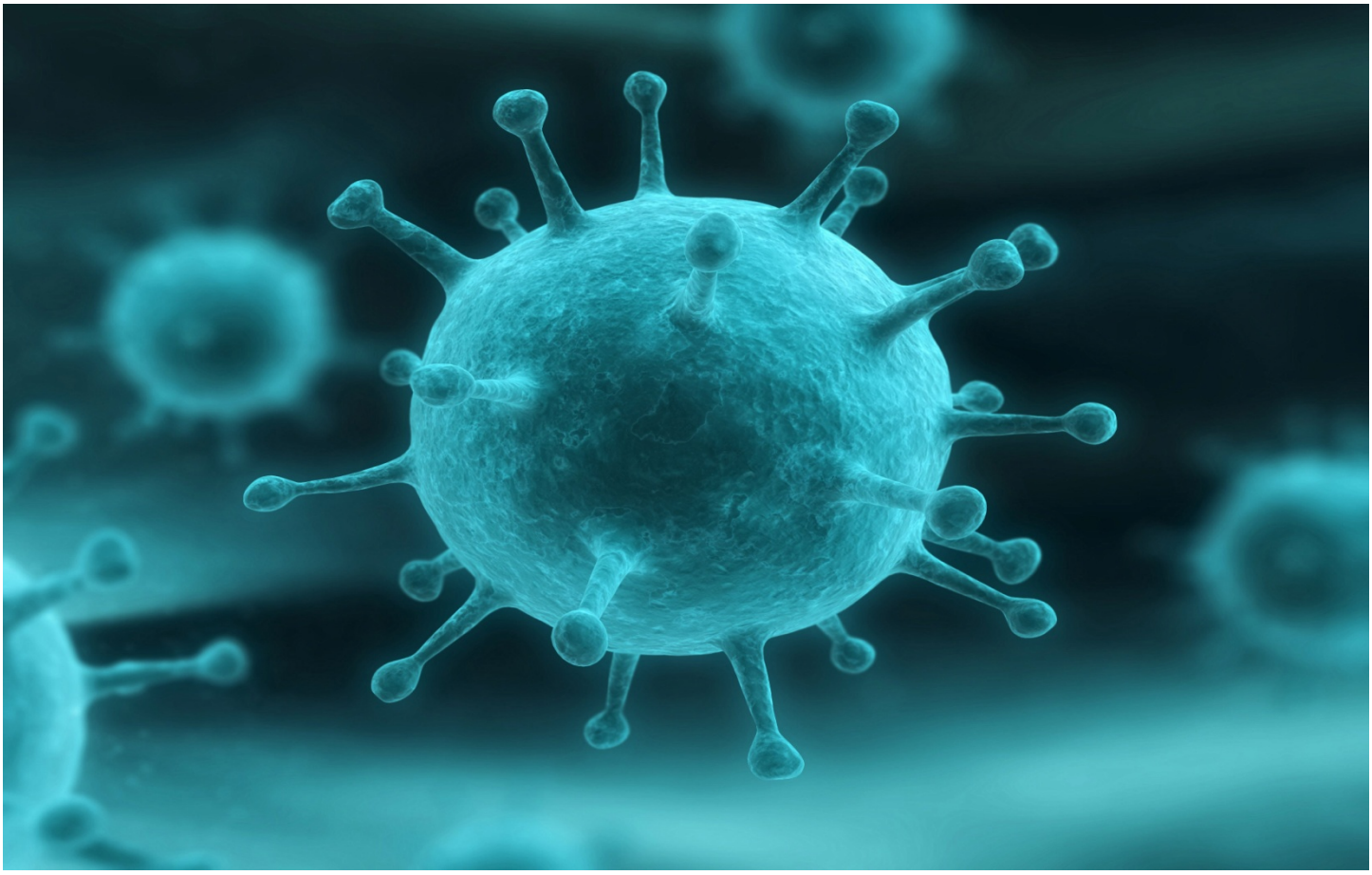
ABBREVIATIONS

RNA	-	Ribonucleic acid
WHO	-	World Health Organization
HLA	-	Human Leukocyte Antigen
MHC	-	Major Histocompatibility Complex
HA	-	Hemagglutinin
NA	-	Neuraminidase
M1	-	Matrix protein 1
RNP	-	Ribonucleoprotein
NEP	-	Nuclear export protein
NS1	-	Non- Structural protein
IFN	-	Interferon
PA, PB	-	Polymerase protein A, B
FDA	-	Food and Drug Administration
APC	-	Antigen presenting cells
PBMC	-	Peripheral Blood Mononuclear Cells
MUSCLE	-	Multiple Sequence Comparison By Log Expectation
AVANA	-	Antigen Variability Analyzer
NCBI	-	National Council of Biological Information

MSA	-	Multiple Sequence Alignment
IEDB	-	Immune Epitope Database
CTL	-	Cytotoxic T- Lymphocytes
BIMAS	-	Bioinformatics and Molecular Analysis Section
DMEM	-	Dulbecco's Modified Eagle Medium
PBS	-	Phosphate Buffer Saline
EDTA	-	Ethylenediaminetetraacetic acid

TABLE OF CONTENTS

INTRODUCTION



Influenza, commonly known as "the flu", is an infectious disease of birds and mammals and is the major cause of sickness and death around the world. Influenza Virus is negative-sense single-stranded RNA virus which belongs to *Orthomyxivirdae* family (Group 5). The several subtypes of influenza are labeled according to an H number (Hemagglutinin subtype) and an N number (neuraminidase subtype). There are 16 different HA subtypes (H1 to H16) and 9 different NA subtypes (N1 to N9).

Influenza A virus subtype H5N1, also known as "bird flu", A(H5N1) or simply H5N1, is a subtype of the influenza A virus which can cause illness in humans and many other animal species. It is enzootic in many bird populations, especially in Southeast Asia. One of the recent strains of influenza, the highly pathogenic avian influenza (HPAI) H5N1, and its variants have been in circulation which leads to 141 human deaths. It is epizootic (an epidemic in nonhumans) and panzootic (affecting animals of many species, especially over a wide area), killing tens of millions of birds and spurring the culling of hundreds of millions of others to stem its spread.

Antigenic variation is the evolutionary mechanism by which viruses evade host immune system. Influenza viruses accomplish this through one of the two mechanisms: (1) Antigenic shift (genetic reassortment between a human and a non-human virus in a non-human host), and (2) Antigenic drift (accumulation of mutations that facilitate evasion of the host-immune response).

Antiviral medicines can be used to treat or prevent influenza. When used as a treatment, the medicine does not eliminate flu symptoms, although it can reduce the severity and duration of symptoms by about one day. Antiviral medicines that are used to treat or prevent the flu include oseltamivir and zanamivir (Neuraminidase inhibitors). Two other antiviral medicines, rimantadine and amantadine (M2 ion channel blockers), were used in the past but are generally no longer effective because most flu viruses are now resistant to them. Antiviral treatment is most effective for seasonal influenza when it is taken within the first 48 hours of flu symptoms. Antibiotics are not useful for treating viral illnesses such as influenza. They are only used if there is a bacterial complication of the flu such as bacterial pneumonia, ear infection, or sinusitis.

Vaccine is considered to be one of the most effective way to control influenza virus . Currently available vaccines used against influenza induce immunity against a specific and closely related antigenic viral strain. Vaccines have been formulated against several of the avian H5N1 influenza varieties. Vaccination of poultry against the ongoing H5N1 epizootic is widespread in certain countries. Some vaccines also exist

for use in humans, and others are in testing, but none have been made available to civilian populations, nor produced in quantities sufficient to protect more than a tiny fraction of the Earth's population in the event of an H5N1 pandemic.

Although the antiviral vaccines are available, which are effective against most recipients, Influenza remains a serious respiratory disease. Therefore, there is a call for the development of a vaccine, which would be protective against different virus strains.

Peptide based vaccines can be designed to represent subdominant epitopes, thus, elicit immunity. The general idea behind the peptide vaccines is based on the chemical approach to synthesize the identified peptide containing B-cell and T-cell epitopes that are immunodominant and can induce specific immune responses. The first epitope-based vaccine was created in 1985 by Jackob et al. They introduced recombinant DNA and express epitopes against cholera in Escherichia coli.

Since the experimental methods to detect epitopes are expensive and time consuming so various computational tools can be employed to facilitate the process of epitope detection by reducing this experimental effort. Therefore, the new developments in Immunoinformatics field may reduce the time and number of wet laboratory experiments to identify the peptide for vaccine target. Immunoinformatics research stresses mostly on the design and study of algorithms for mapping potential B- and T-cell epitopes, which speeds up the time and lowers the cost needed for laboratory analysis of pathogen gene products.

In this study, our approach is to find the immunogenic peptides containing T and B epitopes from conserved peptide sequences of Hemagglutinin (HA) of H5N1 based on immuninformatics approach.

REVIEW OF LITERATURE

2.1 INFLUENZA

Influenza is a viral infection that is caused by segmented negative sense single stranded RNA viruses of the family Orthomyxoviridae that affects mainly the nose , throat, bronchi and occasionally lungs. Influenza may produce nausea and vomiting, particularly in children, but these symptoms are more common in the unrelated gastroenteritis, which is sometimes inaccurately referred to as "stomach flu" Influenza is transmitted through the air by coughs or sneezes, creating aerosols containing the virus. Influenza can also be transmitted by direct contact with bird droppings or nasal secretions, or through contact with contaminated surfaces. Flu can occasionally lead to pneumonia, either direct viral pneumonia or secondary bacterial pneumonia, even for persons who are usually very health. Influenza spreads around the world in seasonal epidemics, resulting in about three to five million yearly cases of severe illness and about 250,000 to 500,000 yearly deaths. New influenza strains appear when an existing flu virus spreads to humans from another animal species. An avian strain named H5N1 raised the concern of a new influenza pandemic after it emerged in Asia in the 1990s, but it has not evolved to a form that spreads easily between people.

2.1.1 TAXONOMY OF INFLUENZA VIRUS

Influenza virus is a part of Group **V** of RNA virus that includes five genera: Influenza virus A, Influenza virus B, Influenza virus C, Isa virus and Thogoto virus. Influenza viruses make up three of the five genera of the family Orthomyxoviridae.

Influenza A viruses are further classified, based on the viral surface proteins hemagglutinin (HA or H) and neuraminidase (NA or N). Sixteen H subtypes (or serotypes) and nine N subtypes of influenza A virus have been identified. The type A viruses are the most virulent human pathogens among the three influenza types and cause the most severe disease.

TABLE 2.1 Characteristics of Influenza Virus

Features	Influenza Type A	Influenza Type B	Influenza Type C
Membrane Protein	Hemagglutinin (HA) Neuraminidase (NA) Matrix Protein 1 (M1)	Hemagglutinin (HA) Neuraminidase (NA) Matrix proteins	Hemagglutinin- Estrase Fusion Protein (HEF) Matrix protein 2 (M2)
RNA segment	8	8	7
Epidemiology	Pandemic	Epidemic	Epidemic
Antigenic Variation	Antigenic shift and Antigenic drift	Antigenic drift	-
Host	Human, bird, pig, horse, seals	Human, seal	Human, pig

2.1.2 STRUCTURE OF INFLUENZA VIRUS

Influenza viruses are roughly spherical particles, ranging from 80 to 120 nm in diameter. A characteristic feature of influenza virus particles is their external layer of approximately 500 spike-like projections. These spikes represent the envelope glycoproteins HA (which has a rod-like shape) and NA (which is mushroom-shaped). The HA spike is a trimer, consisting of three individual HA monomers, while the NA spike is a tetramer. These glycoproteins are wrapped around a central core. The central core contains the viral RNA genome and viral proteins that protect this RNA.

Underlying the envelope, there' is a layer of matrix protein M1) which provides structural integrity to the viral particle. Each genome segment is coated with multiple copies of nucleoprotein (NP) as well as with one cop y of each of the three polymerase genes, PA, PB1 and PB2, together these four components make up the ribonucleoprotein complex (RNP). The RNPs interact with M1 protein early in infection and nuclear export protein (NEP) to facilitate packaging late in infection (Fig 2.1).

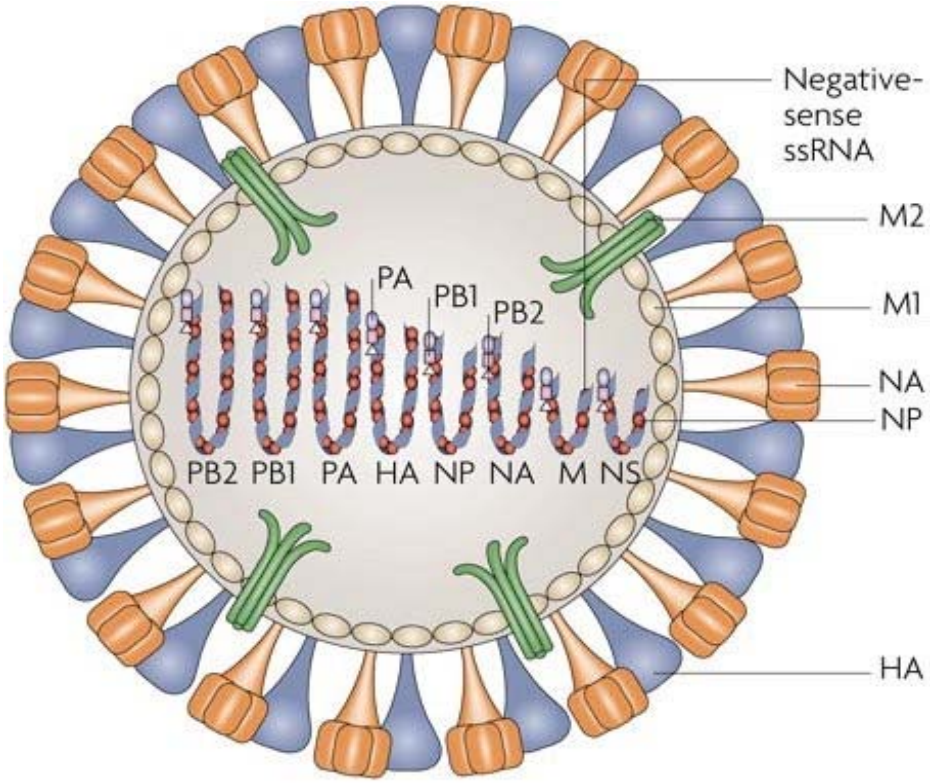


Fig. 2.1 The structure of typical Influenza Virus

The major envelope glycoprotein HA is synthesized in the infected cell as a single polypeptide chain (HA0) with a length of approximately 560 amino acid residues, which is subsequently cleaved into two subunits, HA1 and HA2. These subunits remain covalently linked to one another through disulphide bonds. Cleavage of HA0 is essential for the molecule to be able to mediate membrane fusion between the viral envelope and the host cell membrane. HA belongs to the first proteins for which the entire three-dimensional structure has been elucidated.

The second envelope glycoprotein NA has enzymatic activity, cleaving sialic acid residues from glycoproteins or glycolipids. Since sialic acid functions as a receptor for attachment of influenza virions, the neuraminidase activity of NA, cleaving such receptors, mediates the release of newly formed virus particles from the surface of infected cells. NA is the target for the antiviral drugs oseltamivir and zanamivir. These drugs are sialic acid analogues, which inhibit the enzymatic activity of NA, thus slowing down the release of progeny virus from infected cells.

2.1.3 FUNCTIONS OF VARIOUS INFLUENZA PROTEINS

HA	**Hemagglutinin** envelope protein capable of being recognized by the immune system
NA	**Neuraminidase** binds to plasma membrane
NP	**Nucleoprotein** forms a complex with the viral RNA genome and packages the RNA into a helical ribonucleoprotein core
M1	**Matrix protein 1** surrounds the ribonucleoprotein core
M2	**Matrix protein 2** forms an ion channel during lysosome fusion
PA, PB1,PB2	**Polymerase proteins** are components of viral RNA transcriptase
NS1,NS2	**Non structural proteins** that are involved in the control of nucleo cytoplasmic transport

2.1.4 LIFE CYCLE OF INFLUENZA VIRUS

The influenza virus life cycle can be divided into the 5 stages :

1. Entry of virus particle into the host cell.

2. Entry of viarl ribonucleoprotein into the nucleus.

3. Transcription and replication of viral genome.

4.Export of vRNPs from the nucleus

5.Assembly and budding at the host cell plasma membrane.

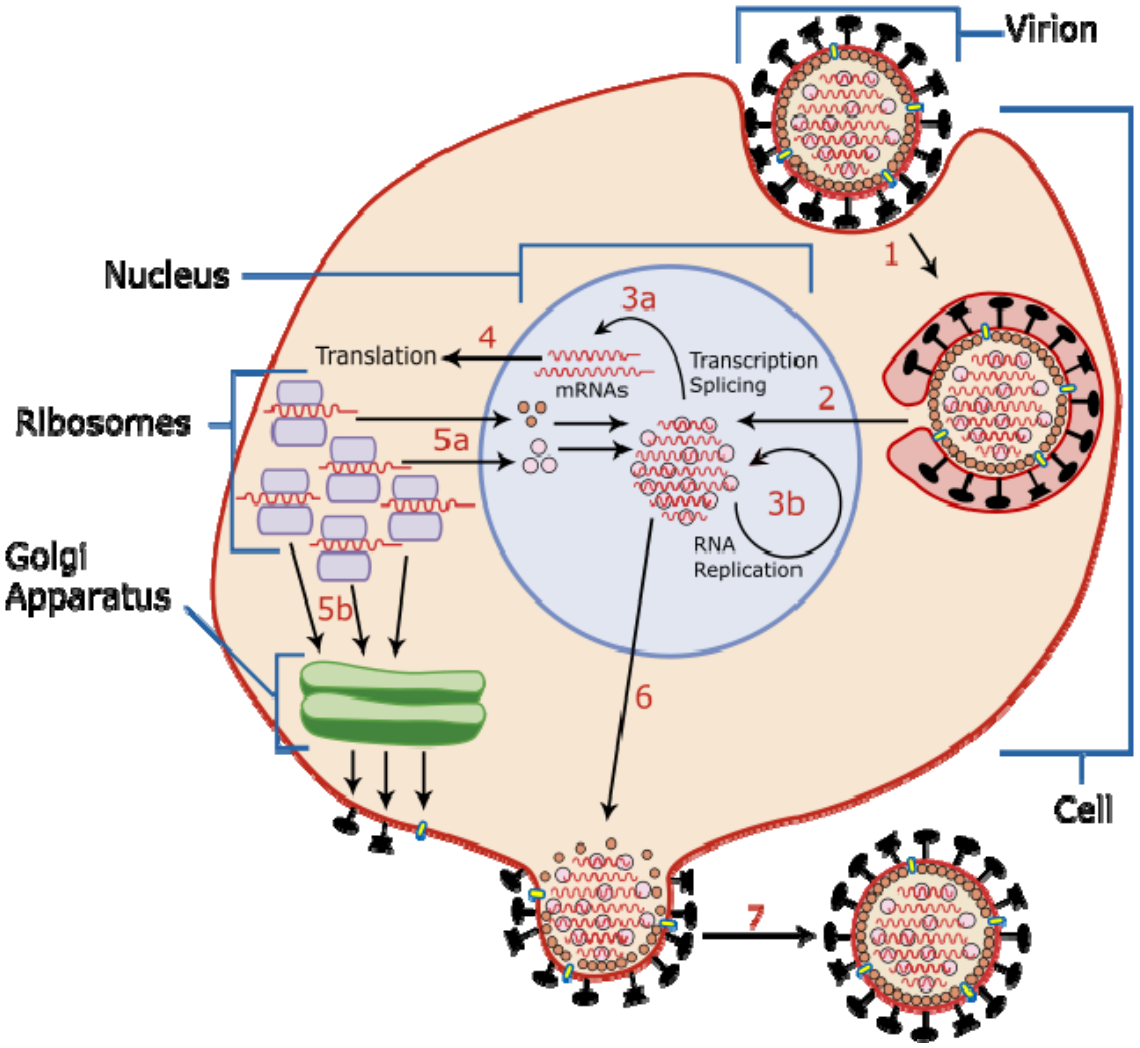


Fig. 2.2 Life Cycle of Influenza Virus

Viruses can replicate only in living cells. Influenza infection and replication is a multi-step process : (Stage 1 to stage 7 ; Fig.2.2)

Stage 1: Influenza viruses bind through hemagglutinin onto sialic acid sugars on the surfaces of epithelial cells, typically in the nose, throat, and lungs of mammals, and intestines of birds . After the hemagglutinin is cleaved by a protease, the cell imports the virus by endocytosis. The acidic conditions in the endosome cause hemagglutinin protein to fuse the viral envelope with the vacuole's membrane, then the M2 ion channel allows protons to move through the viral envelope and acidify the core of the virus, which causes the core to dissemble and release the viral RNA and core proteins.

Stage 2: The viral RNA (vRNA) molecules, accessory proteins and RNA-dependent RNA polymerase are then released into the cytoplasm. The viral proteins that make up the vRNP are NP, PA, PB1, and PB2. All of these proteins have known nuclear localization signals (NLSs) that can bind to the cellular nuclear import machinery and, thus, enter the nucleus.

Stage 3a and 3b: These core proteins and vRNA form a complex that is transported into the cell nucleus, where the RNA-dependent RNA polymerase begins transcribing complementary positive-sense vRNA.

Stage 4: The vRNA either is exported into the cytoplasm and translated or remains in the nucleus.

Stage 5a and 5b : In case of neuraminidase and hemagglutinin, newly synthesized viral proteins are either secreted through the Golgi apparatus onto the cell surface or transported back into the nucleus to bind vRNA and form new viral genome particles . Other viral proteins have multiple actions in the host cell, including degrading cellular mRNA and using the released nucleotides for vRNA synthesis and also inhibiting translation of host-cell mRNAs.

Stage 6: Negative-sense vRNAs that form the genomes of future viruses, RNA-dependent RNA polymerase, and other viral proteins are assembled into a virion. Hemagglutinin and neuraminidase molecules cluster into a bulge in the cell membrane. The vRNA and viral core proteins leave the nucleus and enter this membrane protrusion.

Stage 7: The mature virus buds off from the cell in a sphere of host phospholipid membrane, acquiring hemagglutinin and neuraminidase with this membrane coat. As before, the viruses adhere to the cell through hemagglutinin; the mature viruses detach once their neuraminidase has cleaved sialic acid residues from the host cell. After the release of new influenza viruses, the host cell dies. ((http:// www. proteopedia. Org /wiki /index.php/influenza_hemagglutinin#structure)

2.2 HEMAGGLUTININ PROTEIN

HA is a homotrimeric integral membrane glycoprotein. It is shaped like a cylinder, and is approximately 13.5 nanometres long. The three identical monomers that constitute HA are constructed into a central α helix coil; three spherical heads contain the sialic acid binding sites. HA monomers are synthesized as precursors that are then glycosylated and cleaved into two smaller polypeptides: the HA1 and HA2 subunits. Each HA monomer consists of a long, helical chain anchored in the membrane by HA2 and topped by a large HA1 globule**.**

Since hemagglutinin is the major surface protein of the influenza A virus and is essential to the entry process, it is the primary target of neutralizing antibodies. Neutralizing antibodies against flu have been found to act by two different mechanisms, mirroring the dual functions of hemagglutinin:

1. Inhibition of attachment to target cells
2. Inhibition of membrane fusion (entry)

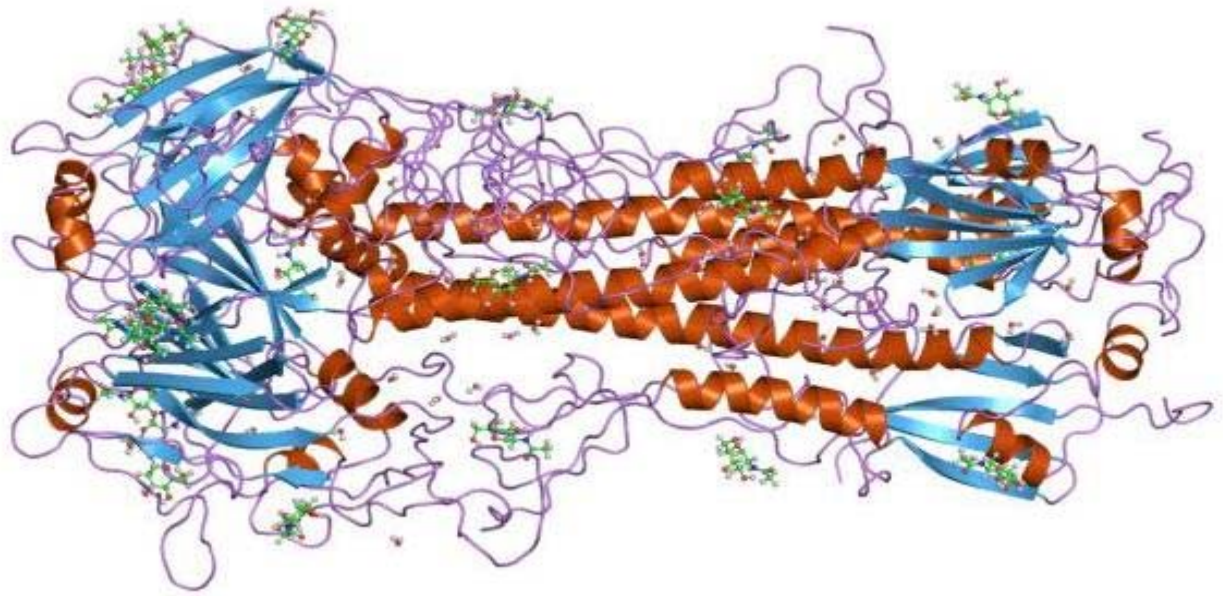


Fig. 2.3 Structure of Hemagglutinin protein

HA has two functions. Firstly, it allows the recognition of target vertebrate cells, accomplished through the binding to these cells' sialic acid-containing receptors. Secondly, once bound it facilitates the entry of the viral genome into the target cells by causing the fusion of host endosomal membrane with the viral membrane. (Varghese et al; 1983)

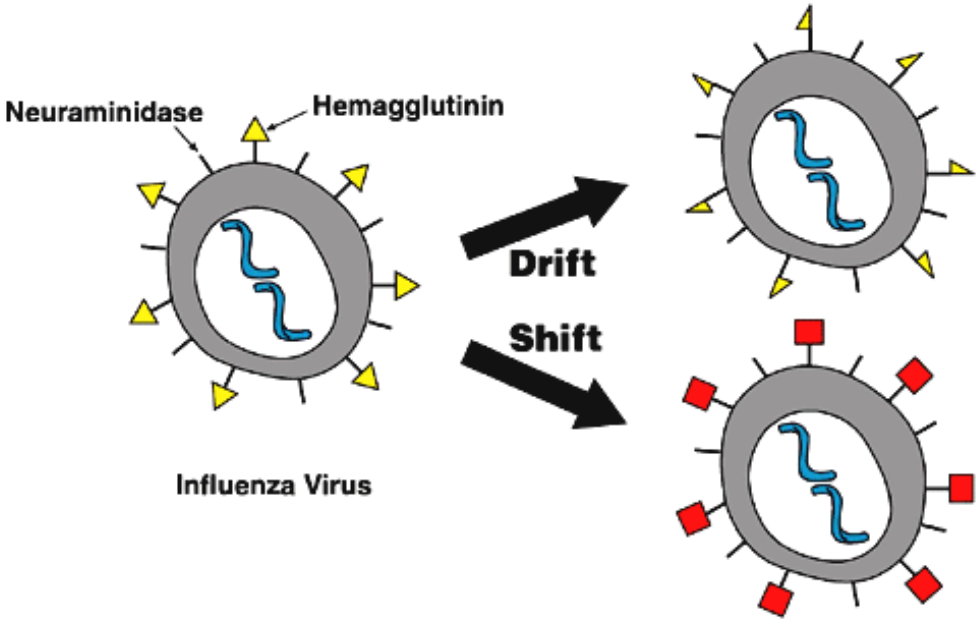
HA binds to the monosaccharide sialic acid which is present on the surface of its target cells, which causes the viral particles to stick to the cell's surface. The cell membrane then engulfs the virus and the portion of the membrane that encloses it pinches off to form a new membrane-bound compartment within the cell called an endosome, which contains the engulfed virus. The cell then attempts to begin digesting the contents of the endosome by acidifying its interior and transforming it into a lysosome. However, as soon as the pH within the endosome drops to about 6.0, the original folded structure of the HA molecule becomes unstable, causing it to partially unfold and release a very hydrophobic portion of its peptide chain that was previously hidden within the protein. This so-called "fusion peptide" acts like a molecular grappling hook by inserting itself into the endosomal membrane and locking on. When the rest of the HA molecule refolds into a new structure (which is more stable at the lower pH), it "retracts the grappling hook" and pulls the endosomal membrane right up next to the virus particle's own membrane, causing the two to fuse together. Once this has happened, the contents of the virus, including its RNA genome, are free to pour out into the cell's cytoplasm.

2.3 ANTIGENIC SHIFT AND ANTIGENIC DRIFT

Influenza virus has a remarkable ability in escaping host defense mechanisms by altering its antigenic character. Continuous and extensive antigenic variation shown by Hemagglutinin (HA) and Neuraminidase (NA) is the main reason of tremendous variability of influenza virus. (Young *et al*; 1980)

Antigenic drift: It is constantly occurring in both types A and B viruses. The HA and/or NA of the new strain are sufficiently different to evade the pre-existing human immunity. This leads to the seasonal epidemics. This subtle process involves point mutations within antibody-binding sites in the HA protein, the NA protein, or both, which potentially occur each time the virus replicates.

Antigenic shift: It occurs only in type A virus. It describes the emergence of an entirely new virus sub type. When this new sub type emerges, it causes a pandemic because there is no pre-existing immunity in humans. The novel assortment of HA or NA proteins in a shifted virus creates a new influenza A subtype. Because people have little or no immunity to such a new subtype, their appearance tends to coincide with severe flu epidemics or pandemics.



2.4 H5N1

H5N1 is the world's major influenza pandemic threat.Influenza A virus subtype H5N1, also known as "bird flu", A(H5N1) or simply H5N1, is a subtype of the Influenza A virus which can cause illness in humans and many other animal species. A bird-adapted strain of H5N1, called HPAI A(H5N1) for "highly pathogenic avian influenza virus of type A of subtype H5N1", is the causative agent of H5N1 flu, commonly known as "avian influenza" or "bird flu". It is enzootic in many bird populations, especially in Southeast Asia. One strain of HPAI A(H5N1) is spreading globally after first appearing in Asia. It is epizootic (an epidemic in nonhumans) and panzootic, killing tens of millions of birds and spurring the culling of hundreds of millions of others to stem its spread.

Due to the high lethality and virulence of HPAI A(H5N1), its endemic presence and increasingly large host reservoir, the H5N1 virus is the world's largest current pandemic threat. At least 12 companies and 17 governments are developing pre-pandemic influenza vaccines in 28 different clinical trials that, if successful, could turn a deadly pandemic infection into a nondeadly one.

Low pathogenic avian influenza H5N1 (LPAI H5N1) also called "North American" H5N1 commonly occurs in wild birds. In most cases, it causes minor sickness or no noticeable signs of disease in birds. It is not known to affect humans at all. The only concern about it is that it is possible to be transmitted to poultry where it mutates into a highly pathogenic strain.

There is not effective treatment for H5N1 flu, but oseltamivir (commercially marketed by Roche as Tamiflu), can sometimes inhibit the influenza virus from spreading inside the user's body. Animal and lab studies suggest that Relenza (zanamivir), which is in the same class of drugs as Tamiflu, may also be effective against H5N1.

There are several H5N1 vaccines for several of the avian H5N1 varieties, but the continual mutation of H5N1 renders them of limited use to date. These vaccines can sometimes provide cross-protection against related flu strains, the best protection would be from a vaccine specifically produced for any future pandemic flu virus strain.

2.5 TREATMENT FOR INFLUENZA VIRUS

2.5.1 Anti-viral drug therapy

Neuraminidase Inhibitors

Neuraminidase cleaves the neuraminic acid component of sialic acid residue to which hemagglutinin is bound .This cleavage is necessary for the virus to get released from the host cell. The neuraminidae drug inhibitors are Zanamivir and Oseltamivir. (McNicholl *et al* ; 2001)

1) Zanamivir:- Zanamivir, an orally inhaled powder , is competitive inhibitor of the neuraminidase glycoprotein, that is essential in the infective cycle of Influenza viruses. It closely mimics sialic acid that is the natural substrate of neuraminidase. It is currently approved in 19 countries for the treatment of , and in two countries for Prophylaxis of Influenza A and B.

2) **Oseltamivir:-** Oseltamivir is a potent inhibitor of influenza neuraminidase. Oseltamivir carbox ylate acts by selective inhibition of influenza A and B viral neuraminidase. A lipophilic side chain of the active drug binds to the virus enzyme, blocking its ability to cleave sialic acid residues on the surface of the infected cell and resulting in an inability to release progeny virions. It si approved by FDA for use as a treatment for Influenza A and B in persons 18 years or older.

Ion Channel Blockers

These drugs inhibit the function of M2 protein and thus stop the replication process. These drugs are effective only against Influenza A, and not Influenza B, because Influenza B does not have an M2 protein , but a substitute protein called NB protein that is not effected by these drugs.Ion channel blocker drugs are Amantadine and Rimantadine.

1) Amantadine:- Amantadine inhibits the replication of influenza A viruses by interfering with the uncoating of the virus inside the cell. It is an M2 inhibitor which blocks the ion channel formed by the M2 protein that spans the viral membrane (Sugrue et.al., 1991). Amantidine is only effective against influenza A, and some naturally occurring strains of influenza , The influenza virus enters the host cell by receptor-mediated endocytosis. Thereafter, acidification of endocytotic vesicles is required for the dissociation of the M1 protein from the ribonucleoprotein complexes.

2) Rimantadine:- Rimantadine is an M2 channel inhibitor which specifically inhibits the replication of influenza A viruses by interfering with the uncoating process of the virus. M2 inhibitors block the ion channel formed by the M2 protein that spans the viral membrane (Sugrue et.al., 1991). It is approved by the FDA for the treatment and prophylax is of influenza A infection in children of one year or older. (Stephenson et al ; 2001 , Jefferson T et al; 2004)

2.5.2 Vaccines

A vaccine is a biological preparation that improves immunity to a particular disease. A vaccine typically contains an agent that resembles a disease-causing microorganism, and is often made from weakened or killed forms of the microbe, its toxins or one of its surface proteins.

Vaccination is an acquisition of protective immunity in advance by admin istration with viral antigenic glycoproteins, such as HA and NA. Current available influenza vaccine contain antigens from two A subtypes, H3N2 and H1N1 and one type of B virus. There are different types of vaccines such as live attenuated vaccines, inactivated vaccines, subunit vaccines, peptide vaccines, DNA vaccines etc.

1) Epitope Based Vaccines

The resistance in vaccine during the past decade has led to several new approaches of vaccine development. One of these approaches is synthetic peptides for eliciting protective immunity against infection. Synthesizing peptides for use as vaccines requires identification of those epitopes in the protein antigen that stimulate protective immunity. Both B and T cell epitopes must be included in the peptide so that arms of the immune system humoral and cell-mediated are stimulated.

In these vaccines, immune response induced by an immunogen is only directed against specific region (sequence) of protein called Epitopes. An epitope, also known as antigenic determinant, is the part of an antigen that is recognized by the immune system, specifically by antibodies, B cells, or T cells. Epitopes may be continuous (linear) or discontinuous (conformational/assembled). Epitope based vaccines can be long-range and broad-spectrum vaccines. Antibody mediated immunity in epitope based vaccine is currently not feasible because the structural determinants of B-cell immunity are highly complex and there is no effective means for predicting the antibody epitope structure.

2.6 IMMUNOINFORMATICS

Immunoinformatics or computational immunology is a field of science that provides fundamental methodologies in the study of immunomics , that is immune related genomics and proteomics.

The integration of immunoinformatics with system biology approaches may lead to a better understanding of immune related diseases at various system levels. Immunoinformatics is an emerging specialization of bioinformatics that focuses upon structure, function and interactions of the molecules involved in immunity (Evans M.C.,2008).

One of the key goals of immunoinformatics is development of computer aided vaccine design. Vaccine informatics is an emerging research area in Immunoinformaticsthat focuses on development and applications of bioinformatics methods that can be used to facilitate every aspect of vaccine development.

Many immunoinformatics algorithms and resources have been developed to predict T cell and B cell immune epitopes for epitope vaccine development and protective immunity analysis (Yongqun He *et al*., 2010). These method may reduce the time and number of wet laboratory experiments to identify the peptide for vaccine targets. These computational methods reduce the time and effort involved in screening potential epitopes.

Its main aim is to convert immunological data into computational problems, solve these problems using mathematical and computational approaches and then convert these results into immunologically meaningful interpretations.

2.6.2 T- Cell Epitope Prediction

The major histocompatibility complex plays a pivotal role in regulating immune response. During T-cell activation peptides derived from protein antigens are presented by MHC molecules. Only a small fraction of the derived peptides are involved in eliciting host immune response (Viret C. and Janeway C.A., 1999). The short antigenic peptides, derived from parent molecules have been proposed as interesting targets of vaccine design (Ishioka G.Y. *et al.,* 1999).

The primary objective of these prediction methodologies is the calculation of MHC peptide binding because high affinity binding often correlates with immunogenicity. A crucial step towards the rational design of synthetic peptide vaccines is the identification of T-cell epitopes from disease causing antigen proteins.

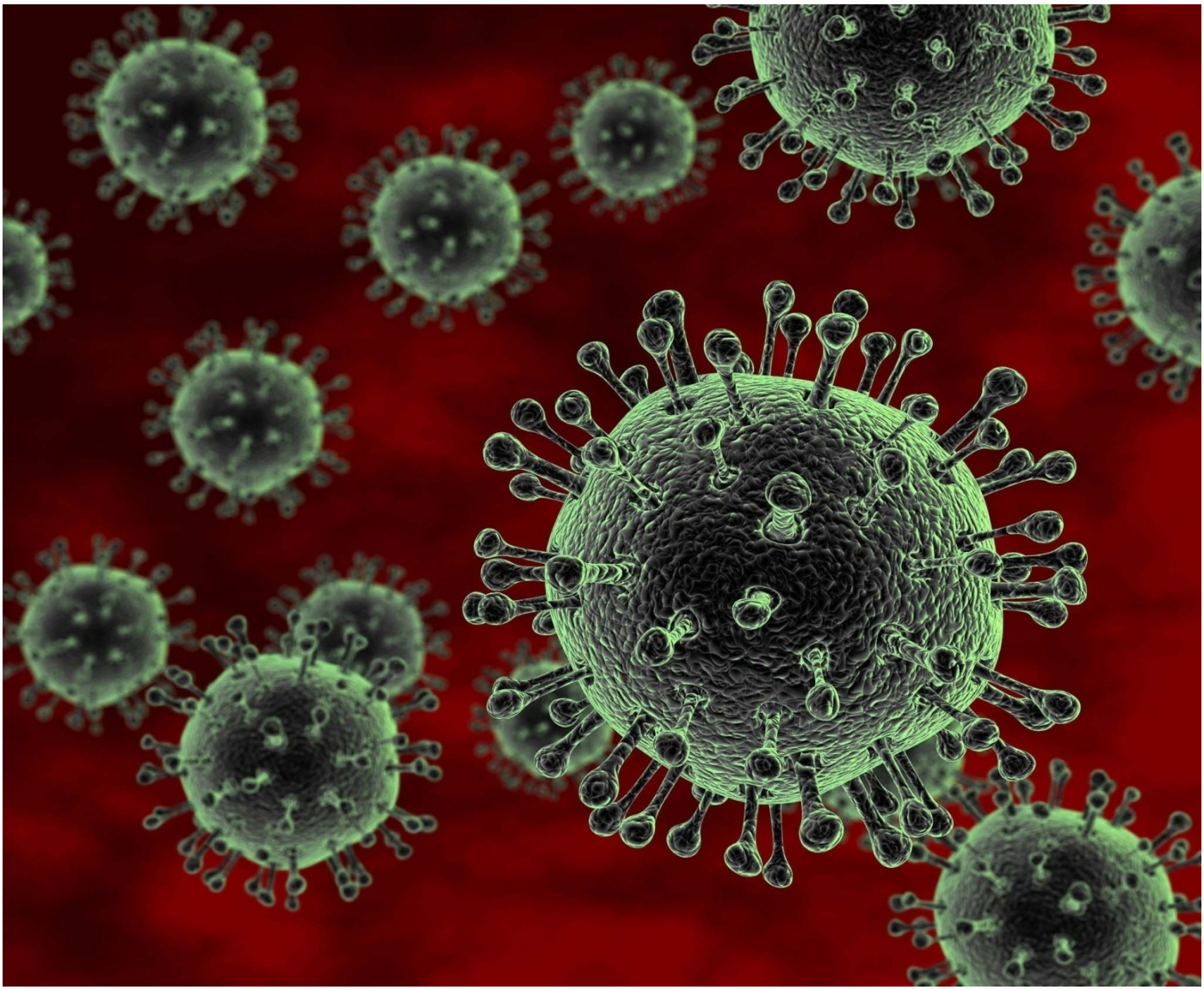
Sequence based methods rely on the primary sequence of peptides that are known to bind specific MHC allele using binding assays. The information on sequence anchors that are deterministic in binding are encoded into a binding motif , a position-dependent matrix , Hidden Markov Model (HMM), Support Vector Machine (SVM), stepwise discriminant analysis, or an Artificial Neural Network (ANN). These models are developed using peptide data in large databases derived from naturally bound peptides.

2.6.3 B-Cell Epitope Prediction

Identification of epitopes that invoke strong responses from B-cells is one of the key steps in designing effective vaccines against pathogens. Because experimental determination of epitopes is expensive in terms of cost, time, and effort involved, there is an urgent need for computational methods for reliable identification of B-cell epitopes. Although several computational tools for predicting B-cell epitopes have become available in recent years, the predictive performance of existing tools remains far from ideal. Antigen-antibody interactions play a pivotal role in the humoral immune response. Antibodies bind to antigens at specific sites which correspond to the antigenic determinants or B-cell epitopes. Identification and characteristics of B-cell epitopes in target antigens is one of the key steps in epitope-driven vaccine design, immunodiagnostic tests, and antibody production.

Add one section on recent literature report of epitope based vaccine design on influenza virus or others with literature

OBJECTIVE



Current influenza virus vaccines are seasonal; thus regular immunization with updated vaccine formulations is necessary to guard against the virus hallmark remodeling of regions that mediate neutralization. Development of a broadly protective influenza vaccine would mark a significant advance in human infectious diseases research (Wang et al., 2010).

The main objective of the present study carried out was to predict a conserved immunogenic peptide containing T and B cell epitopes of Hemagglutinin protein of H5N1 Influenza virus based on immunoinformatics approach.

Work plan included the following objectives:

1) Finding out conserved peptide regions of Hemagglutinin, from all the available strains of H5N1 that has been sequenced.

2) Prediction of immunogenic peptides containing overlapping epitopes in conserved peptide sequences , using different immunoinformatics tools.

3) Optimization of MTT assay (Cell proliferation assay) to estimate cell proliferation of Peripheral Blood Mononuclear cells (PBMC).

MATERIAL AND METHODS

4.1 SEQUENCE RETRIEVAL

The sequences of Hemagglutinin protein of H5N1 were retrieved from NCBI Influenza databse (http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi) from January 1918 to December 2012. Full length sequences were taken and the identical sequences were collapsed using the option available in the database search engine. The sequences were downloaded in fasta format and open ed with WordPad and then these sequences were transferred to a word file.

Some sequences have sometimes an invalid letter code "J " which does not represent any amino acid, and thus these sequences need to be corrected by replacing the "J" with "X".

4.2 CONSERVANCY ANALYSIS

Two bioinformatics tools were used to find out the conserved regions from all retrieved sequences of HA protein of H5N1 influenza virus, MUSCLE to align the sequences and AVANA to find out the conserved regions in the aligned sequences.

4.2.1 MUSCLE (http://www.ebi.ac.uk/Tools/muscle/index.html)

MUSCLE stands for **M**ultiple **S**equence **C**omparison by **L**og-**E**xpectation. It is one of the multiple sequence alignment tool provided by European Bioinformatics Institute (Edgar R.C., 2004).

MUSCLE is a computer program for creating multiple alignments of protein sequences. It is claimed to achieve both better average accuracy and better speed than Clustal W2 or T-Coffee. Elements of the algorithm include fast distance estimation, progressive alignment and refinement using tree-dependent restricted partitioning. MUSCLE uses two distance measures for a pair of sequences: a Kmer distance (for an unaligned pair) and the kimura distance (for an aligned pair).

The limitation of MUSCLE is that it can only align 500 sequences of more than 350 amino acid length at a time. As the number of sequences obtained for HA protein was more than that, so it poses a

limitation to the use of this tool. But this limitation can be solved by grouping the sequences according to their year of isolation .

Multiple alignments of protein sequences are important in many applications, including phylogenetic tree estimation, secondary structure prediction and critical residue identification. Many multiple sequence alignment (MSA) algorithms have been proposed.

4.2.2 AVANA

Antigen Variability Analyzer tool (AVANA) was subsequently used to extract alignments of several subsets of the collected sequences, based on annotation values, such as viral subtype host, and year of isolation. This tool finds conserved regions based on information entropy analysis (Khan AM *et al*., 2008). It also compares alignments using mutual information identifying the mutations that characterize specific sequence sets.

Information entropy is a measure of the uncertainity associated with a random variable or in case of protein mutations occurring in the protein sequence. AVANA uses Shannon Entropy Analysis in account to measure the informational entropy (Shannon C.E., 1948).

Applying Shannon's formula, the nonamer peptide entropy H(x) at any given point x in the alignment is computed by

$$H(x) = -\sum_{i=1}^{n(x)} p(i, x) \log_2 p(i, x)$$

where p(I, x) is the probability of a particular nonamer peptide i being centered at position x. The entropy value increases with n(x), the total number of peptides observed at position x; it is also sensitive to the relative frequency of the peptides; such that it decreases when one peptide is clearly dominant (i.e. the position is conserved). Sites which are highly conserved have lower entropy because entropy is degree of randomness.

The threshold for conservation was fixed to 80% and based on this threshold, the conserved region were obtained using AVANA for all the groups of HA protein. The next step is to find

regions that are conserved in all the groups of a protein which was done manually. These sequences were than further analyzed to predict epitopes.

4.3 EPITOPE PREDICTION

Different immunoinformatics programs were used for both B-cell and T-cell epitopes.

4.3.1 B-CELL EPITOPE PREDICTION

There are different immunoinformatics tools for epitope prediction but only one of them was optimized and used to identify MHC class I epitope.

IEDB Analysis Resource were used for predicting B-Cell epitopes, the software used is **(**http://tools.immuneepitope.org/tools/bcell/iedb_input)

IEDB provides different B cell prediction methods:

1) Chou and Fasman beta turn prediction
2) Emini surface accessibility scale
3) Karplus and Schulz flexibility scale
4) Kolaskar and Tongaonkar antigenicity scale
5) Parker Hydrophilicity Prediction
6) Bepipred Linear Epitope Prediction

Parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes. This has led to a search for empirical rules that would allow the position of continuous epitopes to be predicted from certain features of the protein sequence. All prediction calculations are based on propensity scales for each of the 20 amino acids. Each scale consists of 20 values assigned to each of the amino acid residues on the basis of their relative propensity to possess the property described by the scale.

1) Chou and Fasman beta turn prediction:

The method was developed with the specific aim of predicting only a few peaks for each protein (two or three). Check it It leads to a high level of accurate prediction (70% of correct prediction of known epitopes). This method is useful for selecting protein regions to be synthesized in order to produce anti-peptide antibodies cross-reacting with the parent protein. Chou and Fasman scale is commonly used to predict beta turns.

2) Emini surface accessibility scale

The calculation was based on surface accessibility scale on a product instead of an addition within the window. The accessibility profile was obtained using the formulae $Sn = (\prod^{6} dn+4+i) (0.37)^{-6}$ where Sn is the surface probability, dn is the fractional surface probability value, and i vary from 1 to 6. A hexapeptide sequence with Sn greater than 1.0 indicates an increased probability for being found on the surface.

3) Karplus and Schulz flexibility scale

In this method, flexibility scale based on mobility of protein segments on the basis of the known temperature B factors of the a-carbons of 31 proteins of known structure was constructed. The calculation based on a flexibility scale is similar to classical calculation, except that the center is the first amino acid of the six amino acids window length, and there are three scales for describing flexibility instead of a single one.

4) Kolaskar and Tongaonkar antigenicity scale

A semi-empirical method which makes use of physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes was developed to predict antigenic determinants on proteins. Application of this method to a large number of proteins has shown by the authors that the method can predict antigenic determinants with about 75% accuracy which is better than most of the known methods.

5) Parker Hydrophilicity Prediction

In this method, hydrophilic scale based on peptide retention times during high-performance liquid chromatography (HPLC) on a reversed-phase column was constructed. A window of seven residues was used for analyzing epitope region. The corresponding value of the scale was introduced for each of the seven residues and the arithmetical mean of the seven residue value was assigned to the fourth, (i+3), residue in the segment.

6) Bepipred Linear Epitope Prediction

BepiPred predicts the location of linear B-cell epitopes using a combination of a hidden Markov model and a propensity scale method.

4.3.2 T-CELL EPITOPE PREDICTION

BIMAS tool were used to predict T-Cell epitopes ($CD8^+$ cell binding epitopes) **(**http://www-bimas.cit.nih.gov/molbio/hla_bind/).

4.3.2.1 BIMAS

BIMAS stands for **B**ioinformatics and **M**olecular **A**nalysis **S**ection. BIMAS , a tool of the National Institute of Health (Parker *et al.,* 1994), is the most popular prediction algorithm of peptide-MHC interaction on World Wide Web. The BIMAS tool ranks potential peptides on the basis of predicted half -time of disassociation from HLA class I molecules, which in turn is based on co efficient tables deduced from the published literature. In this method, the contribution to binding from each amino acid at each peptide position with the binding groove is quantified.It is assumed that each position within the peptide contributes independently in binding to a MHC molecule, and a residue located at a given peptide position contributes an equal amount to binding, even within different peptides.

Conserved regions of H5N1 were taken as input. Conserved regions were taken FASTA format and analysis was carried out for all HLA Class 1 molecules. Score was selected in the form of $T_{(1/2)}$ (estimate of

half time dissociation of a molecule containing this subsequence). Different parameters were set as follows :

1) Length of Epitope = ≥ 9 amino acid
2) Predicted $T_{(1/2)}$ = ≥ 1 till 500

4.4 OPTIMIZATION OF THE PROTOCOL FOR MTT CELL PROLIFERATION ASSAY

	REQUIREMENTS	COMPANY
1	Powdered RPMI Media	Himedia
2	Sodium bicarbonate	Himedia
3	L-Gluamine	Himedia
4	Foetal Bovine Serum	Sigma Aldrich
5	Penicillin Sodium	Himedia
6	Streptomycin sulphate	Himedia
7	Amphotericin	Himedia
8	Hisep LSM 1073	Himedia
9	Trypan Blue	Himedia
10	MTT	Sigma Aldrich
11	Dimethyl Sulfoxide	SRL
12	ConA	Sigma Aldrich
13	Sodium chloride	Himedia
14	Potassium chloride	Himedia
15	Sodium hydrogen phosphate	Himedia
16	Potassium dihydrogen phosphate	Himedia

1) **PREPARATION OF POWDERED RPMI MEDIA: (Roswell Park Memorial Institute media)**

10.3 g of powder RPMI media was suspended in 900 ml distilled water and constantly, stirred gently until the powder was completely dissolved and autoclaved for 15 minutes at 121° C and 15 lps pressure in an autoclave. After autoclaving allow it to cool to room temperatue and then add 26.7 ml of 7.5% sodium bicarbonate solution and 10.3 ml of 200 mM L-glutamine solution to 1 liter of medium and stirred until dissolved. pH was adjusted to 4.0 using 1N HCl or 1N NaOH pH of the medium was adjusted ±0.2 below the desired pH since the pH tends to rise during filtration. The final volume was made upto 1000 ml with double distilled water. The medium was immediately sterilized by filtering through a sterile membrane filter with porosity of 0.22 micron or less, using positive pressure rather than vacuum to minimize the loss of carbon dioxide. Liquid medium was stored at 2-8° C and in dark till use. 10% heat inactivated fetal bovine serum (57° C for 30 minutes) and filter sterilized antibiotics [Streptomycin (100µg/ml), Penicillin(100 IU/ml) , Amphotericin (2.5 µg/ml)] were added to media before culturing of cells.

2) PREPARATION OF PBS:

One litre of 1X PBS was prepared by adding 8 g of NaCl, 0.2 g of KCl . 1.44g of Na_2HPO_4. 0.24 g of KH_2PO_4 was added in 800 ml of distilled water. pH was adjusted to 7.4 using HCL and NaOH. Volume was made up to 1 litre by distilled water. PBS was autoclaved for 20 minutes at 121 ^{0} C. After autoclaving PBS was stored at 4^0 C temperature.

3) ISOLATION OF PERIPHERAL BLOOD MONONUCLEAR CELLS FROM WHOLE BLOOD:

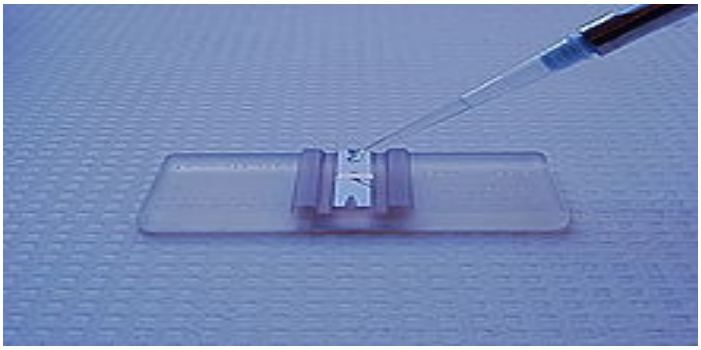
Blood was drawn from a healthy person with the help of vacutainer system (EDTA coated, Becton Dickinson). Blood was diluted in 1:1 ratio with PBS. Then blood sample was layered carefully over Hisep LSM 1073 in the ratio 4:3 and it was centrifuged at 700xg for 40 minutes at 25^0 C. Plasma was removed and then the Buffy coat layer was transferred to a clean centrifuge tube with the help of micropipette. Buffy coat layer was washed twice with 3 volumes of PBS by centrifuging at 1000 rpm for 12 minutes at 25^0C. Supernatant was discarded and pellet of PBMC was suspended in 1 ml of cell culture medium [RPMI + 10% FBS + Penicillin (100units/ml) + Streptomycin (100 µg/ml)].

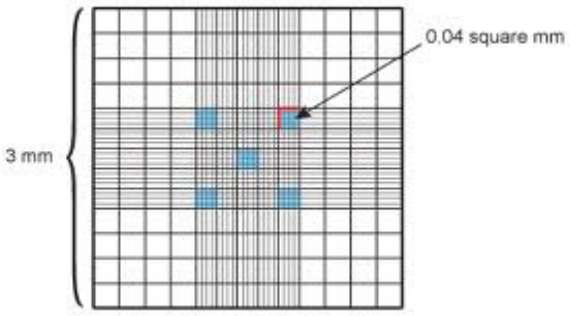
CELL COUNTING AND VIABILITY TESTING:

Cell counting was done with the help of hemocytometer using trypan blue as a stain. Trypan blue is a stain that penetrates through the cell wall of dead cells and stains them in blue color while live cells remain unstained. 10 µl of cell suspension, 80µl of media and 10µl of trypan blue were mixed. Now cell suspension was diluted 10 times to the original cell suspension, and this diluted suspension with trypan blue was loaded on hemocytometer. Hemocytometer was focused on using the 10X objective of the microscope and cells were counted in all 4 sets of squares of hemocytometer using 40 X objective of the microscope.

Cell count was calculated using the formula:

$$\text{Cell count} = \frac{\text{Total number of cells counted}}{\text{Number of chambers counted}} \times \text{Dilution Factor}$$





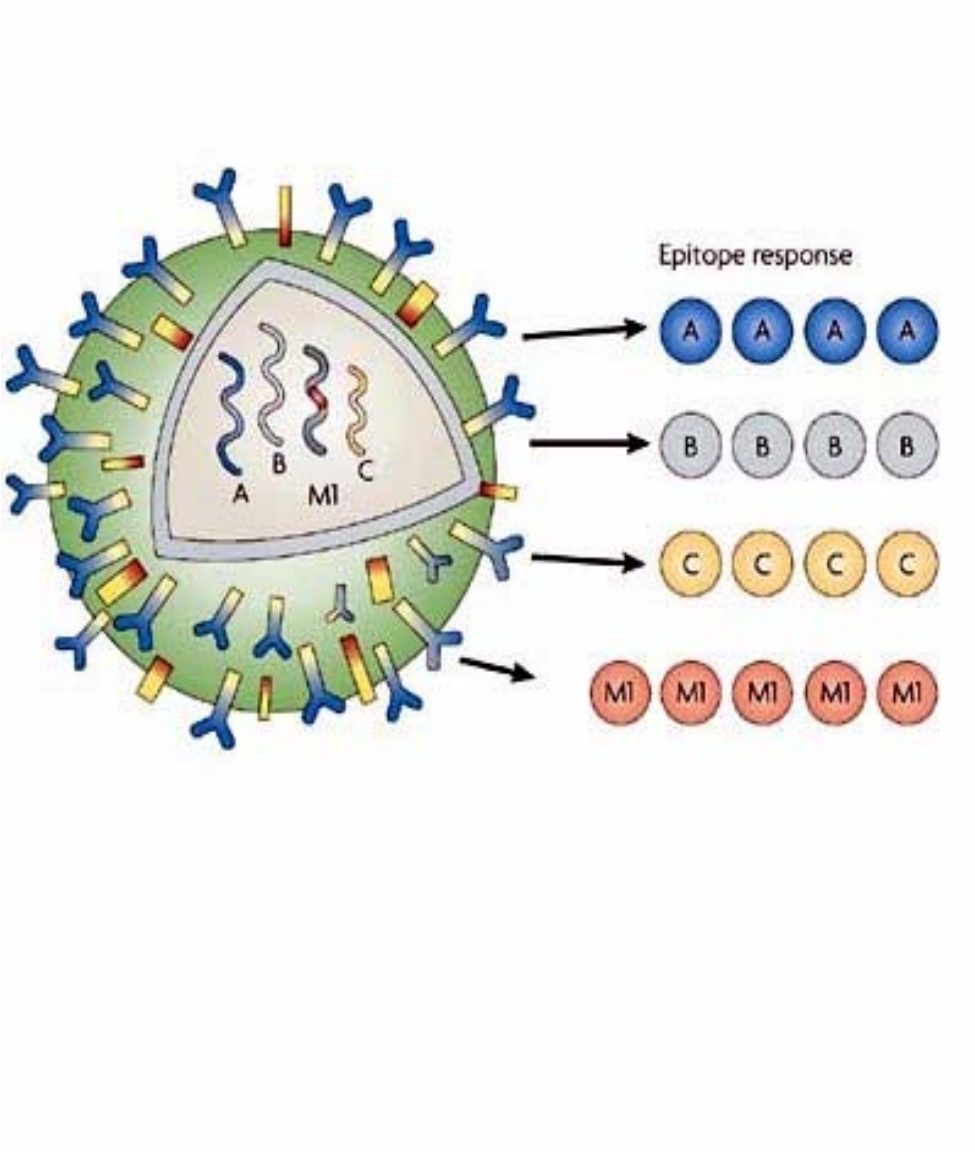
HEMOCYTOMETER

MTT ASSAY:

MTT assay is a calorimetric assay that measures the reduction of yellow 3-(4,5-dimethythiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) by mitochondrial succinate dehydrogenase. The MTT enters the cells and passes into the mitochondria where it is reduced to an insoluble, coloured (dark purple) formazan product. The cells are then solubilised with an organic solvent and the released, solubilised formazan reagent is measured spectrophotometrically. Since reduction of MTT can only occur in metabolically active cells the level of activity is the measure of the viability of the cells. Therefore cell proliferation was tested using MTT assay.

For this assay, freshly isolated lymphocytes ($10x10^5$ cells to $1x10^5$ cells /200 µl media per well) wereseeded in 96-well flat bottom microtiter plate taking ConA (5-10 µg/ml) as positive control (Concanavalin-A is a plant mitogen which stimulates mouse T-cell subsets giving rise to four functionally distinct T-cell populations and hence leads to cell proliferation). Plate was incubated at 37^0 C and 5% CO_2 concentration for 5 days. After 5 days, 10µl of MTT (5 mg/ml) was added to each well and an incubation of another 4 hours was given for reduction of MTT to formazan. Media was removed carefully and purple formazan crystals were dissolved in 100µl of DMSO. Absorbance was recorded at 570 nm by microtiter plate reader (Thermo

RESULTS AND DISCUSSION



5.1 Conserved regions of Hemagglutinin Protein in H5N1 virus

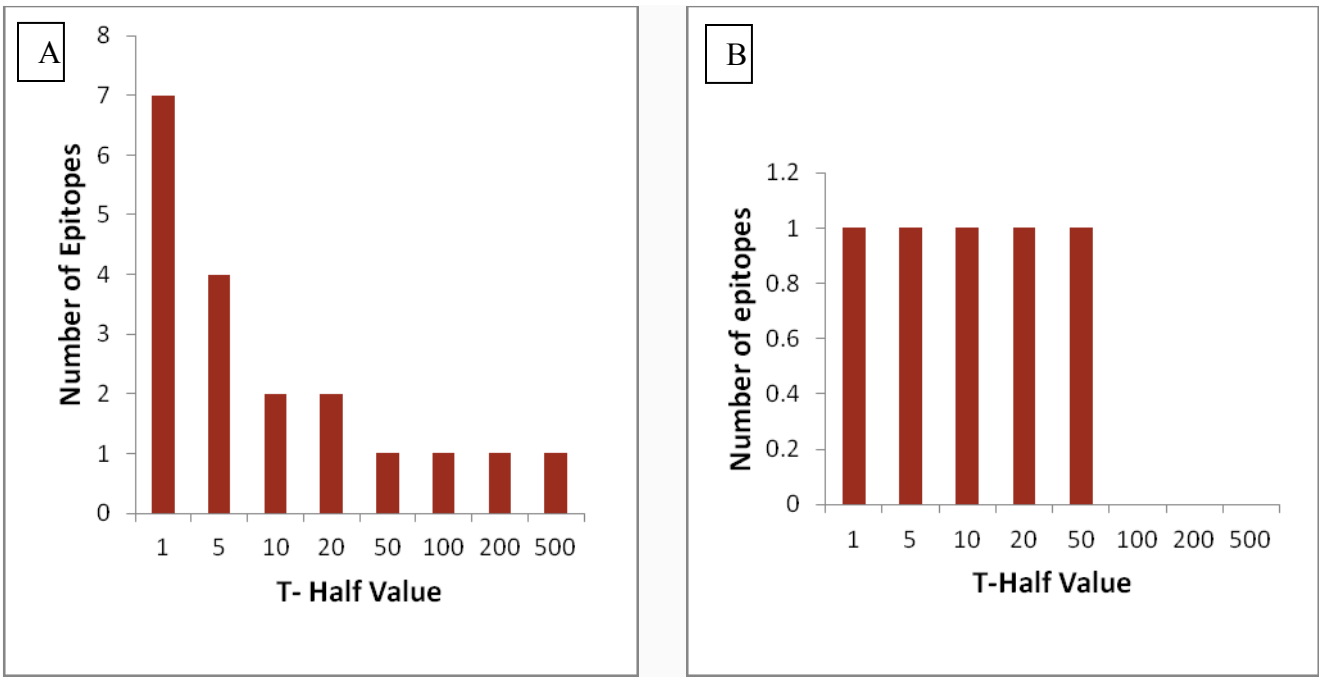
515 HA protein sequences of H5N1 strain were taken from NCBI database These sequences were used to find out the conserved regions using two tools: MUSCLE and AVANA. Nine conserved peptide sequences were obtained which range from 10-135 amino acids, in length (Table)

TABLE 5.1 Conserved Peptides of Hemagglutinin protein in H5N1 influenza virus

Sr.No	Conserved Peptide	Start position	End position
A1	MNPNQKIITIGSICMV	9	24
A2	LAGNSSLCPI	93	102
A3	SKDNSIRIGS	109	118
A4	GDVFVIREPFISCSHLECRTFFLTQGALLNDKHSNGTV DRSPHRTLMSCPVGEAPSPYNSRFESVAWSASACHDG SWLTIGISGPDNGAVAVLKYNGIITDTIKSWRNNINNIL TQESECACVNGSCFTVMTDGPS	120	255
A5	KIFKMEKGKV	262	271
A6	APNYHYEECSCYPDAGEITCVCRDNWHGSNRPWVSF QNLEYQIGYICSGVFGDNPRPNDGTGSCGP	279	345
A7	NGAYGVKGFSFKYGNGVWIGRTKSTNSRSGFEMIWD NGWT	349	389
A8	FSVKQDIVAITDWSGYSGSFVQHPELTGLDCIRPCFW ELIRGRPKESTIWTSGSSISFCGVN	395	457
A9	WSWPDGAELPFTIDK	463	477

5.2 PREDICTION OF T-CELL EPITOPES (MHC Class 1)

BIMAS tool was used to find out the class I MHC binding peptides. BIMAS predict T cell epitope based on $T_{(1/2)}$ which estimate the half time of dissociation of a molecule. We have considered different $T_{(1/2)}$ values to evaluate the T cell epitopes in all the nine conserved peptide sequences obtained.. Number of epitopes were predicted at different $T_{(1/2)}$ values were shown in Table and fig. We found that the number of predicted epitopes is decreasing with increase in $T_{(1/2)}$ values. Considering $T_{(1/2)}$ value equal to 100 and 200 is more stringent as less number of epitopes predited while at 1, 5, 10 and 20, a large number of epitopes were predicted. $T_{(1/2)}$ at 50 and 100, predicted epitopes were comparative. In other study done in our lab on Breast cancer antigen, $T_{(1/2)}$ at 50 is found to be optimal for prediction of epitopes. Finally we have considered the epitopes which were predicted at $T_{(1/2)}$ value equal to 50 as given in table. Seven peptide containing overlapping epitopes were identified (Table)



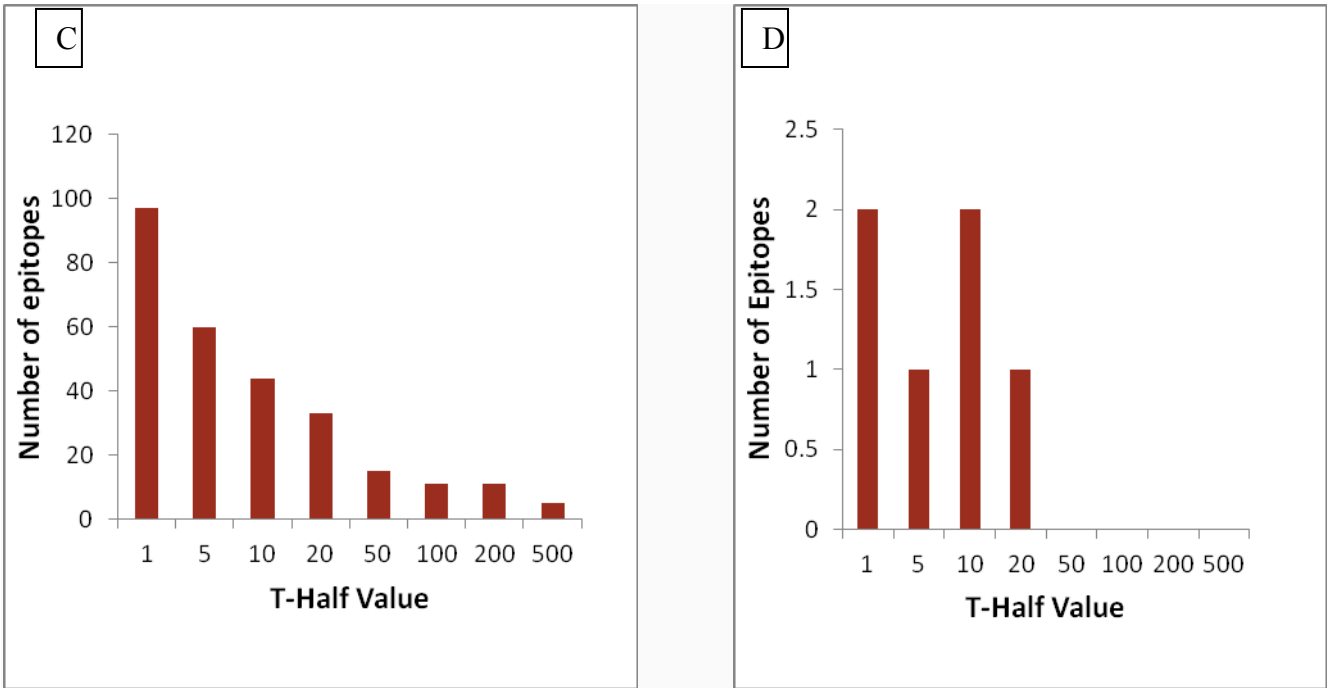


Fig 5.2. Number of predicted epitopes for (A) conserved peptide A1 (B) conserved peptide A2 (C) conserved peptide A4 (D) conserved peptide A5

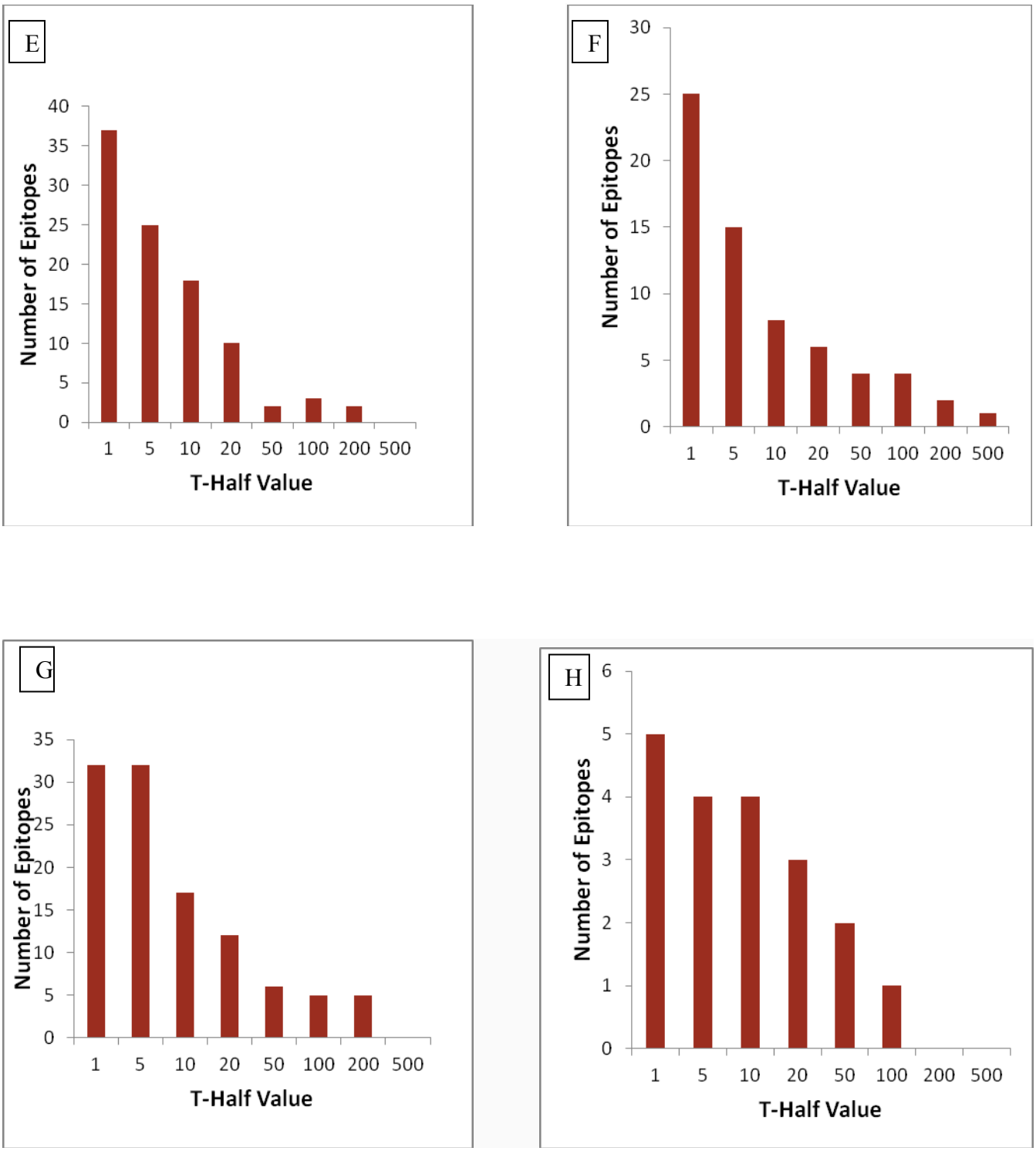


Fig.5.3 Number of predicted epitopes for (A) conserved peptide A1 (B) conserved peptide A2 (C) conserved peptide A4 (D) conserved peptide A5

Table5.2 : Final EPITOPES PREDICTED FOR $T_{1/2\ at}$ 50

CONSERVED PEPTIDE	EPITOPES	OVERLAPPING EPITOPES
A1	NPNQKIITI	------
A2	AGNSSLCPI	------
A4	DRSPHRTLM, EPFISCSHL, GPDNGAVAV, HRTLMSCPV, IREPFISCS, ITDTIKSWR, LRTQESECA, NSRFESVAW ,SHLECRTFF SRFESVAWS, TQGALLNDK, TVKDRSPHR, VAVLKYNGI ,WRNNINNIL	WRNNINNILRTQESECA TVKDRSPHRTLMSCPV GPDNGAVAVLKYNGII
A6	CRDNWHGSN, IGYICSGVF	-------
A7	AYGVKGFSF, GRTKSTNSR, KSTNSRSGF, SRSGFEMIW	GRTKSTNSRSGFEMIW
A8	GRPKESTIW, GSFVQHPEL, IRGRPKEST, IRPCFWVEL, RPCFWVELI, VQHPELTGL	GSFVQHPELTGL IRPCFWVELIRGRPKESTIW
A9	DGAELPFTI, WSWPDGAEL	WSWPDGAELPFTI

5.3PREDICTION OF B-CELL EPITOPES

IEDB Analysis Resource was used to predict B-Cell epitopes. Six different methods based on different parameters were involved in IEDB for epitope prediction. We have found a different number of epitopes using different methods as shown in table (5.3.1, 5.3.2, 5.3.3). Finally, we have selected only those epitopes which are predicted by atleast three tools. We have found 11 epitopes which are predicted by more than three methods (Table). Finally, two peptides containing overlapping epitope were identified (Table).

TABLE 5.3.1 : Immunogenic peptides containing B-cell epitopes predicted by methods Kolaskar & Tongaonkar Antigenicity and Bepipred)

CONSERVED SEQUENCES	Kolaskar and Tongaonkar Antigenicity	Bepipred Linear Epitope Prediction
	EPITOPE	EPITOPE
A1 (16)	KIITIGSI	
A2 (10)		
A3 (10)		SKDNSI
A4 (135)	FVIREPFISCSHLECRTFFLTQGALLND	KHSNGTVKDRSPH
	HRTLMSCPVGEAP	VGEAPSPYNSRF
	ESVAWSASACHDG	
	GAVAVLKYN	ISGPDNGA
	SECACVNGSCFTV	
A5 (10)		
A6 (66)	YEECSCYPDA	APNYHY
	EITCVCRD	SCYPDAG
	EYQIGYICSGV	WHGSNRPWV
A7 (40)	YGVKGFSFK	TKSTNSRS
	GNGVWIG	
A8 (62)	KQDIVAIT	GRPKESTIWTS
	SGSFVQHPELTGLDCIRPCFWVELI	
A9 (14)	AELPFT	WSWPDGAELP

TABLE 5.3.2 : Immunogenic peptides containing B-cell epitopes predicted by methods Chou and Fasman Beta turns and Emini surface accessibility prediction :

CONSERVED SEQUENCES	Chou and Fasman Beta-turn Prediction	Emini surface accessibility prediction
	EPITOPE	EPITOPE
A1 (16)		
A2 (10)		
A3 (10)		
A4 (135)	DKHSNGTVKDRSPH	GTVKDRSPHRT
	PVGEAPSPYNSR	APSPYNSRF
	SACHDGTS	TIKSWRNN
	ISGPDNG	ILRTQE
	ACVNGSC	
	IKSWRNNIN	
A5 (10)		
A6 (66)	NWHGSNRP	NWHGSNRP
	VFGDNPRPNDGTGS	DNPRPNDG
A7 (40)	KSTNSRS	RTKSTNSRSG
	FKYGNGV	
A8 (62)	DWSGYSGSFV	DWSGYSGS
	WTSGSSISFC	RGRPKESTIWTSGSSI
A9 (14)		

.

TABLE 5.3.3 : Immunogenic peptides containing B-cell epitopes predicted by methods Karplus and Schulz Flexibility Prediction :

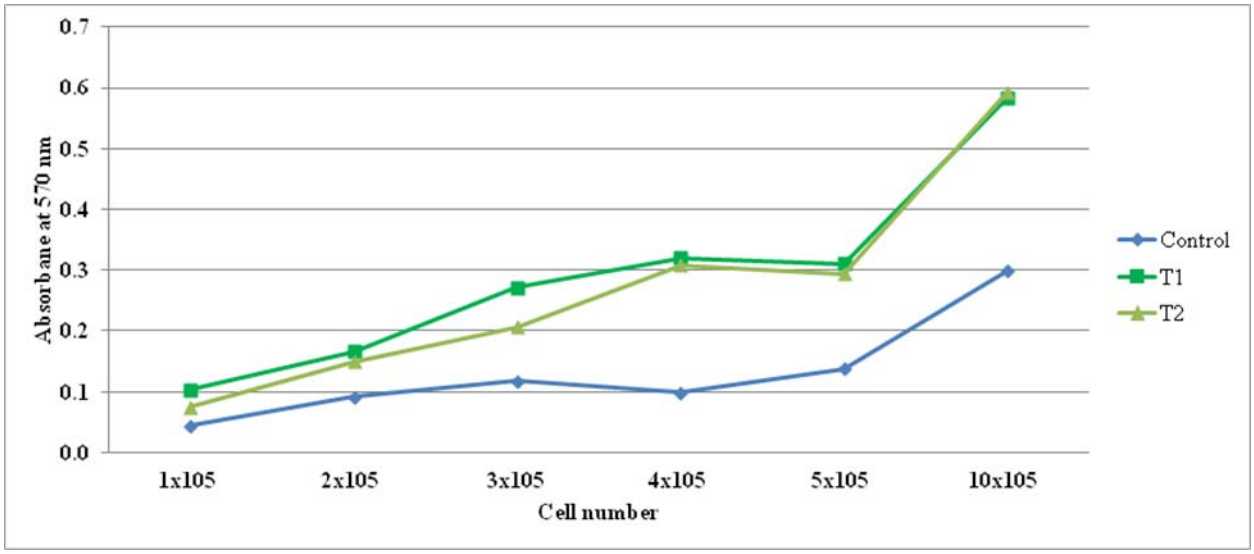
CONSERVED SEQUENCES	Karplus ans Schulz Flexibility Prediction	Parker Hydrophilicity Prediction
	EPITOPE	EPITOPE
A1 (16)		
A2 (10)		
A3 (10)		
A4 (135)	DKHSNGTVKDRSPHR	KHSNGTVKDRSPH
	GEAPSPYNSRFE	PVGEAPSPYNSRFE
	ISGPDNG	ACHDGTS
	RTQESE	SGPDNGA
	DTJKSWRNNI	LRTQESECACVN
A5 (10)		
A6 (66)	HGSNRP	HYEECSCYPDAGEI
	GDNPRPNDGTG	PRPNDGTGS
		NWHGSN
A7 (40)	RTKSTNSRSG	TKSTNSRSG
A8 (62)	DWSGYSGS	KQDIVAIT
	RGRPKESTIWTSGSSI	SGSFVQHPELTGLDCIRPCFWVELI
A9 (14)		AELPFT

TABLE : 5.3.4 Common Epitopes predicted in more than three methods of IEDB

CONSERVED PEPTIDE	COMMON EPITOPES	OVERLAPPING EPITOPES
A4	**GTVKDRSPH** **APSPYNSRF** **SACHDG** **SGPDNG** **ACVNGSC** **KSWRNNI**	**TVKDRSPHRTLMSCPV** **GPDNGAVAVLKYNGII**
A6	**HGSNRP** **PRPNDG**	**------**
A7	**KSTNSRS**	**------**
A8	**DWSGYSGS** **WTSGSSI**	**-------**

5.4 Optimiztion of protocol for PBMC Proliferation Assay at preliminary steps (MTT Assay

In order to opmize the effect of mitogen, we have done the lymphocyte proliferation assay. We have taken different cell numbers at two different concentration (5µg/ml and5µg/ml)of concavillin A mitogen (Fig no). We have found that with increase in cell number, there is more proliferation in compared to control. We did not find observed significant difference in two different concentration of mitogens.



(T1= 5µg/ml , T2=10ug/ml)

Fig: Proliferation assay of peripheral blood mononuclear cells using MTT assay

We have found that with increase in cell number, there is more proliferation in compared to control. We did not find observed significant difference in two different concentration of mitogens.

CONCLUSION

It was concluded that H5N1 Strain of influenza virus A shows a high level of conservancy of Hemagglutinin protein (HA). In this study, nine peptide sequences of 10-135 amino acids were found to be conserved in HA protein from 1918 to 2012.

Thirty and eleven MHC Class I T-cell epitopes and B-cell epitopes respectively were predicted **Seven and two** immunogenic peptides were selected by considering the overlapping epitopes for T and B cell response respectively**.**

These selected immunogenic peptides containing T and B cell epitopes can be considered for further study to evaluate the binding affinity of immunogenic peptide to MHC molecule by structural analysis and molecular modelling. Further it will interesting to estimate the immunogenic potential of these peptides in the PBMC by T-cell proliferation assay (MTT assay) and cytokine production assay.

Hence, these immunogenic peptides may be considered as an interesting candidate in designing vaccine against H5N1 influenza virus.

.

SUMMARY

Influenza A viruses belongs to one of the best studied viruses. Influenza viruses circulating in animals pose threats to human health. Humans can become ill when infected with viruses from animal sources, such as avian influenza virus subtypes H5N1 and H1N1. The primary risk factor for human infection appears to be direct or indirect exposure to infected live or dead animals or contaminated environments.. Current influenza virus vaccines protect mostly against one particular strain. Hence great challenge in the field of influenza virus research is to design a universal vaccine.

The resistance in vaccine during the past decade has led to several new approaches of vaccine development. One of these approach is synthetic peptides for eliciting protective immunity against infection. Synthesizing peptides for use as vaccines requires identification of those epitopes in the protein antigen that stimulate protective immunity. Both B and T cell epitopes must be included in the peptide so that arms of the immune system humoral and cell-mediated are stimulated.

In the course of our study, we have used various bioinformatics tools to determine peptide. sequences in Hemagglutinin (HA) viral protein that act as epitope. Our approach was to find a stretch of immunogenic peptide from conserved peptide sequences as target for vaccine design which may be effective against current and future H5N1 influenza viruses .

The first step of our methodology was to retrieve various sequences of HA protein of H5N1 influenza virus from NCBI influenza database. The sequences were aligned using multiple sequence alignment tool MUSCLE. Nine conserved peptide sequences were identified out from aligned sequences using AVANA. Then, these conserved peptide sequences were further considered for epitope prediction using immunoinformatics tools. BIMAS and IEDB Analysis Resource were used for T-Cell and B-Cell epitopes respectively.

Thirty and eleven MHC Class I T-cell epitopes and B-cell epitopes respectively were predicted based on our criteria in immunoinformatics tools. Finally we have selected **seven and two** immunogenic peptides by considering the overlapping epitopes for T and B cell respectively**.**

Hence it will be interesting to carry out structural analysis and molecular modeling to estimate the binding affinity of selected immunogenic peptide to different MHC molecules. Further, these peptides can be selected for chemical synthesis and then carried out lymphocyte cell proliferation assay (MTT assay) and cytokine production to assess their immunogenic response.

If these peptides show good response in the in vitro system then can be considered interesting candidate for vaccine design against H5N1 influenza virus.

REFERENCES

Nardin E, Calvo-Calle J, Oliveira G, Nussenzweig R, Schneider M, Tiercy J, Loutan L, Hochstrasser D, Rose K: A totally synthetic polyoxime malaria vaccine containing *Plasmodium falciparum* B cell and universal T cell epitopes elicits immune response in volunteers of diverse HLA types. *J Immunol* 2001, 166:481-489.

Parker J, Guo D, Hodges R: New hydrophilicity scale derived from High-Performance Liquid Chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 1986, 25:5425-5432.

Chou P, Fasman G: Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* 1978:45-148.

McSparron H, Blythe M, Zygouri C, Doytchinova I, Flower D: Jen-Pep: a novel computational information resource for immunobiology and vaccinology. *J Chem Inf Comput Sci* 2003, 43(4):1276-1287.

Pellequer J, Westhof E, Van Regenmortel M: Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunol Lett* 1993, 36:83-99

Barlow D, Edwards M, Thornton J, et al: Continuous and discontinuous protein antigenic determinants. Nature 1986, 322:747-748.

Langeveld J, martinez Torrecuadrada J, boshuizen R, Meloen R, Ignacio C: Characterisation of a protective linear B cell epitope against feline parvoviruses. Vaccine 2001, 19:2352-2360.

Van Regenmortel M: What is a B-cell epitope? Methods in molecular biology (Clifton, NJ) 2009, 524:3. Reineke U, Schutkowski M: . Epitope mapping protocols. Preface Humana Press, 2 2009.

EL-Manzalawy Y, Dobbs D, Honavar V: Predicting linear B-cell epitopes using string kernels. J. Mol. Recognit. 2008, 21:243-255.

EL-Manzalawy Y, Dobbs D, Honavar V: Predicting linear B-cell epitopes using evolutionary information. IEEE International, Conference on Bioinformatics and Biomedicine 2008

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 54, 905–921

DeLano, W. L. (2002) *The PYMOL User's Manual*, DeLano Scientific, San Carlos, CA, USA

Fernandez-Fuentes, N., Zhai, J., and Fiser, A. (2006) ArchPRED: a template based loop structure prediction server. *Nucleic Acids Res.* 34, W173–W176

Elton, D., Medcalf, E., Bishop, K., and Digard, P. (1999) Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. *Virology* 260, 190–200

Prokudina, E. N., Semenova, N. P., Chumakov, V. M., and Rudneva, I. A. (2004) Transient disulfide bonds formation in conformational maturation of influenza virus nucleocapsid protein (NP). *Virus Res.* 99, 169–175

24. Albertini, A. A., Wernimont, A. K., Muziol, T., Ravelli, R. B., Clapier, C. R., Schoehn, G., Weissenhorn, W., and Ruigrok, R. W. (2006) Crystal structure of the rabies virus nucleoprotein- RNA complex. *Science* 313, 360–363