Conserved envelope peptides containing multiple HLA-restricted T cell epitopes as vaccine candidate against Zika virus

A Dissertation

Submitted in the partial fulfillment of the requirement for

the award of the degree of

MASTER OF SCIENCE

IN

BIOTECHNOLOGY

By
Tavleen Kaur
301501019

Under the guidance of
Dr. Manoj Baranwal
Associate Professor

DEPARTMENT OF BIOTECHNOLOGY

THAPAR UNIVERSITY, PATIALA (PUNJAB)-147004

JULY 2017

CANDIDATE'S DECLARATION

I, hereby declare that the work presented in the thesis entitled **"Conserved envelope peptides containing multiple HLA-restricted T cell epitopes as vaccine candidate against Zika virus"** in the partial fulfillment of the requirement for the award of the degree of Master of Science in Biotechnology, Department of Biotechnology, Thapar University, Patiala, is an authentic record of my work during the period of six months from Jan,2017 to July,2017 under the guidance of Dr. Manoj Baranwal, Assistant Professor, Thapar University, Patiala. I have not submitted the matter embodied in this thesis for the award of any other degree or diploma.

Tavleen Kaur

Tavleen Kaur
(301501019)

Place: Patiala
Date: 17/07/2017

CERTIFICATE

This is to certify that the thesis entitled **"Conserved envelope peptides containing multiple HLA-restricted T cell epitopes as vaccine candidate against Zika virus"** submitted by Tavleen Kaur in partial fulfillment of the requirement for the award of Degree of Master of Science in Biotechnology to Thapar University, Patiala, in record of student's own work carried out by her. The report has not been submitted for the award of any other degree or certificate in this or any other University or Institute.

Dr. Manoj Baranwal
Associate Professor, Supervisor
Department of Biotechnology, TU
Patiala

ACKNOWLEDGEMENT

This is a golden opportunity for me to convey my sincere regards to all those people who enabled me to accomplish my project successfully. It is my privilege to express my sincere gratitude to my supervisor, **Dr. Manoj Baranwal**, Assistant Professor, for his encouragement and valuable advice. I am thankful that he gave me a brilliant opportunity to conduct my M.Sc thesis under his guidance. His confidence imbibing attitude, continuous support, valuable advice and scientific approach helped me to a great extend to accomplish my work. I express my sincere thanks to **Dr. Moushami Ghosh**, Professor and Head, Department of Biotechnology, Thapar University, Patiala, for providing me with the best laboratory facilities during my research pursuit. Special thanks to **Mr. Sahil Jain** research scholar (DBT), for his immense help and valuable suggestions throughout my study. I would also like to thank my dear friends **Mehendi Goyal** and **Harlovleen Kaur**, for their constant encouragement as it motivated me to work hard and give my best efforts during my research project. Last but not the least I am grateful to my **family** for their love, concern guidance and support throughout my project.

Tavleen Kaur

(301501019)

Table of Contents

ABSTRACT

Envelope protein in Zika virus play important role in assembly of virions, host cell binding and membrane fusion and cover majority of the virion surface, which makes them ideal candidate for development of peptide vaccines against Zika virus. In the present study, conserved peptides containing multiple T cell epitopes were identified by different epitope predictions algorithms and docking studies were carried out to establish the binding affinity of these peptides with different HLA molecules. Employment of six epitope prediction tool and BLAST analysis resulted into seven non-self peptide fragments containing multiple HLA restricted T cell epitopes. The average population coverage for immunogenic response of these selected peptides was found out to be 76.37% and 98.13% for class I and II HLA specific respectively. Further, docking studies reveal that $CD8^+$ T cell epitopes and CD4+ peptides which are part of selected were have shown good binding affinity with different HLA molecules. Six peptides were found to be non-allergic and non-toxic by *in-silico* analysis. Further, some of these peptides were found to be conserved in other viruses of flaviviridae family infecting human. Hence, it is suggested that these peptides may be considered for synthetic peptide vaccine design against Zika virus.

LIST OF TABLES

LIST OF FIGURES

Chapter 1 INTRODUCTION

Zika virus (ZIKV) infection, a mild febrile disease which is primarily spread to people through the bite of an infected *Aedes* species mosquito (Sikka *et al*., 2016). A Zika virus infection in human was majorly reported in Africa, Asia, and Oceania with sporadic cases in America and Europe (Plourde *et al*., 2016). Due to its rapid widespread, Zika virus infection declared as a public health emergency of international concern by World Health Organization in February 2016 (Roberts *et al*., 2016). Neurological disorders such as Guillain-Barre syndrome in adults and microcephaly in babies born to infected mothers was found to be associated with this virus infection (Plourde *et al*., 2016).

ZIKV is related to other flaviviruses responsible for causing diseases such as Japanese encephalatitis, west Nile, dengue and yellow fever (Lanciotti *et al*., 2007). Zika virus genome structure consists of a positive single stranded RNA virus encoding a polyprotein which undergoes proteolytic cleavage into capsid protein(C), precursor membrane protein (prM) and Envelope protein (E) as structural and seven non- structural proteins. E- Protein covers majority of the virion surface and consists of three structural domains (DI, DII, and DIII) (Goo *et al*., 2017). It plays an important role in assembly of virions, host cell binding and membrane fusion (Faye *et al*., 2014). Virus enters via cellular receptors of skin cells, and migrates to the lymph node and the bloodstream. Studies have reported that ZIKV might enter through human skin fibroblasts, keratinocytes, and immature dendritic cells (Hamel *et al*., 2015). ZIKV replication is enhanced in skin fibroblast due to presence of several adhesion factors (e.g., AXL receptor tyrosine kinase) which facilitate infection and cellular autophagy needed for flaviviral replication (Hamel *et al*., 2015). It was reported that ZIKV infection evades the human type 1 interferons response (Winkler *et al*., 2017). Further studies conducted in mice have shown the critical role of CD4+ and CD8+ T cells in ZIKV infection (Manangeeswaran *et al*., 2016 and Pardy *et al*., 2017). Currently, there is no approved vaccine against ZIKV but the development of vaccine is much in tune due to global concern of this virus. Scientists at NIAID's Vaccine Research Centre (VRC) have developed a DNA-based vaccine which has entered in Phase 1 clinical trial (www.niaid.nih.gov). NIAID's laboratory has developed a live attenuated vaccine against Zika virus and dengue virus infection. This vaccine is currently being evaluated in a large phase 3 study in Brazil (www.niaid.nih.gov).

Peptide based vaccine focusing on conserved epitopes fragments have shown significant progress with some peptide vaccine is in different clinical stages (Nardin., 2010). Using experimental methods for epitope identification can prove to be a very expensive and time consuming therefore immunoinformatics driven approach for identification of epitopes against T and B cell epitopes came into picture (Oyarzun *et al*., 2015). Immunoinformatics identified peptides enriched epitopes have proven to be effective in *in vitro* and *in vivo* system (Rojas-Caraballo *et al*., 2014, Lohia *et al*., 2017). In the present study, conserved peptides containing multiple T cell epitopes were identified by different epitope predictions algorithms and docking studies were carried to establish the binding affinity of these peptides with different HLA molecules.

Chapter 2 REVIEW OF LITERATURE

2.1 Zika Virus: History of the disease

Zika was firstly discovered in a rhesus monkey in Zika backwoods of Uganda in 1947 (Sikka *et al*., 2016). In 1952, the first ever case of human infection due to Zika virus was reported in Nigeria (Dick *et al*., 1952). From that point onwards, spread of Zika has been noticed in tropical Africa, Southeast Asia and the Pacific Islands. In 2007, a noteworthy episode of Zika infection was registered in Federated States of Micronesia (Duffy *et al*., 2009). Zika virus came into limelight in year 2015-16 due to its epidemic spread in American continents with extensions to Europe. A Public Health Emergency of International Concern in regard of Zika infection was pronounced by the World health organization on February 1, 2016(Sikka *et al*., 2016). In addition, 29 different nations in America have reported massive Zika infection transmission, including Puerto Rico and US Virgin Islands (Hennessey *et al*., 2016).

2.2 Taxonomy of Zika virus:

Zika virus (ZIKV) has a place with the family *Flavivirdae* and the genus *Flavivirus.* It is spread by daytime-dynamic *Aedes* mosquitoes (*A. aegypti and A. Albopictus*) (Malone *et al*., 2016). In the late 1800s or early 1900s, two distinct lineages of Zika virus emerged from East Africa named as African and Asian lineages respectively. The virus migrated from Africa to Southeast Asia (Malaysia) that led to the origin of Asian lineage (Gatherer *et al*., 2016) .The phylogenetic analysis so performed suggested that the African strains are of two types-the ZIKV MR 766 model strain categorized as the Uganda and Nigeria group and the ZIKV 2007 Micronesian and Malaysian strains constituting the Asian clade (Faye *et al*., 2014). Both epidemiologic and grouping information are in support of the theory that the prevailing strains of ZIKV arose due to hereditary change in the Asian ancestry infection. The classification of ZIKA virus is mentioned in table.

Table 1: Taxonomic classification of ZIKA virus

GROUP	FAMILY	GENUS	SPECIES	ORDER
Group IV ((+)ss RNA)	*Flaviviridae*	*Flavivirus*	Zika virus	Unspecified

2.3 Zika virus infection:

It is the most essential arboviral ailment in terms of mortality (www.nejm.org). The infection has been found to incite a life–long defensive insusceptibility. ZIKV contamination can be asymptomatic or a self-constrained, intense febrile ailment (www.who.int). Because of absence of immunization and antiviral medications, the sole measure of control is constraining the *Aedes* mosquito vectors (www.cdc.gov).

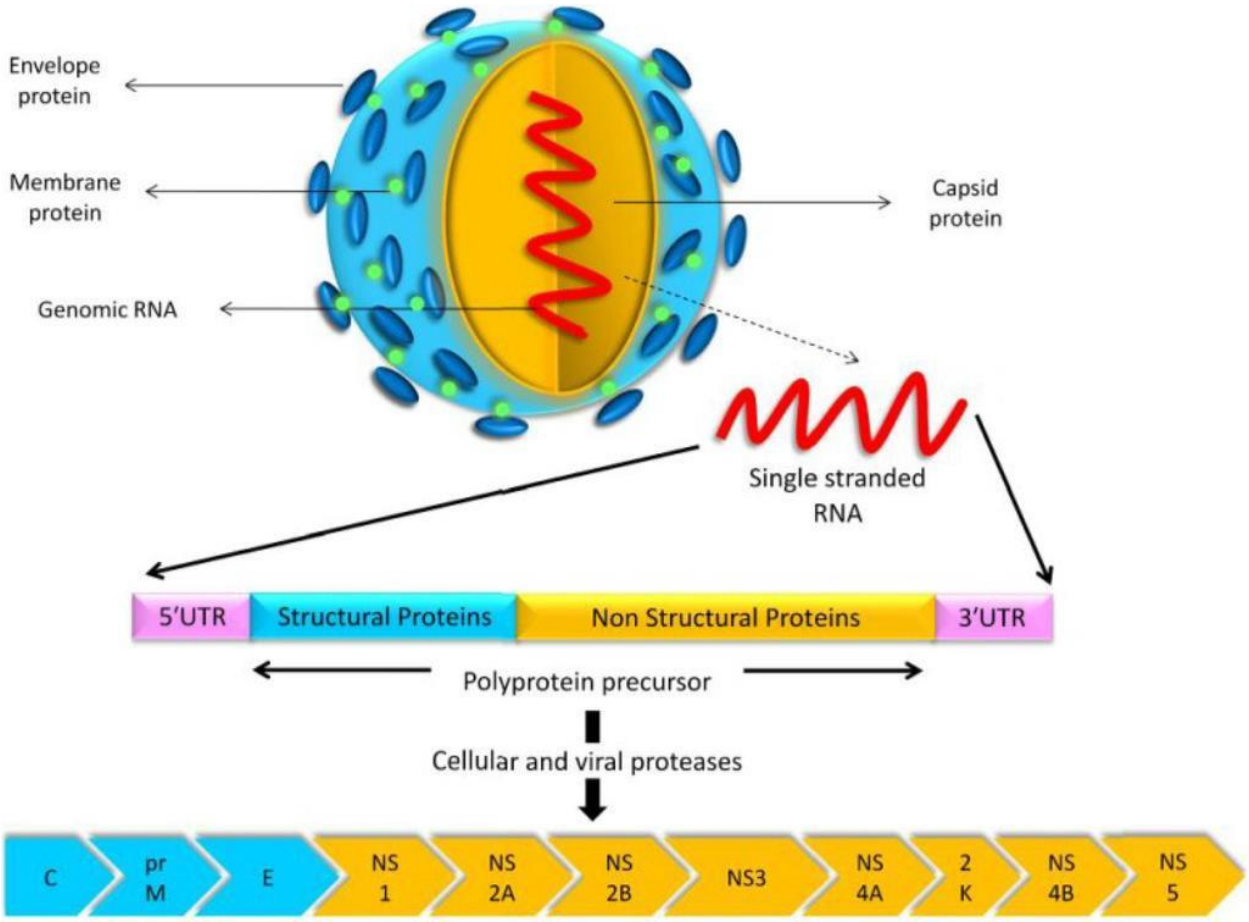


Figure 1: Structure of Zika virus (mdpi.com)

2.4 Zika virus structure

Genome structure of Zika infection involves a positive single stranded RNA which consists of 10,974 bases and 2 non-coding regions (NCR) existing as 5'NCR and 3'NCR. The virion diameter is around 40 nm and it also contains surface projections of 5-10 nm orchestrated in an icosahedral symmetry. Further, the infection likewise includes a nucleocapsid incorporated by a lipid bilayer, containing the prM and E envelope protein (Kostyuchenko *et al*., 2016). The open reading frame (ORF) in Zika infection genome is 5'- C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-3' and this ORF codes for a polyprotein that is separated into capsid (C), precursor membrane (prM) , envelope (E) and non-basic proteins (NS) (Chambers *et al*., 1990). Amongst the non-structural proteins, some are large in size and highly conserved like NS1, NS3,NS5 whereas some are small and hydrophobic in nature like NS2A, NS2B, NS4A, and NS4B (Chambers *et al*., 1990). In the 3'

NCR region of genome, 428 nucleotides are located which play a very important role in translation, cyclization, genome stabilization, and recognition (Faye *et al*., 2014). A loop structure is formed by the 3'NCR region. The 5'NCR permits translation by means of a methylated nucleotide cap (Flaviviridae, Swiss Institute of Bioinformatics).

2.5 Structural proteins

Capsid (C), precursor membrane (prM), envelope (E) is the order of the genes for the structural proteins starting from the 5' end (viralzone.expasy.org). Selective cleavage of the precursor polypeptide results in the origin of these proteins.

2.5.1 Nucleocapsid or Core protein(C)

The first polypeptide to be synthesized during translation is C protein which has a sub-atomic weight of around 13,500 and it is enriched with lysine and arginine amino acid deposits (~ 25%). This protein can associate with the virion RNA. C protein does not have a N-terminal (hydrophobic flag sequence) which recommends that its synthesis occurs on non-layer-bound ribosomes. At the carboxy end, there exists a stretch of hydrophobic amino acids which likely goes about acting as transmembrane signal for the neighboring M-protein precursor, prM (Henchal *et al*., 1990).

2.5.2 <u>Membrane associated protein (M)</u>

During virus development, proteolytic break down of a glycosylated prM precursor (22,000) with furin protease brings about the arrangement of the 8,000-dalton M protein. This proteolysis takes place under acidic conditions inside Golgi apparatus and the virus particles are released from the cell owing to decreasing levels of prM (Yu *et al*., 2008 and Heinz *et al*., 1994). The generation of M from prM is a pivotal and terminal stage in virion morphogenesis. It brings about a substantial increment in virus infection and a rearrangement of the virion surface structure, which in turn is made out of E-prM heterodimers in juvenile virions (Zybert *et al*., 2008, Randolph *et al*., 1990). The role of M protein in viral development is still unknown.

2.5.3 <u>Envelope protein (E)</u>

The virion envelope glycoprotein whose atomic weight is 51,000 to 60,000 exists as a homotrimer or might be discovered intracellularly in the form of E-prM heterodimers (Henchal *et al*., 1990). Nowak and Wengler in 1987 discovered the area of the disulfide linkages (shaped by cysteine deposits) and gave an auxiliary model to the envelope protein. When correlation between structural properties of epitopes and disulfide bridge assignment was obtained then this model was refined (Mandl *et al*., 1989). The existing model of E protein comprises of three non-overlapping antigenic spaces that are made of minimum 16 distinct epitopes). A, B, and C. Areas A and B contain irregular antigenic determinants whose integrity relies on intact disulfide bridges. On the other hand, the C domain is highly variable and does not contain disulphide bridges and moreover, epitopes in this particular region are not demolished by phenomenon like reduction, carboxymethylation or sodium dodecyl sulfate denaturation.

2.6 Non Structural proteins

2.6.1 NS1

NS1 is a monomeric and hydrophilic glycoprotein which is produced in the rough endoplasmic reticulum. It plays a very significant role in replication, virus morphogenesis and immune evasion (Winkler *et al*., 1989 and Rice *et al*., 1986).

2.6.2 NS2

NS2 is a membrane-associated cysteine autoprotease that mediates cleavage of its own carboxyl terminus from NS3. Its coding region codes for proteins NS2a and NS2b respectively. NS2B is a small protein (~14 kDa) with a central hydrophilic part (residues 49–89) involved in binding to NS3. The function of NS2a is assembly and replication and that of NS2b is that it acts as a serine protease co-factor (Bollati *et al*., 2010).

2.6.3 NS3

The proteolytic role of the NS3 protease has been well understood and documented (Murray *et al*., 2008). It acts as a serine protease, helicase, replication and RNA phosphatase (Bazan *et al*., 1989).

2.6.4 NS4

It's coding region codes for two proteins v.i.z, proteins NS4A and NS4B. They are highly hydrophobic transmembrane proteins accountable for the membrane arrangements leading to viral replication complex formation. This complex is a prerequisite for the viral life cycle (Nemesio *et al*., 2012). NS4a is involved in replication, assembly, induction of membrane rearrangements (Chambers *et al*., 1989).

2.6.5 NS5

Flavivirus NS5 consists a C-terminal RNA-dependent-RNA polymerase (RdRp) domain as well as a N-terminal MTase. Flavivirus NS5 MTase belongs to SAM-dependent Mtases family (Zhou *et al*., 2007). It acts as a methyl transferase and RNA dependent RNA polymerase (Murray *et al*., 2008).

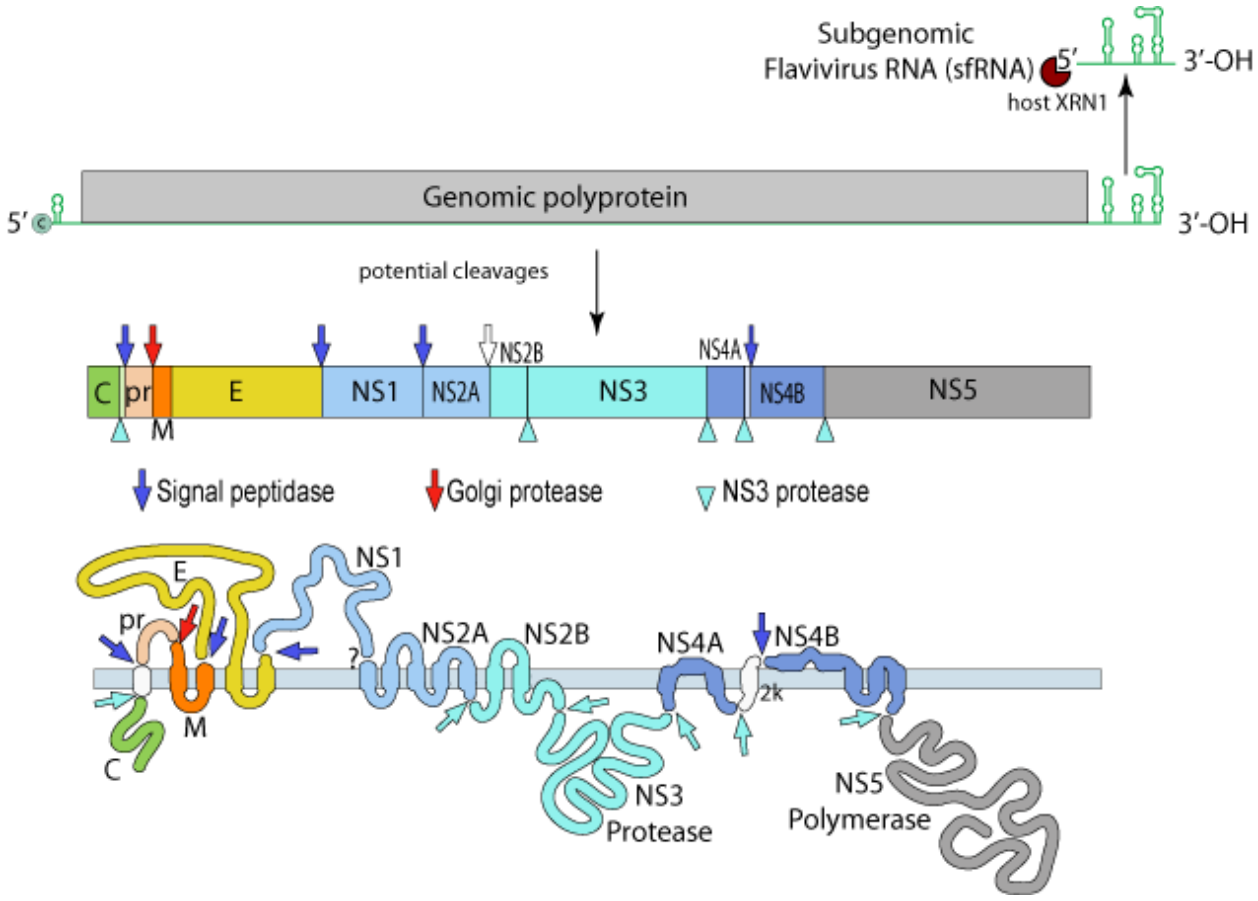


Figure 2: Zika virus proteins (viralzone.expasy.org)

2.7 Life Cycle of Zika Virus

First and foremost, the virus life cycle involves the arrival of the virion through the adhesion molecules on the cell surface and skin receptors followed by its embedment through endocytosis. Under the influence of acidic conditions of the endosome, the existing viral glycoproteins arbitrate integration of viral and cellular membranes, which further allows dismantling of the virion and liberation of vRNA into the cytoplasm. Then there is translation of vRNA into a polyprotein which is then subsequently cleaved by means of viral and cellular proteases followed by replication of the RNA genome utilizing NS proteins. Finally, there is assembly of virus which happens at the endoplasmic reticulum (ER) membrane, and it is here that the C protein and vRNA are encapsulated by the ER membrane and glycoproteins ultimately resulting in the formation of immature virus particles. After this, these virus particles are conveyed through the secretory pathway, and in the low pH region of the trans-Golgi network (TGN), furin-mediated break down

of prM protein handles the development of the virus. Eventually**,** adult virus is liberated from the cell (www.researchgate.net).

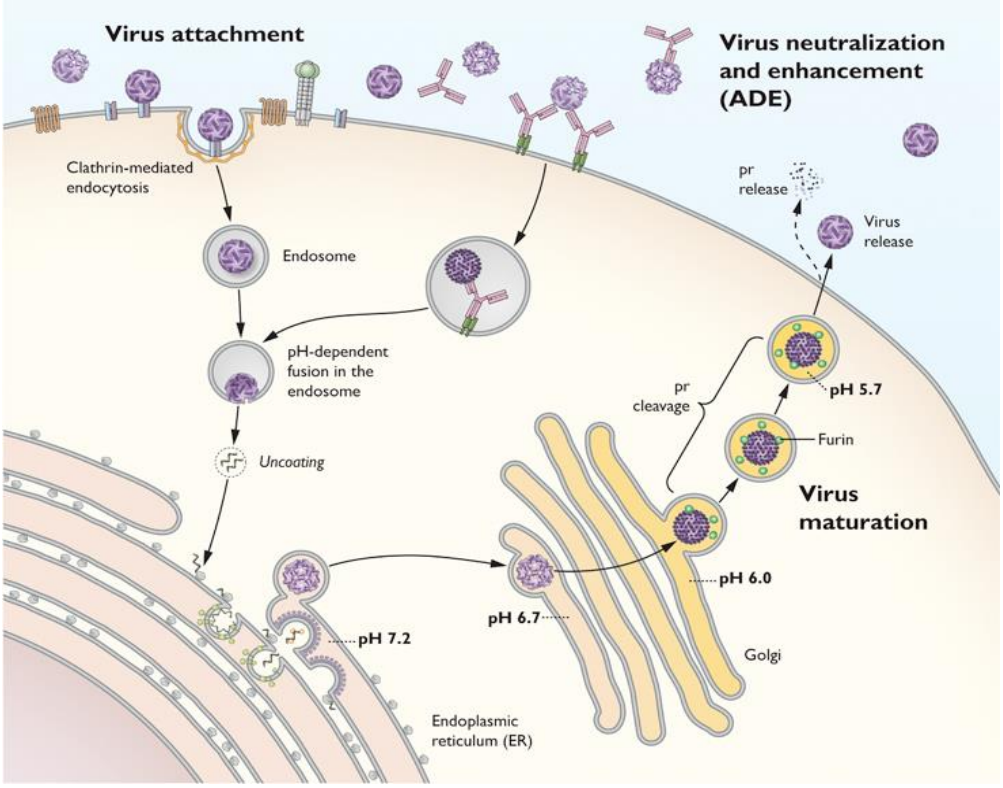


Figure 3: Zika virus lifecycle (blogs.discovermagazine.com)

2.8 Mutation in Zika Virus

Zika virus has experienced critical hereditary changes in the previous 70 years and these progressions have empowered it to infect and attack new tissues that give defensive specialties to viral proliferation, or evade the immune system, prompting viral constancy". The strain of the viral infection is now spreading across south and Central America and the Caribbean is not the same as that first seen in Africa (www.nbcnews.com). As per the group at UCLA (University of California, Los Angeles) and the Chinese Academy of Medical Sciences and Peking Union Medical College, the Zika infection is now capable of causing serious birth defects and paralyzing disorders in its sufferer descending from a strain that spread in Asia before moving across the Pacific (www.nbcnews.com).

2.9 Diagnosis of Zika

2.9.1 Molecular Test for Zika Virus

Zika virus RNA detection at initial stages of illness is possible for symptomatic victims. For this, test serum is gathered in the initial two weeks following the manifestation inception. RNA NAT testing also possible on urine samples that have been gathered 14 days after appearance of indications. If the test comes out to be positive Zika virus infection is confirmed. However, if the test turns out to be negative then it does not guarantee the absence of infection. IgM antibody (serological) testing should be directed for facilitating investigation on serum. RNA NAT testing in case of asymptomatic pregnant women is done using serum and urine samples within 14 days of last exposure. RNA NAT testing is additionally suggested in case of pregnant women who are IgM positive. Standard IgM testing ought to be done in the event of pregnant women as a piece of obstetric care in the first and second trimester living in zones of dynamic Zika infection transmission (www.cdc.gov).

2.9.2 Trioplex Real-time RT-PCR Assay

Trioplex RT-PCR is used to detect to detect Zika, dengue or chikungunya infection, a research facility test named trioplex RT-PCR is accessible where the viral RNA is diagnosed. This test has not been affirmed by FDA and hence, can be utilized under an Emergency Use Authorization (www.cdc.gov).

2.9.3 Serologic Test for Zika Virus

IgM antibodies particular for Zika infection begin to develop in first week after infection. Positive IgM levels is seen on around fourth day after the onset of manifestations. Hence, to confirm the non-existence or to detect the presence of infection, this test is performed in case a negative result is obtained for RNA NAT (www.cdc.gov).

2.9.4 Zika MAC-ELISA

Zika MAC-ELISA kit is used to determine IgM antibodies in serum or cerebrospinal fluid to detect Zika infection. Sometimes, results might be hard to decipher because of cross-response with different flaviviruses and possible nonspecific reactivity.

2.10 Treatment and prevention for Zika

There is no particular medication or vaccine for Zika infection till now. The patient is instructed to take rest and drink a lot of fluids to avoid dehydration. Acetaminophen (Tylenol®) is one of the medicine which is prescribed to diminish fever and pain. As Zika infection is transmitted by mosquitoes, it is recommended to utilize bug repellent and wearing full sleeved-garments in infection prone regions. Moreover, recent investigations concluded that there is a chance of sexual transmission and hence, safe sex practices are encouraged to avoid the infection. In addition, Zika infection can likewise be transmitted through blood transfusion, care must be taken to evade transfusion of contaminated blood from the patient to the healthy individual (www.cdc.gov).

2.11 Vaccine Development against Zika

Since Zika virus infection epidemics are highly uncommon, innovations in vaccine development were not eagerly pursued. The recent outbreaks have alerted the scientists to the grim situation and hence, various measures to fight Zika infection are being considered. These include recombinant live attenuated vaccines, purified inactivated vaccines (PIVs), DNA vaccines, and viral vector vaccines. Bortezomib and mycophenolic acid are examples of some of the established anti-flaviviral drugs that have been found to inhibit Zika virus infection (Barrows *et al*., 2016). Zika virus infection increases caspase-3 activity which can be inhibited by emricasan which is a pan caspase inhibitor. Spread of infection is hindered by ten random inhibitors of cyclin-dependent kinases. Zika virus replication can likewise be repressed by Niclosamide, which is a category B anthelmintic medication affirmed by the US Food and Drug Administration (Xu *et al*., 2016).

Continuous endeavors are being made to fasten Zika virus vaccine development and assessment with a definitive objective of diminishing time to licensure. The National Institute of Allergy and Infectious Diseases (NIAID), along with the National Institutes of Health, have propelled a clinical

trial of a vaccine candidate which is proposed to avert Zika virus infection. The study conducted at early stages should evaluate the safety concern of and immunogenic potential in participants. Researchers at NIAID's Vaccine Research Center (VRC) have built up a DNA-based vaccine. This Zika virus vaccine candidate entered Phase 1 clinical trials in August 2016. NIAID's research facility has built up a live attenuated vaccine against Zika infection and dengue infection disease. This vaccine is as of now being assessed in a huge stage 3 think about in Brazil. Advancement of some mRNA based vaccines is also in progress by NIAID.

Walter Reed Army Institute of Research (WRAIR) has built up a sanitized inactivated Zika vaccine called ZPIV. WRAIR has utilized methodologies similar to the ones they used to design vaccines against Japanese Encephalitis and dengue infection. Phase 1 trials testing ZPIV have started at WRAIR. Hereditarily designed variant of vesicular stomatitis is being considered as a plausible Zika infection treatment target (www.niaid.nih.gov).

2.12 Peptide Vaccine

The idea of peptide vaccines is based on identification of T and B cell epitopes which are immunodominant and can induce specific immune responses. One of the approaches used to design peptide vacccines that combines the immunoinformatic prediction with rigorous experimental validation and hence, is considered as an interdisciplinary approach (Flower *et al*., 2013). Peptides that are used as vaccine candidates are usually 20-30 amino acids long. These peptide lengths are considered appropriate for the activation of cellular and humoral responses. B-cell antigenic determinant of a target molecule can also be combined with a T-cell epitope to obtain a vaccine candidate with higher immunogenicity. In epitope based vaccine, immune response induced by an immunogen is not directed against the whole molecule, but it is against a specific region (sequence) of protein which is an epitope. The primary epitope-based immunization was made in 1985. They presented recombinant DNA and communicated epitopes in *Escherichia coli* against cholera (Dermime *et al*., 2004) (Meloen *et al*., 2001).A recent study has identified conserved peptides containing multiple T-cell epitopes of matrix 1 protein in H1N1 flu using immunoinformatics and then validated immunogenic response in PBMC (Lohia *et al*., 2015, Lohia *et al*., 2017).Thus, development of epitope based vaccines for T and B lymphocytes is feasible. Moreover, peptide-based epitope ensemble vaccines offer various advantages over traditional

whole-organism vaccines. They solely focus on the immune response of relevant epitopes and hence, help in avoiding non desirable effects like autoimmunity. Another advantage of such vaccines is that they induce allergen specific tolerance (Hailemichael *et al*., 2013). These vaccines represent no danger of change or inversion and almost no risk of contamination by harmful or pathogenic substances. They are comparatively easy to deliver than a regular vaccine and are also chemically stable.

Peptide-based vaccines have proved to be a boon for the treatment of life threatening diseases like hepatitis C and cancer (Wang *et al*., 2013, Hailemichael *et al*., 2013). In a study, metadherin based peptide vaccine against cancer has been carried out based on immunoinformatics approach (Dhiman *et al*., 2016).Even though epitope-based vaccines can be long-range and broad spectrum immunizations, yet with the utilization of small molecules as peptides, an issue of low immunogenicity as compared to traditional vaccines is there.

2.13 Immunoinformatics:

Immunoinformatics (sometimes referred to as **computational immunology**) is a sub-branch of bioinformatics that makes use of data management and computational tools to solve immunological problems (Brusic *et al*., 2004).Conventional methods involve pathogen cultivation and subsequent protein extraction which is very tedious, cost ineffective and needs lots of time. Immunoinformatics investigation stresses on finding the potential B-and T-cell epitopes. This approach accelerates the time and brings down the cost required for research center examination of pathogen related gene products (Patronov *et al*., 2013). By utilizing these computational tools and data, an immunologist can analyze the sequence areas with potential sites for binding, which thus prompts the advancement of new vaccines. . The approach of reverse vaccinology is used to identify the potential immunogenic peptide regions in a pathogen (Tomar *et al*., 2010).

2.14 HLA Polymorphism and Vaccine Design

The human leucocyte antigen (HLA) is an array of genes which are arranged inside a long continuous region of DNA on chromosome 6 in humans. HLA genes are usually arranged into

areas encoding for three classes of molecules. Class I HLA (A, B and C) genes encode glycoprotein communicated on the surface of about every nucleated cell; the significant capacity of the class I HLA is introduction of intracellular antigens to T cytotoxic cells. Class II HLA genes (DP, DQ, DR, DM, DOA, DOB) encode glycoprotein communicated fundamentally on antigen-displaying cells, where they introduce prepared extracellular antigenic peptides to T-cells. In addition to other products Class III HLA genes encode for different emitted proteins that have resistant capacity, including components of the complement system and molecules which are involved in inflammation (Kuby Immunology, Chapter 6).

The loci constituting the HLA are exceptionally polymorphic; that is, numerous option types of the genes, or alleles, exist at every locus amongst the population. The genes of the HLA loci lie near one another. HLA class I gene contains three distinctive HLA loci A, B and C each of which is coding for α chain polypeptides of HLA class I molecule. The class II gene contains three noteworthy HLA loci, DP, DQ and DR; each of these loci is coding for α and β chain polypeptides of HLA class II molecule. Most people acquire the alleles encoded by these firmly connected loci as two sets, one from each parent. Each arrangement of alleles is usually referred to as a haplotype. The alleles are co predominantly communicated; that is, both maternal and fatherly gene products are expressed in same cells. The HLA polymorphism amongst species produces diversity of binding specificities and along these lines distinctive examples of responsiveness to antigens. So T cell epitopes which tie to different HLA molecule will be a successful vaccine targets as it can give better scope of the population.

T cells can identify antigen just when it is presented by HLA molecules. HLA class II is present on particular cell types which include proficient APCs while class I HLAs are found on each cell having nucleus. HLA class II proteins tie oligopeptide sections which are determined through the protein break down of pathogen antigens, and present them onto the surface of cell for recognition by CD4+ T cells. If triggered, T cells may unleash a versatile resistant reaction specifc to the pathogen. The HLA I molecule ties to a peptide of about 9 amino acids long inside a shut section. In contrast, because the antigen-binding groove has it’s both ends open, the HLA II molecules can therefore present much longer peptides, lying in the range of 12 to 25 amino acids, out of which nine occupy the binding groove. This contrast between HLA I and HLA II are critical for the advancement of these prediction algorithms (Larson *et al*., 2006).

2.15 Epitopes prediction algorithms

One of the key issues in T-cell epitope prediction is forecasting HLA binding of identified epitopes as it is viewed as a pre-requisite for T cell acknowledgment. All T-cell epitopes are good HLA binders, yet not all great HLA binders are T-cell epitopes. Deducing the peptide binding inclinations which have been exhibited by this substantial set of alleles is past the present limit of experimental techniques, requiring the expansion of prediction methodologies in bioinformatics. The best prediction techniques for T-cell epitopes created till date have been mostly data driven. Epitope prediction includes characterizing the peptide restricting specificity of particular class I or class II HLA alleles and after that foreseeing epitopes in silico. Such methods include motif-based systems, support vector machines (SVMs), and hidden Markov models (HMMs), quantitative structure–activity relationship (QSAR) analysis and structure-based approaches (Patronov *et al*., 2013). Utilizing sequence information of peptides, experimentally determined affinity data have been used as a part of the development of numerous T-cell epitope prediction algorithms. Contrasting to T cell epitope prediction calculations, the B cell epitope prediction has been more confounded, particularly for the conformational B cell epitopes in light of the fact that, the 3D-structure of protein should likewise be considered.

The prediction algorithms for linear B cells are like that of T-cells. The precision of essential arrangement based algorithm is quite low and altered calculations based on machine learning were subsequently developed, for instance, ABCpred (Saha *et al*., 2006) and BepiPred (Larsen *et al*., 2006) with significant improvements in their accuracy. Prediction algorithms for conformational B cell epitopes in light of 3D structure are additionally accessible to the steadily expanding 3D structure of antigen-antibody complex data. Some expectation servers in view of these calculations are available, for instance Disco Tope and CEP (http://bioinfo.ernet.in/cep.htm) (Kulkarni *et al*., 2005). These strategies make utilization of data conveyed in the structure of antibodies directed against proteins which are important to uncover the 3D collapsing of target proteins.

Chapter 3: OBJECTIVES

- Identification of conserved peptides containing CD4+ and CD8+ T cell epitopes of Zika virus envelope protein by employing various epitope prediction tools.
- Analysis of peptide-HLA interaction based on *in silico* approach.

Chapter 4: MATERIALS AND METHODS

4.1 Sequence

The Zika virus envelope protein consists of 504 amino acids. Envelope protein sequences of Zika virus available till March 2017 were downloaded from Zika virus database

(www.viprbrc.org).Complete protein sequences with Human as host were downloaded in FASTA format and saved in Microsoft office word. Partial or incomplete protein sequences were avoided.

4.1.2 Multiple sequence alignment

Multiple sequence alignment is done utilizing MUSCLE tool (http://www.ebi.ac.uk/Tools/msa/muscle).

MUSCLE acronym stands for Multiple Sequence Comparison by Log-Expectation. MUSCLE is a computer based program for making numerous arrangements of protein sequences. Components of the algorithm incorporate quick separation estimation utilizing kmer tallying, progressive alignment utilizing the log-desire score, and refinement utilizing tree dependent restricted partitioning. MUSCLE is asserted to accomplish both better accuracy and good speed over ClustalW2 or T-Coffee (Edgar RC., 2004).

The constraint of this tool is that it can only align 500 sequences of more than 350 amino acid length at a time. The number of sequence retrieved was more than 500 and hence, the tool was used multiple times. The results of MUSCLE were saved in FASTA format.

4.1.3 Conservancy Analysis

Conservancy analysis was carried out to find out the conserved region present in Envelope protein of Zika virus. The regions of envelope protein showing = 100% conservancy were selected. AVANA tool was used for conservancy analysis (http://avana.sourceforge.net/).

AVANA

The Antigen Variability Analyzer (AVANA) device utilizes entropy information to quantify fluctuation in protein sequence alignment. It additionally looks at arrangements utilizing common data, distinguishing the changes that describe particular sequence sets (http://avana.sourceforge.net/).Alignment result of MUSCLE (FASTA format) was used as input for AVANA software. Parameters were set to 100% conservancy and minimum length of 9 amino acids as a threshold value in AVANA. Conserved regions were searched in the alignments. The overlapping regions found in the conserved regions were merged together.

4.2 CD8+ T-cell Epitope prediction

CD8+ cells recognize the peptides presented by Class I HLA molecules on the cell surface which are either tumorous in nature or viral infected. The target cells are killed by the release of perforins and granzymes by CD8+ cells. For the identification of CD8+ specific epitopes, three different tools were used.

4.2.1 IEDB: It stands for Immune Epitope Database and analysis resource. It is a physically curated database of tentatively portrayed immune epitopes (Zhang *et al*., 2008).The IEDB contains information identified with antibody and T-cell epitopes for humans, non-human primates, rodents and other species. This tool anticipates IC50 values for peptides binding to particular MHC molecules. It takes in an amino acid sequence or set of sequences and decides every subsequence's capacity to bind to a particular MHC Class 1 molecule (tools.iedb.org).

4.2.2 SYFPEITHI

SYFPEITHI makes the predictions for both HLA class I and class II alleles. The algorithm identifies the anchor and auxiliary residues in the input sequence by comparing it with the database of naturally occurring epitopes. It assigns particular score to each residue depending upon their position and generates overall score for predicted epitope (Rammensee *et al*., 1999).

SYFPEITHI is a database containing more than 7000 peptide successions known to bind class 1 and class 2 MHC molecules (www.syfpeithi.de).The score is ascertained by assigning amino acids to some peptide a specific value depending on the fact if they are anchor, auxiliary anchor or preferred. So, here anchors are given 10 points; 6-8 points are given to unusual anchors, auxiliary anchors 4-6 and preferred deposits 1-4 points. The amino acids that are responsible for hindering the binding ability are assigned score between -1 and -3. The cut off value taken was 20 for the predicted epitopes (Dhiman *et al*., 2016).

4.2.3 netCTL 1.2

netCTL 1.2 prompts the epitope prediction on the premise of C-terminal proteasomal cleavage, TAP transport productivity and HLA binding. It identifies the epitopes for 12 HLA-A and HLA-B super types (Larsen *et al*., 2007). The cut off value >0.75 was used for the epitope prediction. The cut off 0.75 signifies the 80% sensitivity and 0.97 specificity score of the predicted epitopes. The three combined scores result in the identification of the potential immunogenic epitopes from the given sequence.

The common epitopes predicted by all the three tools were selected. The selected epitopes were overlapped to generate the peptide fragments which consisted of multiple epitopes.

4.3 CD4+ T-cell Epitope prediction

HLA class II molecules usually bind peptides and present them to CD4+ T cells. Peptides of class II HLA–peptide complexes are generally 13–18 amino acid residues long, longer than the nonameric peptides of class I molecules but the core sequence is still 9 amino acid in length. The peptide-binding split in class II molecules is open at both closures, enabling

longer peptides to stretch out past the ends. To recognize the CD4+ particular epitopes three unique tools v.i.z., MHC2pred (Joachim's, 1999; Cristianini and Shawe-Taylor, 2000), Propred (Singh *et al*., 2001) and IEDB agreement (Wang *et al*., 2008) were utilized.

4.3.1 MHC2pred

MHC2Pred is a support vector machine (SVM) based strategy for prediction of unbridled binders for 42 MHC class II alleles. The binders and non-binders for all the alleles have been obtained from MHCBN and JenPep database (Bhasin *et al*., 2003 and Blythe *et al*., 2002).All the peptides which haveIC50 value under 500nm are considered as binders and peptides with IC50 value more prominent than 500nm are considered as non-binders. Peptides containing fewer than 9 amino acids have been erased from the dataset. The binding core of 9 amino acids has been obtained from the binders of variable length without considering MHC binding motifs utilizing Matrix Optimization Techniques (MOT). For advancement of MHC binder prediction method, a machine learning procedure SVM (Joachim's, 1999; Cristianini and Shawe-Taylor, 2000) has been used.SVM has been prepared on the parallel contribution of single amino sequence.

4.3.2 Propred

ProPred is a graphical online server for predicting HLA class II binding areas in antigenic protein sequences. The server actualizes Quantitative Matrix based prediction calculation, utilizing amino-acid position/position coefficient table concluded from literature. The predicted binders can be pictured either as peaks in graphical interface or as shaded deposits in HTML interface (Singh H. *et al*., 2001). ProPred measures the percentage score with respect to the best score for that particular allele. The threshold value for the epitope prediction was taken as 3%. This means that only those peptides shall be obtained in results which are compatriots to the 3% best scoring natural peptides.

4.3.3 IEDB consensus

The IEDB consensus enables the user to apply multiple algorithms such as NN-align, SMM-Align and combinatorial method for the prediction of epitopes. The performance for HLA-II binding methods are in order: Consensus > netMHCIIpan > NN-align > SMM-align > Combinatorial Library. Based upon the epitope predicted by different algorithms a percentile rank is generated for the finally predicted epitopes (Wang *et al*., 2008). The threshold values taken for IEDB consensus was IC 50 value of less than 500.

The common epitopes predicted by all the three tools were selected.

4.4 Blast Screening

In order to avoid any similarity of the peptides with functional human protein, BLAST analysis (Altschul *et al*., 1990) was performed for epitopes predicted to bind HLA class I and II respectively. The peptides showing similarity in seven consecutive amino acids without gap or mismatch were eliminated (Tan *et al*., 2012) thus ruling out any possibility of autoimmune response against any human functional protein. The peptides screened after BLAST showing overlaps were merged together to generate peptide fragments.

4.5 Population Coverage Analysis

HLA alleles are highly polymorphic and diversity of alleles is found in the global population (Zúñiga *et al*., 2013). Different individuals of the worldwide population may respond in different ways against particular antigen. The IEDB database enables predicting the possible world population capable of responding to particular immunogen. The database contains variety of class I and II HLA alleles frequently found in different world population of geographical areas. It compares the epitopes specific alleles with the alleles in database and find out percentage population capable of responding. In our study IEDB population coverage analysis tool was carried out for 16 different geographical areas distributed globally (Huynh-Hoa Bui *et al.*, 2006).

4.6 Mapping of peptide fragments

The Discovery Studio v3.5 visualizer tool was used to locate the selected peptide fragments on the three-dimensional structure of ZIKA virus envelope protein. The protein structures for envelop protein available in the PDB database were used as the model structures to visualize the predicted peptide fragments.

4.7 Prediction of peptide allergenicity and toxicity

Allergenicity of the predicted peptides was determined using the online tool AlgPred, which is based on screening IgE epitopes in query protein sequence and Motif Alignment & Search Tool (Saha *et al*., 2006). An expressed protein/peptide is designated as a potential allergen if it possesses 35% identity using a window of 80 amino acids or has six contiguous amino acids with an established allergen (www.fao.org.).Therapeutic peptides are often associated with toxicity towards eukaryotic cells.

ToxinPred is an *in silico* tool that helps in screening and identifying toxic peptides in a dataset of non-toxic or random peptides obtained from SwissProt and the translation of EMBL nucleotide sequence databases (Gupta *et al*., 2013).

4.8 HLA-Peptide interaction analysis via Docking

In computer aided drug designing and in structural molecular biology molecular docking has been found to play a very important role. Docking enables the prediction of the predominant binding mode of a ligand with a protein of known three-dimensional structure. High dimensional space searching is used in some of the successful docking methods and it also makes use of a scoring function that correctly ranks candidate dockings (Morris *et al*., 2008).

4.8.1 Structure generation of peptides/epitopes

The structure of predicted epitopes was elucidated using peptide structure prediction server PEP-FOLD (Thevenet *et al*., 2012). The models generated by PEPFOLD are assorted either by using the coarse grain energy of the PEPFOLD or by predicted Tm score. For the peptides of length 36 residues the coarse grain energy of the PEPFOLD is used to sort the predicted models whereas Tm scores for the Peptides with the residues more than 36.

4.8.2 Separation of naturally bound peptide from HLA molecule

The structure of twelve HLA each of class I and II bound to their natural peptide were retrieved. The native peptides (Ligands) and HLA structure (Receptor) were separated using the discovery studio visualization tool 3.5. The generated peptide free HLA structures were further used for docking using Auto dock Vina (Trott & olson ., 2010) tool.

4.8.3 Docking using Auto Dock Vina

The isolated native peptides were redocked with their respective HLA via Auto Dock vina (Trott *et al*., 2009) tool. The binding energy of these naturally bound ligands was determined. The anticipated epitopes were docked with the HLA class I and II structure and their binding energy was compared with the binding energy of the redocked naturally bound peptides.

Briefly, the process followed for carrying out the above analysis was as follows:-

Ligand Separation

Open the PDB structure of HLA molecule in discovery studio visualization tool 3.5

↓

Go to the scripts option and select the water molecules

↓

Go to edit and delete

↓

Go to scripts and select protein chains

↓

Go to edit and delete

↓

Save as ligand.pdb file

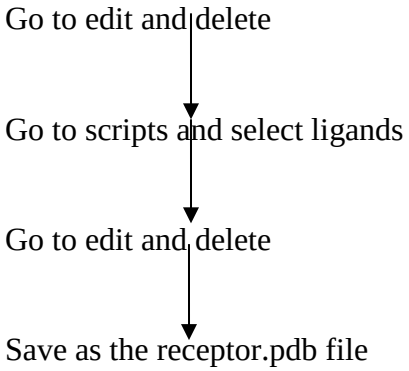
Receptor Separation

Open the PDB structure of HLA molecule in discovery studio visualization tool 3.5

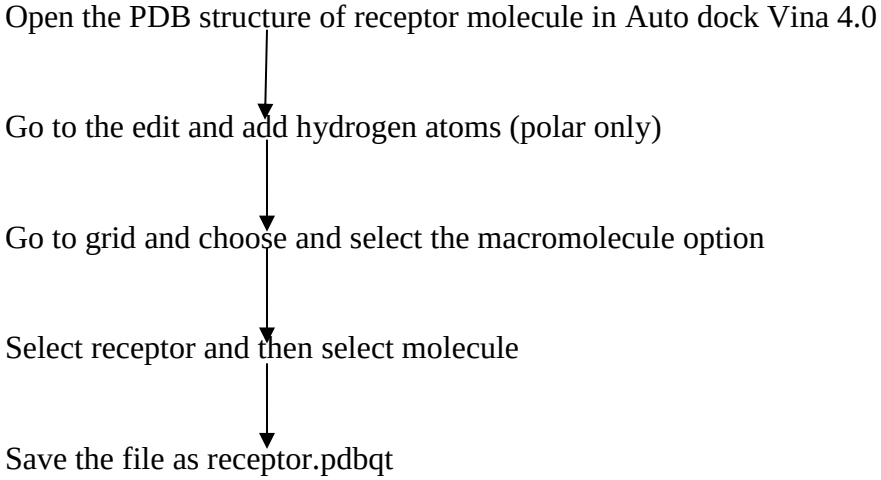
↓

Go to the scripts option and select the water molecules

↓

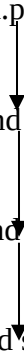


Receptor preparation for docking



Ligand preparation for docking

Open the ligand.pdb file in the Auto dock Vina 4.0



Go to Ligand and choose the molecule

Select the Ligand molecule

Go to output and save the molecule as ligand.pdbqt

4.8.3 Preparation of configuration file

To perform the docking in auto dock vina, we are supposed to generate a text file named as config containing the following commands. The config file should be created for the each molecule separately depending upon their grid box selection.

receptor = “receptor name”.pdbqt
ligand = “ligand name”.pdbqt
out = out.pdbqt
center_x = “x” (molecule specific value obtained after grid box optimization)
center_y = “y” (molecule specific value obtained after grid box optimization)
center_z =“z” (molecule specific value obtained after grid box optimization)
size_x = “x” (molecule specific value obtained after grid box optimization)
size_y = “y” (molecule specific value obtained after grid box optimization)
size_z = “z” (molecule specific value obtained after grid box optimization)
Exhaustiveness = 16.

Chapter 5 RESULTS

5.1 Conserved regions of Envelope protein in Zika Virus

Fifty envelope protein sequences of different genetic variants of Zika virus infecting human from 1947 were obtained from the Zika virus database. MUSCLE (alignment) and AVANA (conservation analysis) tool were done to find the conserved regions. Fourteen peptide fragments ranging from 9 to 85 amino acids were found to be 100% conserved (Table 2). These fragments were then considered for further epitope prediction.

Table 2: Conserved peptide fragments of Zika virus envelope protein

	Conserved sequences	**Length**
CSE1	NRDFVEGMSGGTWVD	15
CSE2	VLEHGGCVT	9
CSE3	ASDSRCPTQGEAYLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAK FACSKKMTGKSIQPENLEYRIMLSVHGSQHSGM	85
CSE4	VNDTGHETDENRAKVEITPNSPRAEA	27
CSE5	HKEWFHDIPLPWH	22
CSE6	TGTPHWNNKEALVEFK	16
CSE7	QEGAVHTALAGALEAEMDG	19
CSE8	LSSGHLKCRLKMDKLRLKGVSYSLC	25
CSE9	ENSKMMLELDPPFGDSYIVIG	21
CSE10	GEKKITHHW	9
CSE11	RSGSTIGKAFEATVRGA	17
CSE12	RMAVLGDTAWDFGSVGG	17
CSE13	GAAFKSLFGGMSWFSQ	16
CSE14	CLALGGVLIFLST	13

5.2. CD8+ T-cell specific epitope identification

Cytotoxic T lymphocytes ($CD8^+$ T cells) are important for immune defense against intracellular pathogens including virus and bacteria and for tumor surveillance. Three tools used for the prediction of HLA class I restricted epitopes were netCTL, IEDB and SYFPEITHI. Twenty epitopes considered which were commonly predicted by all three tools (Table 3). Blast analysis revealed that none of the predicted epitopes were homologous to human protein. The epitopes that had overlapping amino acid sequences were merged to form a single peptide fragment. Seven overlapping peptide fragments were obtained which contained more than one $CD8^+$ T-cell epitope (Table 4).

Table 3:$CD8^+$ T cell epitopes

CD8+T cell epitopes	
KEWFHDIPL	SIQPENLEY
VHTALAGAL	QSDTQYVCK
QEGAVHTAL	GLFGKGSLV
RLKGVSYSL	CPTQGEAYL
KLRLKGVSY	TQYVCKRTL
GEKKITHHW	LEYRIMLSV
ALGGVLIFL	DTAWDFGSV
LALGGVLIF	VLGDTAWDF
ETDENRAKV	SLFGGMSWF
TPNSPRAEA	KSLFGGMSW

Table 4: Peptides containing overlapping $CD8^+$ T cell epitopes

$CD8^+$ T cell peptide	**Position**	**Number of epitopes**
QEGAVHTALAGAL	1-13	VHTALAGAL,QEGAVHTAL
KLRLKGVSYSL	15-25	RLKGVSYSL,KLRLKGVSY
LALGGVLIFL	2-11	ALGGVLIFL,LALGGVLIF
SIQPENLEYRIMLSV	61-75	SIQPENLEY,LEYRIMLSV
QSDTQYVCKRTL	17-28	QSDTQYVCK,TQYVCKRTL
VLGDTAWDFGSV	4-15	DTAWDFGSV,VLGDTAWDF
KSLFGGMSWF	5-14	SLFGGMSWF,KSLFGGMSW

5.2. $CD4^+$ T-cell specific epitope identification

$CD4^+$T-cells are very important as they play a major role in immune defense by secretion of specific cytokines. Once $CD4^+$ T-cells activated, they divide rapidly to activate other cells of the immune system like $CD8^+$ T-cells and B cells. Three tools used for prediction of HLA class II restricted epitopes were MHC2pred, Propred and IEDB. Nineteen epitopes were considered which were commonly predicted by all three tools (Table 5). Blast analysis revealed that none of the predicted epitopes are homologous to human protein. Five peptide fragments were generated by merging overlapping epitopes (Table 6).

Table 5:$CD4^+$ T cell epitopes

$CD4^+$ specific T cell epitopes	
VEITPNSPR	FGGMSWFSQ
ITPNSPRAE	YRIMLSVHG
WNNKEALVE	LVTCAKFAC
LRLKGVSYS	FACSKKMTG
LKCRLKMDK	IMLSVHGSQ
LKMDKLRLK	LSVHGSQHS
LELDPPFGD	YVCKRTLVD
VLGDTAWDF	LFGKGSLVT
FKSLFGGMS	LEYRIMLSV
LFGGMSWFS	

Table 6: Peptides containing overlapping $CD4^+$ T cell epitopes

$CD4^+$ T cell peptide	**Position**	**Number of epitopes**
VEITPNSPRAE	15-25	VEITPNSPR,ITPNSPRAE
LKCRLKMDKLRLK	6-18	LRLKGVSYS,LKCRLKMDK,LKMDKLRLK
FKSLFGGMSWFSQ	4-16	FKSLFGGMS,LFGGMSWFS,FGGMSWFSQ
LEYRIMLSVHGSQHS	69-81	LEYRIMLSV,YRIMLSVHG,LSVHGSQHS,IMLSVHGSQ
LVTCAKFACSKKMTG	45-59	LVTCAKFAC,FACSKKMTG

5.3. Identification of Peptide containing $CD4^+$ and $CD8^+$ specific epitope

It was interesting to find peptides that contained both $CD4^+$ as well as $CD8^+$ T-cell epitopes. To generate such peptides having overlapping $CD4^+$ and $CD8^+$ T cell, all the predicted $CD4^+$ and $CD8^+$ T cell epitopes were carefully looked. This resulted in generation of seven conserved peptide fragments that contained both $CD4^+$ and $CD8^+$ T-cell epitopes (Table 7). These peptide fragments (PE1-PE7) consist of twelve$CD8^+$T cell epitopes and fifteen $CD4^+$ T-cell epitopes (table 7). These selected peptides were predicted to bind a large number of different HLA class I and II alleles (Table 7).

Table 7: Envelope protein containing peptides enriched with $CD8^+$ and $CD4^+$specific T cell epitopes

$CD8^+$ T cell Peptides	Number		$CD4^+$ T cell Peptides	Number		Peptides containing multiple $CD8^+$& CD^+4 T cell epitopes
	Epitope	HLA I		Epitope	HLAII	
TPNSPRAEA	1	3	VEITPNSPRAE	2	23	VEITPNSPRAEA (PE1)
KLRLKGVSYSL	2	24	LKCRLKMDKLRLKGVSYS	3	134	LKCRLKMDKLRLKGVSYSL (PE2)
VLGDTAWDFGSV	2	8	VLGDTAWDF	1	18	VLGDTAWDFGSV (PE3)
KSLFGGMSWF	2	14	FKSLFGGMSWFSQ	3	41	FKSLFGGMSWFSQ (PE4)
SIQPENLEYRIMLSV	2	14	LEYRIMLSVHGSQHS	4	122	SIQPENLEYRIMLSVHGSQHS (PE5)
QSDTQYVCKRTL	2	7	YVCKRTLVD	1	22	QSDTQYVCKRTLVD (PE6)
GLFGKGSLV	1	5	LFGKGSLVT	1	36	GLFGKGSLVT (PE7)

5.4. Population coverage and HLA distribution Analysis

HLA distribution varies among different ethnic groups around the globe. Population coverage analysis plays an important role for the design of peptide-based vaccine. Due to highly polymorphic nature of HLA, the response against a particular antigen throughout a human population is restricted. Hence, T cell epitope which binds to several HLA alleles is preferred for maximum population coverage. The peptide fragments were found to cover larger number of HLA class I and II alleles. In order to estimate the immunogenic potential of the predicted peptide fragments on the global population, population coverage analysis was carried out using the IEDB population coverage tool. IEDB tool calculated the expected fraction of individuals from Asia, Europe, Africa, North America, South America and Oceania countries that were expected to respond to these peptide fragments. Population coverage was expressed as an average in case of Asia (East Asia, Northeast Asia, South Asia, Southeast Asia, and Southwest Asia), Africa (East Africa, West Africa, Central Africa, North Africa, South Africa) and North America (North America, Central America).Peptides PE4 and PE5 are expected to give a high immunogenic response ranging from 90-100% whereas the response of peptide PE1 and PE2 were in the range of 80-90% in all the populations under consideration. The average population coverage for immunogenic response of predicted peptides were found out to be 76.37% and 98.13% for class I and II HLA specific respectively.

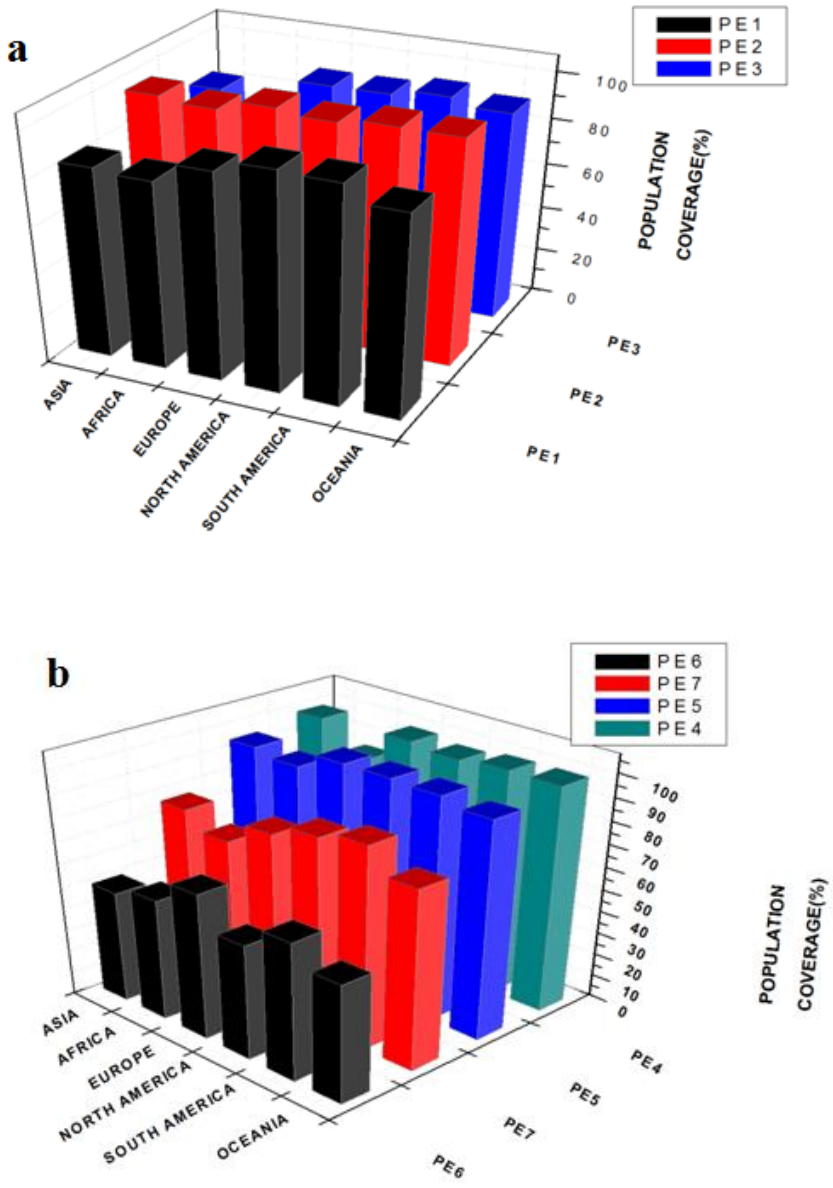


Figure 4: Population coverage analysis of the peptides (a) PE1, PE2 and PE3 (b) PE4, PE5, PE6 and PE7.

5.5. Structural analysis

The envelope protein structure was obtained from PDB and it was having a multimeric structure. Therefore, the PDB structure of the protein was edited to remove the multimeric structures leaving only monomers using Discovery Studio v4.1 visualizer. Five out of seven identified peptides containing $CD4^+$ and $CD8^+$ T cell epitopes were mapped on the three dimensional structure of envelope protein. The peptides PE3 and PE4 could not be labeled as they were located in ribonucleoprotein binding domain of envelope protein hence could not be mapped.

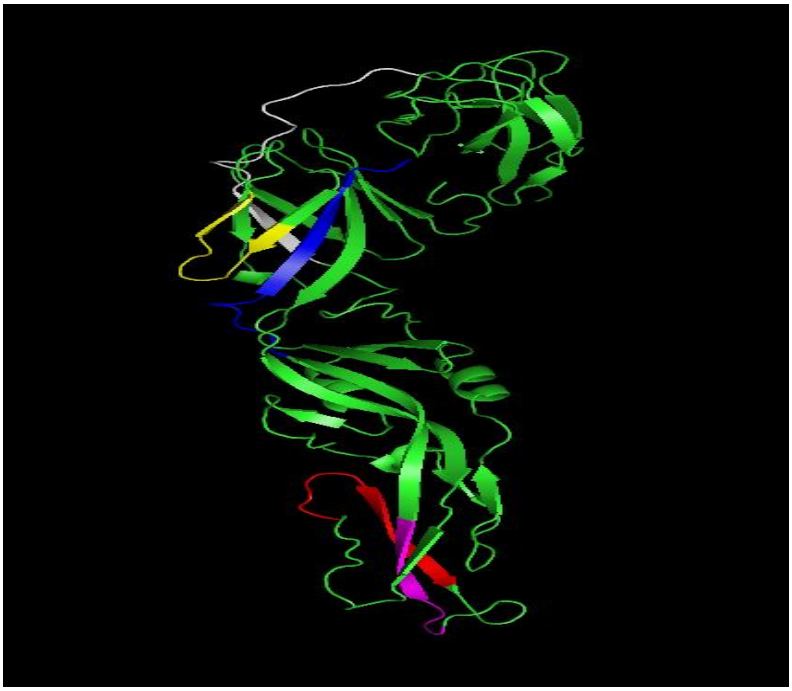


Figure 5: Mapping of selected peptides on three dimensional structure of envelope protein Peptide fragmentsPE1, PE2, PE5, PE6 and PE7are depicted in yellow, grey, blue, red and magenta color.

5.6. Epitope Docking with Class I and II HLA molecules

The immunogenic potential of the predicted epitopes depends upon the fact that how efficiently it can bind to peptide binding pocket of the HLA molecule. The auto dock vina tool has been used to assess the binding affinity of the predicted epitope with HLA class I and II molecules. Twelve peptide structure HLA complex each for class I and II were retrieved from PDB database for docking purpose. PEPFOLD was used for prediction of all seven peptide fragments containing $CD4^+$ specific T-cell epitopes as well as for 12 epitopes specific for $CD8^+$ T cells (Table 7). Pepfold generate structure of the given sequence based upon structure alphabet letters. It predicts the structure of four consecutive amino acids and then combines the series of structural alphabet letter to generate the structure (Thevenet *et al.*, 2012). The best model generated was taken and subjected for epitope docking.

Peptides were extracted from the HLA complex and then redocked with corresponding HLA molecules, using auto dock vina. Twelve epitopes were docked with HLA Class I molecule while seven$CD4^+$ T cell peptides were docked with HLA Class II molecule. The binding energies of the peptides capable of binding to HLA molecule were compared with the naturally bound peptide. The individual binding energies of different epitopes after docking with class I HLA molecule were found to be comparable to the binding energy of the naturally bound peptides. However, some peptide-HLA docked complexes were found to have higher binding energy than native peptides. $CD8^+$T cell epitope DTAWDFGSV for HLA 4JQV and $CD4^+$peptideVLGDTAWDF for all 12 HLA II alleles of PE3 peptide have shown higher binding energies than native peptide (Figure 6 and Figure 7). Similarly, $CD8^+$ T cell epitope QSDTQYVCK(HLA 3RL1) and $CD4^+$ peptides YVCKRTLVD(HLA 1A6A,1BX2,1D5M,1D5Z,1KLU,1UVQ,2FSE,2NNA and 3L6F) of PE6 peptide have shown higher binding energy than native peptide. GLFGKGSLV of PE7 was not able to bind to HLA 3SPV, thus their binding energy is not shown in figure 6. Similarly, VEITPNSPRAE, YRIMLSVHGSQHS, YVCKRTLVD of peptides PE1, PE5 and PE6 were not able to bind to HLA 1UVQ. Best docking pose of selected HLA-peptides are shown in figure.

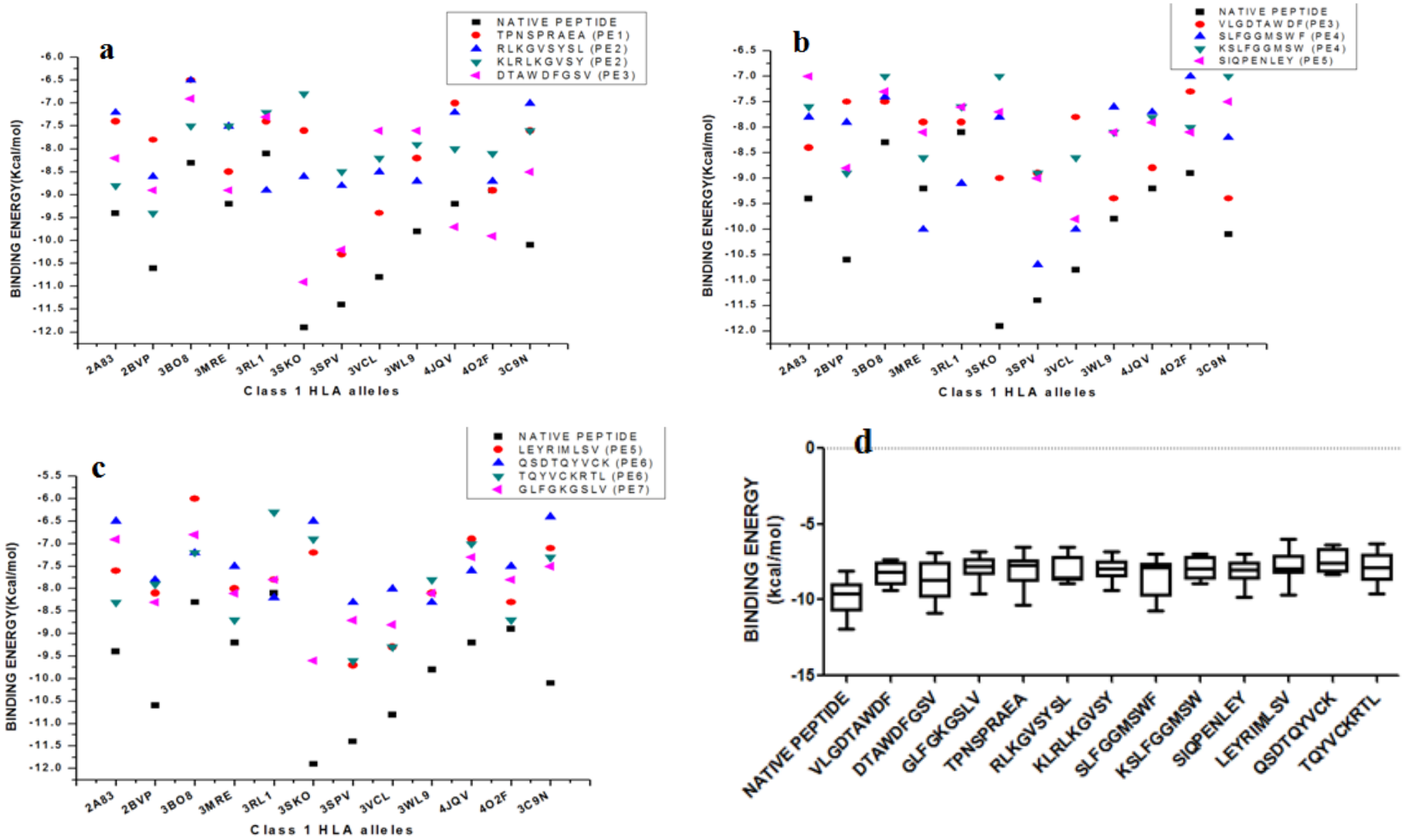
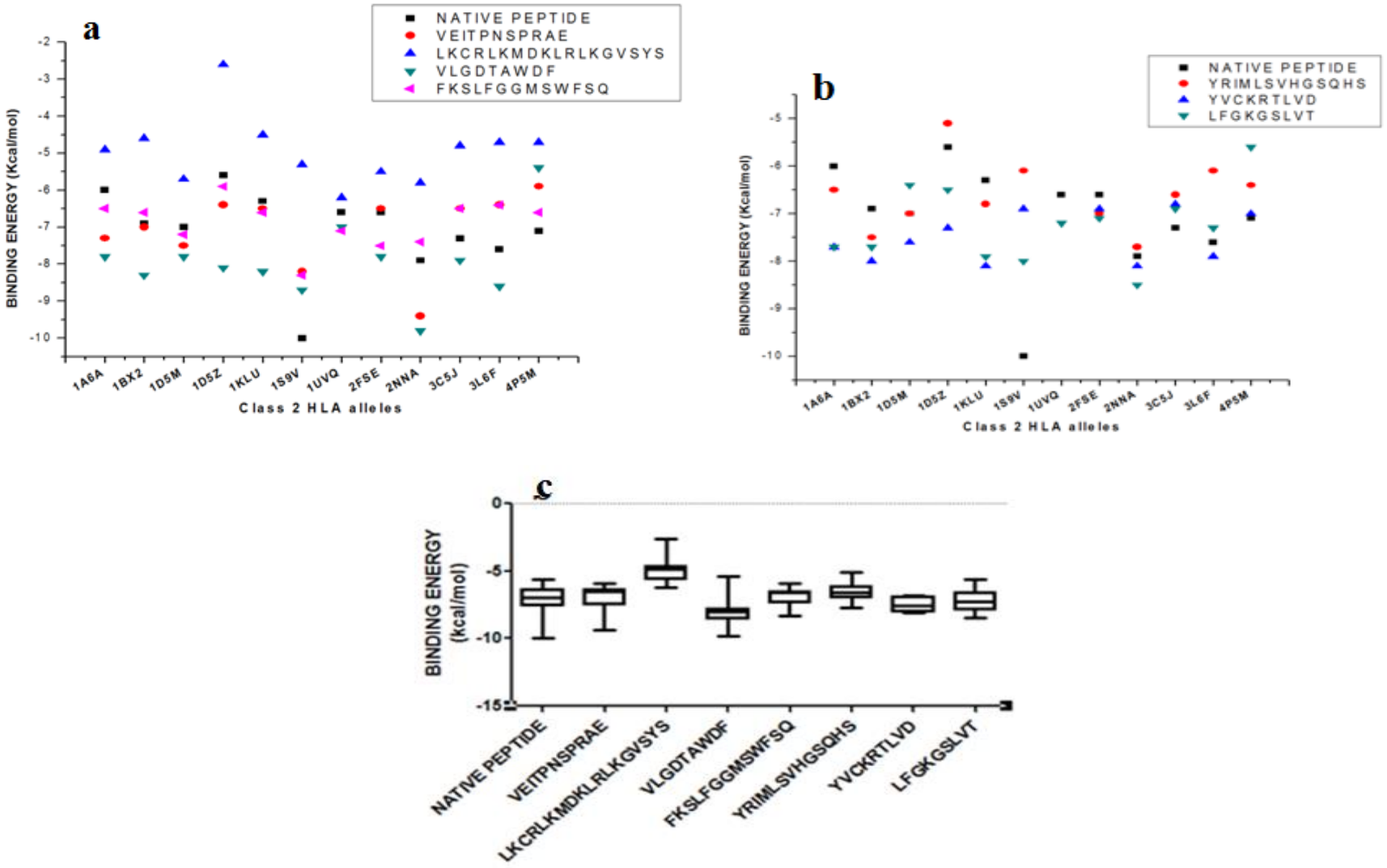


Figure 6: Docking of CD8^{+}T cell epitopes of Envelope protein with HLA class I molecules. Binding energy of naturally bound peptides and epitopes of peptides(a) PE1, PE2, PE3, (b) PE3, PE4, PE5, (c) PE5, PE6, PE7, (d) Combined binding energy of each epitope with twelve HLA molecules**.** The bars represent the range of energy with a horizontal line at the median. Whiskers extend from the smallest value up to the largest.



F**igure 7**: Docking of peptides containing CD4$^+$ T cell epitopes of Envelope protein with HLA class II molecules. Binding energy obtained for peptide fragments of peptide (a) PE1, PE2, PE3, PE4 (b) PE5, PE6, PE7 (c) Combined binding energy of each epitope with twelve HLA molecules. The bars represent the range of energy with a horizontal line at the median. Whiskers extend from the smallest value up to the largest. Most of the peptides have their energies close to the native peptide energy range.

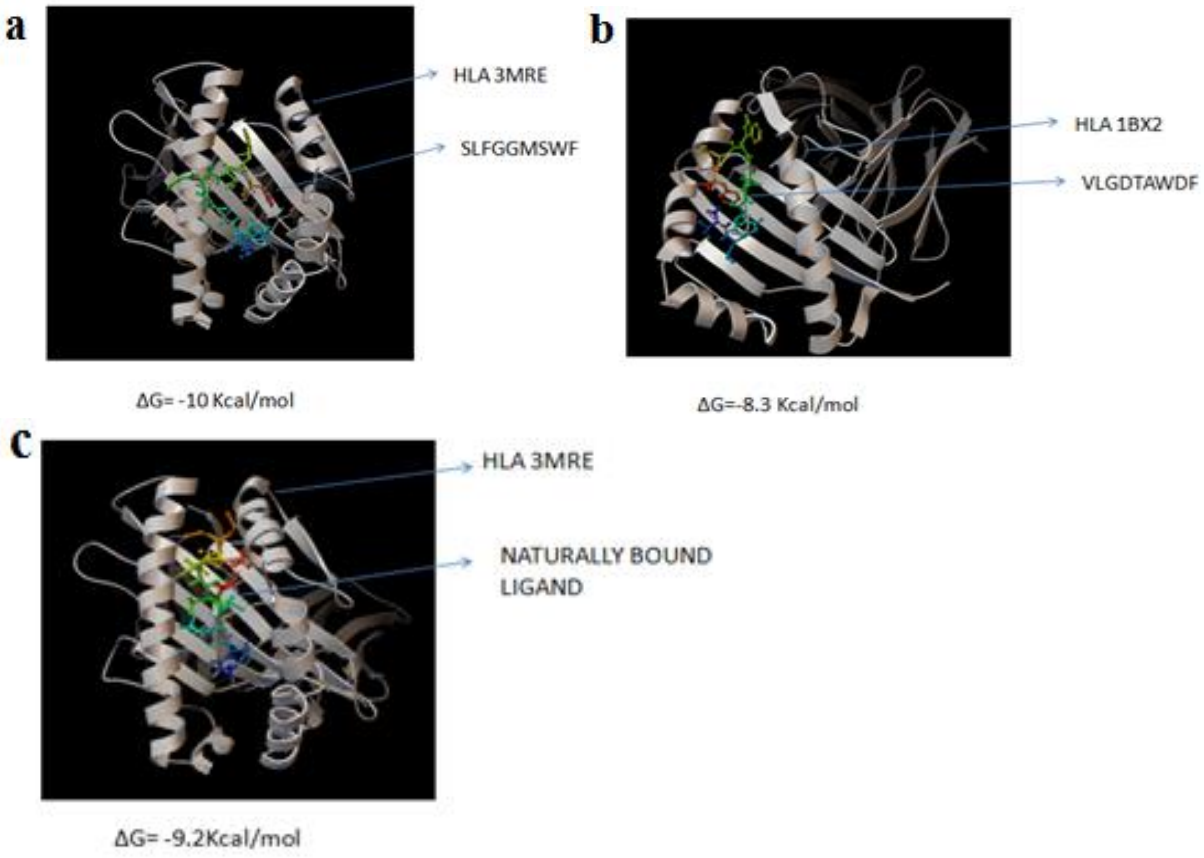


Figure 8: Poses of dockings showing highest binding energy obtained for (a) $CD8^+$ T cell epitope of envelope protein (b) CD4+ T cell peptide of envelope protein (c) naturally bound peptide

5.7. Allergenicity and toxicity analysis of tested peptides

Most of the peptides derived from the envelope protein of Zika virus are considered to be good candidates for development of a universal influenza vaccine. As essential aspect of vaccine development is assessment of allergenicity and toxicity of the vaccine candidates. Allergenicity and toxicity analysis of the peptides conducted showed that one of the peptide VLGDTAWDFGSV was found to be allergen and rest all peptides were non-allergen, none of the peptides under consideration were found to possess any toxicity.

5.8. Conservation analysis of peptides in other viruses infecting human

Blast analysis was conducted in which conservation of these predicted peptides was determined in other viruses infecting humans. Analysis suggested that peptide PE3 was having 100% conservation in Spondweni, Iiheus and St. Louis encephlatitis viruses. Peptides PE2, PE5, PE6 and PE7 were found to be 84%, 67%, 75% and 100% conserved in Dengue virus.

Table 8: Conservation of peptides in other virus infecting human

Peptide Fragments	Dengue virus	Spondweni virus	Iiheus virus	St. Louis encephlatitis virus	Kedougou Virus	West nile Virus
VEITPNSPRAEA (PE1)	-	-	-	-	-	-
LKCRLKMDKLRLKGVSYSL (PE2)	84%	84%	-	-	84%	-
VLGDTAWDFGSV (PE3)	-	100%	100%	100%	-	-
FKSLFGGMSWFSQ (PE4)	-	85%	-	-	-	85%
SIQPENLEYRIMLSVHGSQHS (PE5)	67%	80%	-	-	-	-
QSDTQYVCKRTLVD (PE6)	75%	-	-	-	-	-
GLFGKGSLVT (PE7)	100%	-	-	-	-	-

Chapter 6
DISCUSSION and CONCLUSION

Discussions

Zika virus is the causative agent of the mild febrile disease in humans. Current treatment for Zika is only symptomatic as no proven treatment is yet available. There are no licensed vaccines against Zika, although some potential candidates are under clinical trials.

Immunoinformatics driven vaccine design research is emerging as a new way to lessen time in vaccine design process**.** It bypasses the exhaustive and expensive conventional vaccine design approach and uses the *in silico* vaccine design algorithms providing a practical alternative. The conventional vaccine design approaches always has a risk of reverting back and becoming live again. The current study considered 50 envelope protein sequences of from all available strains of Zika virus (1947 to March 2017) known to infect human population. Zika virus protein and 100% conserved were taken for defining epitope. Peptides enriched epitopes predicted by various bioinformatics tools have proven to be immunogenic both *in vivo* and *in vitro* assays. There have only been very few studies which have aimed at the prediction of epitopes in Zika through *in silico* approach. Majority of them have used only one tool (Gritzapis *et al*., 2010) or two tools (Yuan *et al*., 2012) for epitope prediction. In contrast to this, the present study identification of epitopes is based on six different prediction algorithms that strengthened the potential of putative epitopes for immunogenicity. Considering a consensus approach has remarkable advantages which has discussed in earlier study (Lohia *et al*., 2014). Also to avoid any chance of auto immune response and eliminating the chances of tolerance by the predicted peptides BLAST analysis is of great importance. The BLAST screening suggested that there was no similarity between any of the conserved peptides and human protein. Seven non-self-peptides were identified which are rich in multiple helper and cytotoxic T cells.

HLA alleles are highly polymorphic and diversity of alleles was found in the global population. Different individuals of the worldwide population may respond in different way against particular antigen. Epitope prediction results have shown that all seven selected peptides were associated to large number of HLA alleles. The IEDB analysis enables predicting the possible world population capable of responding to particular immunogen. The database contains variety of HLA class I and

II alleles frequently found in different world population of geographical areas. In our study IEDB population coverage analysis tool was carried out for 16 different geographical areas distributed globally (Huynh-Hoa Bui *et al.*, 2006). Interestingly, peptide fragment LKCRLKMDKLRLKGVSYSL on an average revealed 97.58% population coverage in Asia, Africa, Europe, North America, South America and Oceania which is the highest among all the peptides. These results show that these peptide fragments are capable of inducing an immunogenic response among individuals belonging to different regions of the world population. Molecular docking is another approach to give insight on peptide-HLA interaction. All the twelve $CD8^+$ T cell epitopes and 7 peptides fragments containing $CD4^+$ T cell epitopes which are part of selected peptides were considered for docking studies. All the bound peptides showed comparable binding energies with the natural bound peptide. Similar kinds of results were found in our previous studies of metadherin protein and matrix protein of HIN1 virus (Lohia *et al*., 2015, Dhiman *et al*., 2016).

These peptides were evaluated by *in-silico* tool to have allergic response and toxicity. One peptide was found to be allergic and other were non-allergic and non-toxic in nature. Further, these peptides were looked for conservation in other virus infecting humans. Interestingly, some of the peptides were found to be conserved in different viruses such as Dengue virus, Spondweni virus, Iiheus virus, St. Louis encephlatitis virus, Kedougou virus and West nile virus. Thus, these potential vaccine candidates may be effective in other viruses as well.

Conclusion

To conclude, six conserved peptides which contain multiple HLA-restricted T cell epitopes were obtained on employment of *in-silico* tools. Peptide-HLA interaction revealed that these have the capacity to induce immune response among the different populations of the world and have the potential to bind diverse HLA alleles. These selected peptides are non-self, non-allergic and non-toxic in nature. Thus, these peptides can further be evaluated in *in-vitro* and *in-vivo* response to confirm the immunogenic nature and may be considered as potential vaccine target against ZIKA virus.

REFERENCES

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool . J Mol Biol. 1990;215(3):403-10.

2. Barrows NJ, Campos RK, Powell ST, Prasanth KR, Schott-Lerner G, Soto-Acosta R, Galarza-Muñoz G, McGrath EL, Urrabaz-Garza R, Gao J, Wu P, Menon R, Saade G, Fernandez-Salas I, Rossi SL, Vasilakis N, Routh A, Bradrick SS, Garcia-Blanco MA. A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. Cell Host Microbe. 2016;20(2):259-70.

3. Bazan JF and Fletterick RJ. Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. Virology. 1989;171:637-639.
4. Bhasin M, Singh H, Raghava GP. MHCBN: a comprehensive database of MHC binding and non-binding peptides. Bioinformatics . 2003;19(5):665-6.

5. Blythe MJ, Doytchinova IA and Flower DR. JenPep: a database of quantitative functional peptide data for immunology. Bioinformatics. 2002;18:434–439.

6. Bollati M, Alvarez K, Assenberg R, Baronti C, Canard B, Cook S, Coutard B, Decroly E, de Lamballerie X, Gould EA, Grard G, Grimes JM, Hilgenfeld R,Jansson AM, Malet H, Mancini EJ, Mastrangelo E, Mattevi A, Milani M, Moureau G, Neyts J, Owens RJ, Ren J, Selisko B, Speroni S, Steuber H, Stuart DI,Unge T, Bolognesi M. Structure and functionality in flavivirus NS-proteins: perspectives for drug design. Antiviral Res. 2010;87(2):125-48.
7. Brunsvig P, Dueland S, Moller M, Gjertsen MK, Trachsel S, Aamdal S, Gaudernack G. Telomerase peptide vaccination: a phase I/II study in patients with non-small cell lung cancer. Cancer Immunol Immunother. 2006;1512:1553–1564.
8. Brusic V, Bajic VB, Petrovsky N. Computational methods for prediction of T-cell epitopes--a framework for modelling, testing,and applications. Methods. 2004;34(4):436-43.

9. Buckley A, Gould EA. Detection of virus-specific antigen in the nuclei or nucleoli of cells infected with Zika or Langat virus. J Gen Virol. 1988;69:1913–20.
10. Burges C , Smola A. Advances in Kernel methods—support vector learning, Cambridge, MA, London, England MIIT Press, pp. 169–84.
11. Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organization, expression and replication. Annu. Rev. Microbiol. 1990;44:649-88.
12. Chambers TJ, McCourt DW, and Rice CM. Yellow fever virus proteins NS2a, NS2b, and NS4b: identification and partial N-terminal amino acid sequence analysis. Virology. 1989; 169:100-109.
13. Cristianini N and Shawe J –Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press. 2000.
14. Dermime S, Gilham DE, Shaw DM, Davidson EJ, Meziane el-K, Armstrong A, Hawkins RE, Stern PL. Vaccine and antibody-directed T cell tumour immunotherapy. Biochim Biophys Acta. 2004;1704(1):11-35.
15. Dhiman G, Lohia N, Jain S and Baranwal M. Metadherin peptides containing CD4þ and CD8þ T cell epitopes as a therapeutic vaccine candidate against cancer. Microbiol Immunol . 2016;60:646–652.
16. Diagnostic tests for Zika virus .http://www.cdc.gov 4th April,2016.
17. Dick GW, Kitchen SF, Haddow AJ. Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg. 1952;46:509–20.
18. Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, Lanciotti RS, et al. Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med. 2009;360:2536–43.
19. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. 2004;5:113.
20. Faye O, Freire CC, Iamarino A, Faye O, de Oliveira JV, Diallo M, et al. Molecular evolution of Zika virus during its emergence in the 20(th) century. PLoS Negl Trop Dis. 2014;8:e2636.
21. Flavivirus lifecycle. http://www.researchgate.net 24th April, 2016.

22. Flower DR. Designing immunogenic peptides. Nature Chemical Biology. 2013;9(12):749–753
23. Gatherer D, Kohl A. Zika virus: a previously slow pandemic spreads rapidly through the Americas. J Gen Virol. 2016;97:269–73.
24. Goo L, VanBlargan LA, Dowd KA, Diamond MS, and Pierson TC. A single mutation in the envelope protein modulates flavivirus antigenicity, stability, and pathogenesis. PLoS Pathog. 2017;13(2): e1006178.
25. Gritzapis AG, Fridman A, Perez SA, Monica NL, Papamichail M, Aurisicchio L, et al. HER-2/neu(657-665) represents an immunogenic epitope of HER-2/neu oncoprotein with potent antitumor properties. Vaccine. 2010;28:162-70.
26. Gupta M, Mahanty S, Greer P, Towner JS, Shieh WJ, Zaki SR, Ahmed R, Rollin PE. Persistent Infection with Ebola Virus under Conditions of Partial Immunity. J Virol. 2004;78(2):958-67.
27. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R; Open Source Drug Discovery Consortium, Raghava GP. In silico approach for predicting toxicity of peptides and proteins. PLoS One. 2013;8(9):e73957.
28. Hailemichael Y, Overwijk W. Peptide-based anticancer vaccines The making and unmaking of a T-cell graveyard. Oncoimmunology. 2013;2(7):e24743.
29. Hamel R, Dejarnac O, Wichit S, Ekchariyawat P, Neyret A, Luplertlop N, et al. Biology of Zika virus infection in human skin cells. J Virol. 2015;89:8880–96.
30. Heinz FX, Stiasny K., Püschner-Auer G, Holzmann H, Allison SL, Mandl CW and Kunz C. Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. Virology. 1994;198:109–117.
31. Henchal EA, Putnak JR. The Dengue Viruses. Clinical microbiology reviews. 1990;3(4): p. 376-396.
32. Hennessey M, Fischer M, Staples JE. Zika virus spreads to new areas—region of the Americas, May 2015–January 2016. MMWR Morb Mortal Wkly Rep. 2016;65:55–8.

33. Huynh-Hoa Bui, Sidney J, Dinh K, Southwood S, Newman MJ and Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 2006;7:153.
34. Joachims T. Making large-scale support vector machine learning practical.; Scholkopf B BC, Smola A. Cambridge MA: MIT Press, editor. 1999:169–184 p.
35. Kolaskar A, Tongaonkar P. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276:172–174.
36. Kostyuchenko VA, Lim EXY, Zhang S, Fibriansah G, Ng TS, Ooi JS, et al. Structure of the thermally stable Zika virus. Nature. 2016; 533: 425-8.
37. Kulkarni-Kale, U., Bhosle, S., Kolaskar, A.S. CEP: a conformational epitope prediction server. Nucleic Acids Res. 2005;33:168–171.
38. Lanciotti RS, Kosoy OL, Laven JJ, Velez JO, Lambert AJ, Johnson AJ, et al. Genetic and serologic properties of Zika virus associated with an epidemic. Emerg Infect Dis. 2008;14:1232–9.
39. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics. 2007;8:424.
40. Lohia N, Baranwal M. Immune responses to highly conserved influenza A virus matrix 1 peptides. Microbiol Immunol. 2017;61(6):225-231.
41. Lohia N and Baranwal M. Identification of Conserved Peptides Comprising Multiple T cell Epitopes of Matrix 1 Protein in H1N1 Influenza Virus. Viral Immunol. 2015;28(10): 570–579.
42. Lohia N, Baranwal M. Conserved Peptides Containing Overlapping CD4+ and CD8+ T-Cell Epitopes in the H1N1 Influenza Virus: An Immunoinformatics Approach. Viral Immunol. 2014;27(5):225-234.
43. Malone RW, Homan J, Callahan MV, Glasspool-Malone J, Damodaran L, et al. Zika Virus: Medical Countermeasure Development Challenges. PLOS Neglected Tropical Diseases. 2016;10(3):e0004530.
44. Manangeeswaran M, Ireland DD, Verthelyi D. Zika (PRVABC59) Infection is Associated with T cell Infiltration and Neurodegeneration in CNS of Immunocompetent Neonatal C57Bl/6 Mice. PLoS Pathog. 2016;12(11):e1006004.

45. Mandl CW, Heinz FX, Stöckl E, Kunz C. Genome sequence of tick-borne encephalitis virus (western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. Virology. 1989;173:291-301.

46. Meloen HR, Langeveld JPM, Schaaper WMM, Slootstra JW. Synthetic Peptide Vaccines: Unexpected Fulfillment of Discarded Hope. 2001;29(3):233-236

47. Morris GM, Huey R, Olson AJ. Using AutoDock for Ligand-Receptor Docking . Curr Protoc Bioinformatics. 2008;Chapter 8:Unit 8.14.

48. Murray, C.L., Jones, C.T., and Rice, C.M. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. Nat. Rev. Microbiol. 2008;6:699–708.

49. Nardin E. The past decade in malaria synthetic peptide vaccine clinical trials. Hum Vaccin. 2010;6(1):27-38.

50. Nemésio H, Palomares-Jerez F, Villalaín J. NS4A and NS4B proteins from dengue virus: membranotropic regions. Biochim Biophys Acta. 2012;1818(11):2818-30.

51. Oyarzún P , Kobe B. Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production, human vaccines and immunotherapeutics. 2016;12(3):763–767.

52. Pardy RD, Rajah MM, Condotta SA, Taylor NG, Sagan SM, Richer MJ. Analysis of the T Cell Response to Zika Virus and Identification of a Novel CD8+ T Cell Epitope in Immunocompetent Mice. PLoS Pathog. 2017;13(2):e1006184.

53. Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. Open Biol. 2013;3:120139.

54. Plourde AR, Bloch EM. A Literature Review of Zika Virus. Emerging infectious diseases journal. 2016;22:7.

55. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanović. S.SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics. 1999;50(3-4):213-9.

56. Randolph V.B., Winkler,G. and Stollar,V. Acidotropic amines inhibit proteolytic processing of flavivirus prM protein. Virology. 1990;184:450–458.

57. Rice, C. M., E. G. Strauss, and J. H. Strauss. Structure of the flavivirus genome. In S. Schlesinger and M. Schlesinger (ed.), Togaviruses and flaviviruses. Plenum Publishing Corp., New York. 1989: p. 279-327.

58. Roberts, Michelle. "Zika-linked condition: WHO declares global emergency". *BBC News*. Retrieved 1 February 2016.

59. Rojas-Caraballo J, Lo´pez-Aba´n J, Pe´rez del Villar L, Vizcaı´no C, Vicente B, et al. In Vitro and In Vivo Studies for Assessing the Immune Response and Protection-Inducing Ability Conferred by Fasciola hepatica-Derived Synthetic Peptides Containing B- and T-Cell Epitopes. PLoS ONE. 2014;9(8):e105323.

60. Saha S, Raghava G.P. AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res. 2006;34:202-9.

61. Sikka V, Chattu VK, Popli RK, Galwankar SC, Kelkar D, Sawicki SG, Stawicki SP, Papadimos TJ. The emergence of zika virus as a global health security threat: A review and a consensus statement of the INDUSEM Joint working Group (JWG). Jounal of global infectious diseases. 2016;8:3-15**.**

62. Singh H, Raghava GPS . ProPred: Prediction of HLA-DR binding sites. Bioinformatics. 2001;17(12):1236-37.

63. Symptoms, Testing, & Treatment.http://www.cdc.gov. 12th june, 2017.

64. Tan X. G., Przekwas A , Kannan R , Ott K , Harrigan T , Roberts J , et al. “A comparative study of the human body model under blast loadings,” in ASME Int. Mech. Eng. Congress and Exposition. 2012.

65. Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P. PEP-FOLD: an updated *de novo* structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res. 2012;40:288–293.

66. Tomar N, K R. DeAffiliated with Machine Intelligence Unit, Indian Statistical Institute . Immunoinformatics: A Brief Review. 2014;1184:23-55.

67. Trott O and Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem. 2010;31(2): 455–461.

68. Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B. A Systematic Assessment of MHC Class II Peptide Binding Predictions and Evaluation of a Consensus Approach. PLoS Comput Biol. 2008;4:e1000048.

69. Wang W., Guan S., Sun S., et al. Detection of circulating antibodies to linear peptide antigens derived from ANXA1 and DDX53 in lung cancer. *Tumor Biology*. 2014;35(5):4901–4905.

70. Winkler CW, Myers LM, Woods TA, Messer RJ, Carmody AB, McNally KL, Scott DP, Hasenkrug KJ, Best SM, Peterson KE. Adaptive Immune Responses to Zika Virus Are Important for Controlling Virus Infection and Preventing Infection in Brain and Testes. J Immunol. 2017;198(9):3526-3535.

71. Winkler, G., S. E. Maxwell, C. Ruemmler, and V. Stollar. Newly synthesized dengue-2 virus nonstructural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. Virology. 1989;171:302- 305.

72. World Health Organization (WHO) and Food and Agriculture Organization (FAO) proposed guidelines to assess the potential allergenicity of protein are available from http://www.fao.org/es/ESN/food/pdf/allergygm.pdf. 30th April, 2016

73. Xu M, Lee EM, Wen Z, Cheng Y, Huang WK, Qian X, Tcw J, Kouznetsova J, Ogden SC, Hammack C, Jacob F, Nguyen HN, Itkin M, Hanna C, Shinn P, Allen C, Michael SG, Simeonov A, Huang W, Christian KM, Goate A, Brennand KJ, Huang R, Xia M, Ming GL, Zheng W, Song H, Tang H. Identification of small-molecule inhibitors of Zika virus infection and induced neural cell death via a drug repurposing screen. Nat Med. 2016;22(10):1101-1107.

74. Yu I. M., Zhang W., Holdaway H. A., Li L., Kostyuchenko V. A., Chipman P. R., et al. Structure of the immature dengue virus at low pH primes proteolytic maturation. Science. 2008;319(5871):1834-7.

75. Yuan b, Zhao L, Xian R, Zhao G. Identification of novel HLA-A/0201-restricted CTL epitopes from Pokemon. Cellular Immunology. 2012;274:54-60.

76. 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. 2008;36:513-8.

77. Zhou Y, Ray D, Zhao Y, Dong H, Ren S, Li Z, Guo Y, Bernard KA, Shi PY, Li H. Structure and function of flavivirus NS5 methyltransferase. J Virol. 2007;81(8):3891–903.
78. Zika virus (strain Mr 766).http://viralzone.expasy.org. 5th april, 2017.
79. Zika virus mutation may explain spread: study.http://www.nbcnews.com 24th April, 2016.
80. Zika Virus Vaccines.http://www.niaid.nih.gov 24thApril, 2017.
81. Zika Virus.http://www.nejm.org 12th june, 2017.
82. Zika virus.http://www.who.int. 5th april, 2017.
83. Zúñiga J, Yu N, Barquera R, Alosco S, Ohashi M, Lebedeva T, Acuña-Alonzo V, Yunis M, Granados-Montiel J, Cruz-Lagunas A, Vargas-Alarcón G, Rodríguez-Reyna TS, Fernandez-Viña M, Granados J, Yunis EJ. HLA class I and class II conserved extended haplotypes and their fragments or blocks in Mexicans: implications for the study of genetic diversity in admixed populations. PLoS One. 2013;8(9):e74442.
84. Zybert I. A., van der Ende-Metselaar H., Wilschut J., Smit J. M. **(**2008**).** Functional importance of dengue virus maturation: infectious properties of immature virions. J Gen Virol. 2008;89: 3047–3051.