IN-SILICO PEPTIDE BASED VACCINE AGAINST ZIKA VIRUS

A DISSERTATION PROPOSAL

SUBMITTED BY:

Parul Sharma

(11210184)

In fulfillment of the requirement for the

Award of degree in

Master of technology in biotechnology

Under The Guidance of

Mr. Vikas Kaushik

Assistant Professor,

Department of Biotechnology

SCHOOL OF BIOENGINEERING AND BIOSCIENCES

LOVELY PROFESSIONAL UNIVERSITY PHAGWARA, PUNJAB-144411

(May, 2017)

CERTIFICATE

This is to certify that Parul Sharma has completed M.Tech dissertation proposal titled "In-silico peptide based vaccine against Zika virus" under my guidance and supervision. To the best of my knowledge, the present work is the result of her original investigation and study. No part of the dissertation proposal has ever been submitted for any other degree or diploma.

The dissertation proposal is fit for the submission and the fulfillment of the conditions for the award of M.Tech Biotechnology.

Date: Mr. Vikas Kaushik

DECLARATION

I hereby declare that the dissertation proposal entitled, "In-silico peptide based vaccine against Zika Virus" submitted for the M.Tech degree is entirely my original work, all ideas and references have been duly acknowledged. It does not contain any work for the award of any other degree or diploma.

Date: **Parul Sharma**

11210184

Table of Contents

I. LIST OF FIGURES

II. LIST OF TABLES

III. ABBREVIATIONS

S.no.	Abbreviations	Full Forms
1.	ZIKV	Zika Virus
2.	RNA	Ribonucleic acid
3.	RT-PCR	Reverse Transcription Polymerase chain reaction
4.	ELISA	Enzyme- linked immunosorbent assay
5.	WHO	World health organization
6.	ss	Single-stranded
7.	YFV	Yellow fever virus
8.	DENV	Dengue virus
9.	CHIKV	Chikungunya virus
10.	WNV	West Nile virus
11.	NS	Non-structural
12.	C	Capsid
13.	prM	Pre membrane
14.	E	envelope
15.	ATP	Adenosine triphosphate
16.	GTP	Guanosine-5'-triphosphate

17.	DNA	Deoxy-ribonucleic acid
18.	ER	Endoplasmic reticulum
19.	CNS	Central nervous system
20.	CTL	Cytotoxic T-cell
21.	MHC	Major histocompatibility complex
22.	ACE	Atomic contact energy
23.	PDB	Protein data bank
24.	DIR	directory
25.	*Ae.*	Aedes species
26.	psf	Protein structure file
27.	RMSD	Root Mean Square deviation
28.	VMD	Visual molecular dynamics
29.	NAMD	Nanosacle molecular dynamics
30.	DCD	DesignCAd drawing file format

IV. CHAPTER

PLAN

CHAPTER 1

INTRODUCTION

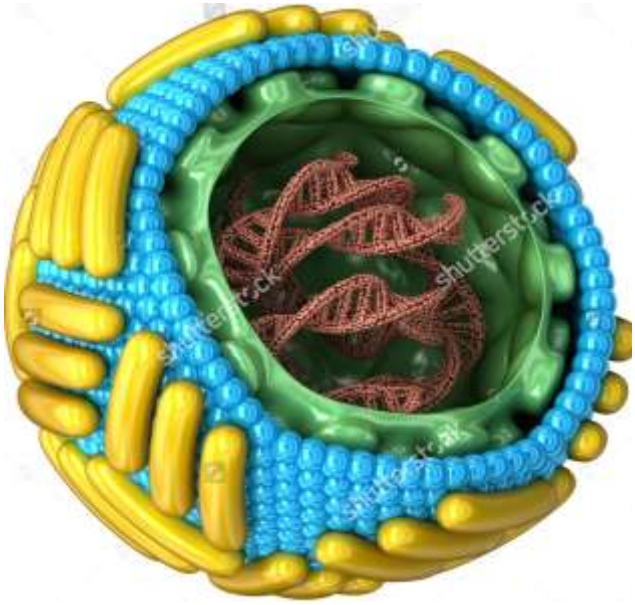


Figure of Zika Virus

Introduction

Zika Virus (ZIKV) exists as mosquito-borne parasite derived out of a sentry rhesus macaque monkey and a group of *Aedes africanus* mosquitoes [1]. It exists as an arthropod-borne virus (arbovirus) which belongs to genus flavivirus. This virus leads to Zika fever (a self –confined feverish disease) and an exanthematic arthralgia syndrome which is exactly resembles dengue fever [2].

It was primarily originated in 1947 from Zika forest situated in Uganda. ZIKV is a positive sense and single stranded RNA virus of family Flaviviridae with genus Flavivirus which has genome of 10,794 nt and is similar to Spondweni virus [3]. ZIKV is transmitted in a zoonotic cycle between non-human anthropoids and arboreal Aedes sp. Mosquitoes present in the forest of Africa and Asia. ZIKV also circulates in sylvatic habitats. According to a serological study done in Nigeria, 40% of citified population acquire neutralizing antibodies to Zika virus [4].

Symptoms and Detection

Some studies showed that ZIKV is native to Africa as well as Southeast Asia. In humans it can be determined by mild fever which ranges from 37.8°C – 38.5°C. Some other symptoms are arthralgia (markedly for small joints of feet and hands), headache, myalgia, conjunctivitis, retro-orbital ache as well as cutaneous maculopapular rash. During acute phase or viremic phase it cannot be diagnosed or misdiagnosed due to some non-specific influenza like symptoms [5].

It can be biologically confirmed by detection of virus RNA within serum with a technique used called RT-PCR (reverse transcription polymerase chain reaction). ELISA is used to detect the IgM against it, but in some laboratories as its diagnosis in laboratory is difficult due to its less cross-reactivity and viremia along added flaviviruses (counting dengue) and that can be confirmed by neutralization assays. The Zika virus RNA can also be detected by real-time RT-PCR with utility of urine [5].

Mortality rate

Thereafter a huge commotion arised in western Pacific region, it has come to attention and includes affected areas like French Polynesia, Easter Island, New Caledonia and Cook Islands.

Infections related to travel have also been broadcasted from western pacific and intermittently in the travelers of additional regions like Indonesia, Senegal and Thailand. A major focus on the current outbreaks is Colombia and Brazil where around 100,000 estimated people have been infected due to which in February 2016; the World Health Organization (WHO) had affirmed Zika Virus becoming a global emergency [2].

CHAPTER 2

TERMINOLOGY

TERMINOLOGY

Virus	any infective agent that consists of a nucleic acid in its protein coat and replicates inside the living cell
Glycoprotein	a class of proteins which consists of carbohydrate groups attached to polypeptide side-chains
Docking	a research technique used for predicting the orientation of any two molecules with each other
Simulation	the formation of any computer model of anything for purpose of some study
Ligand	any molecule or ion that is attached to a metal atom through coordinate bonding
Epitope	any part of antigen that can be recognized by body immune and antibody attaches to it
MHC-I	class of major histocompatibility complex molecules found on the surface of cell required by immune system to recognize foreign invaders
MHC-II	class of major histocompatibility complex molecules found only in antigen presenting cells like dendritic cells
Bioinformatics	a discipline of science where one can collect and analyze complex biological data like genetic codes
Molecular Dynamics	a simulation method for analyzing physical movements of atoms

CHAPTER 3

REVIEW

OF

LITERATURE

REVIEW OF LITERATURE

Virology and Pathogenesis

ZIKV is considered as a positive-sense ssRNA from family Flaviviridae with several other mosquito-borne viruses like YFV, DENV and WNV. ZIKV mostly relates to Spondweni virus [6]. ZIKV genome contains 10,794 nt which encodes for 3.419 aa and also compose of 2 non-coding regions i.e. 5'and 3' which flanks an open reading frame. ZIKV polyprotein is cleaved into capsid, precursor, envelope and 7 other NS proteins [7]. Through its phylogenetic analysis, ZIKV can be classified as Asian and African ancestries which were emerged from east Africa in late 1800s and early 1900s [8]. Its molecular evolution study demonstrated that the viral strains assembled from 4 different countries in West Africa (1947-2007) shows several sites within the genome with strong negative selection force [9]. When the inoculation of mosquito in human host is done, its cellular entry was similar to that of flavivirus whereas when the virus enters the cells of skin via cellular receptors it enables the movement to bloodstream and lymph nodes. Some studies demonstrated about pathogenesis of ZIKV infection. A study showed that the fibroblasts of human skin, some immature dendritic cells and keratinocytes allow the entry of ZIKV [10]. Some adhesion factors and entries assist infection, cellular autophagy and require flaviviral replication which improves ZIKV replication in fibroblasts of skin [00]. After its entry in cell, virus replicates in ER- derived vesicles. The antigens of ZIKV were found in the nucleus of infected cells and also at the location of replication [11].

Genome of Zika Virus

Zika virus is a single-stranded, non-segmented, positive sense RNA genome which nearly relates to the *Spondweni virus* and abides as twin viruses in the Spondweni virus clade. ZikaVirus particles have diameter of 40 nm, also consists of a compressed inner core as well as an external envelope [12]. The Zika virus RNA is long with 10,617-nucleotide containing two non-coding portions called as the 5' and 3' NCR. The Zika Virus open reading frame is read as follows: 5′-C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-3′. The polyprotein inside it is cleaved into envelope (E), capsid (C), precursor membrane (prM) and non-structural proteins (NS). The envelope consists of virion covering as well as is immersed in replication specifically membrane fusion and host cell binding. The NS1, NS3 and NS5 are abundant, highly-conserved

proteins whereas the NS2A, NS2B, NS4A and NS4B are smaller, hydrophobic proteins. The 3' NCR region consists of 428 nucleotides are located which helps in translation, cyclization, genome stabilization, RNA packaging and recognition. This region makes up a loop structure whereas the 5' NCR helps in translation through a methylated nucleotide cap or a protein which is genome-linked [13].

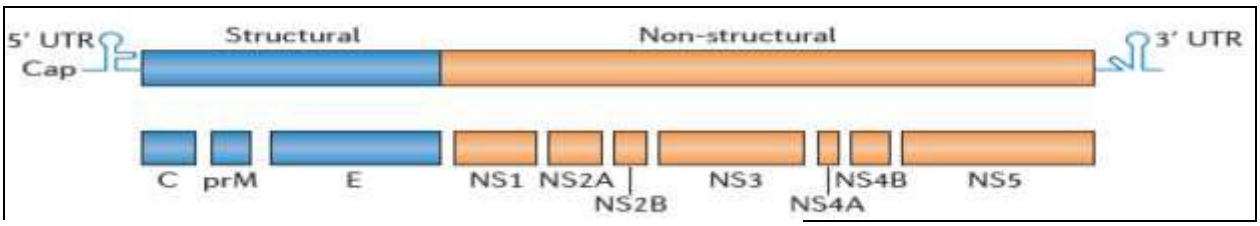


Fig_1: Genome of ZIKV

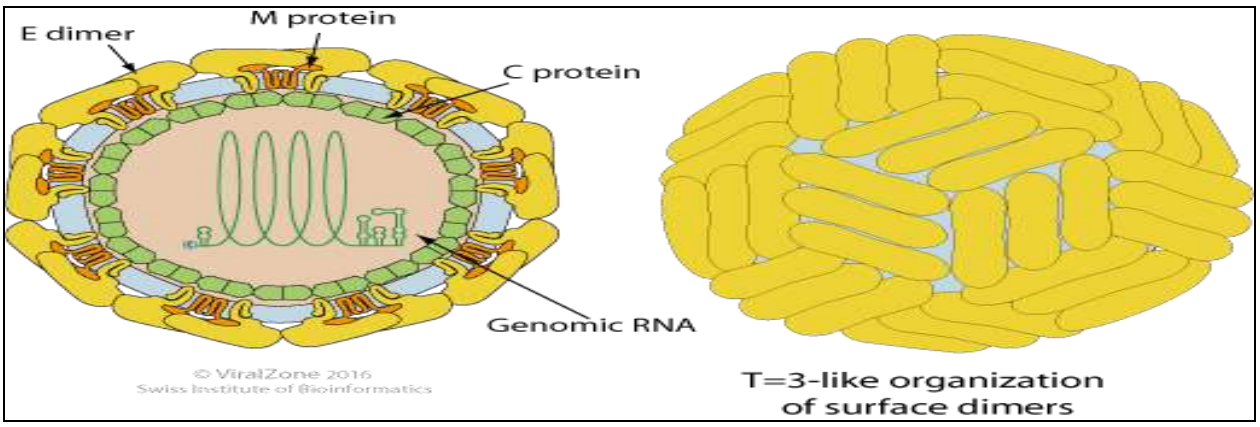


Fig_2: Virion of ZIKV

Zika virus genome contains 39 & 59 un-translated regions flanking a single open reading frame (ORF) that encodes a polyprotein that is cleaved into three structural proteins: the capsid (C), premembrane/membrane (prM), and envelope (E); and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B, and NS5)[14] . The 5′ end of positive strand genomic RNA is modified with a cap-1 structure formed by an RNA triphosphatase, with guanylyltransferase, N7-methyltransferase and 2′-O methyltransferase. The non structural proteins are responsible for these activities. The NS3 protein encodes a RNA triphosphatase within its helicase domain. The N-terminal domain of the non-structural protein 5 (NS5) has both the N7methyltransferase and guanylyltransferase activities necessary for forming mature RNA cap structures. RNA binding

affinity is reduced by the presence of ATP or GTP and enhanced by Sadenosyl methionine [15]. This protein also encodes a 2′-O methyltransferase.

About Proteins and their functions

The NS3 with NS5 proteins are required for viral replicative capacity as they are known for nearly all of the catalytic actions essential for both capping as well as replicating the viral RNA [16].

Protein Name	Property	Function
Envelope Protein	Comprises of majority virion surface	Involved in replication such as host cell binding and membrane fusion.
NS1	Highly conserved protein with species of some epitopes and also contains His tag	Secretes extracellular space as a hexameric lipoprotein which helps in pathogenesis and immune envasion.
NS2A & NS2B	Responsible for cytoplasmic cleavage	Essentially required for virus replication.
NS3	Multi domain protein with N-terminal NS3Pro as well as C-terminal including RNA triphosphatase(NS3RTPase) with RNA helicase(NS3Hel)	Helps viral synthesis as well as capping.
NS4A & NS4B	Small and hydrophobic proteins responsible for microcephaly identification	Induces autophagy

NS5	N-terminal methyltransferase (NS5MTase) domain with C-terminal RNA-dependent RNA polymerase (Ns5RdRp) domain.	Behave as a motif for • Synthesis of ssRNA • Synthesis of viral polyprotein

Table_1: Properties and functions of proteins

Lifecycle of Zika Virus

The replication of ZIKV is similar to that of flavivirus. As effective virion binds with the receptors of host cell membrane through envelope protein, it activates endocytosis. Then the pathogen membrane get intermix with endosomal membarane as well as the single stranded RNA genome is liberated into cytoplasm of host cell where it is interpreted into polyprotein. The polyprotein is splitted into structural as well as non structural proteins. Later the replication begins in the intracellular compartments resulting into double stranded DNA genome which is further transcribed into ssRNA genomes. Then the Assembly occurs in endoplasmic reiticulum as well as the newly formed virions are transferred into Golgi apparatus which after that time was discharged within intracellular space at which they infect fresh host cells [17].

Mode of Transmission

ZIKV is mainly transmitted through mosquito bite of family *Aedes* genus like *Ae. Aegypti* , *Ae. africanus*, *Ae. hensilli*, and *Ae. albopictus* [18]. ZIKV is considered as wild-caught for *Ae. aegypti* mosquitoes as some laboratory experiments have proved that it is capable of transmitting ZIKV [19]. This species of mosquito are globally distributed and have native habitats for several species in warm and sub-tropical regions [20]. The Mosquito attains virus during a blood meal and after the uptake, it replicates and then is transferred to an animal at another meal [21]. Through a study, it was demonstrated that the kinetics of ZIKV infection is via blood-feeding membranes of *Ae. aegypti* mosquito and the viral content is high on the day of feeding and decreased to some levels till the 10^{th} day whereas increases till day 15 and is at high level during 20-60 days. Thus it was demonstrated that the incubation period of mosquito was around 10 days.

Non-vector modes for ZIKV transmission consists of perinatal, congenital and sexual ways [22].the possible way of ZIKV transmission is via blood transfusion, animal bite and laborartory exposure [23]. Transmission through intrauterine of ZIKV RNA is via reverse transcription PCR (RT-PCR) within amniotic fluid of mother during pregnancy and the babies are delivered with microcephaly [21]. ZIKV RNA was detected in the tissue of foetus of women at the time of pregnancy and also in the brain of newly-borned infants who died within 20 hours due to microcephaly. ZIKV viral RNA has been detected in breastmilk but it was not transferrable [24].

Clinical Manifestations

It was demonstrated that the incubation period of ZIKV in humans while mosquito bite was not more than 12 days. Infection doesn't show symptoms in more than 80% cases through mosquito bite [25]. It symptoms were mild, non-specific and self-controlable like with other arbovirus infections like DENV and CHIKV. Mostly reported symptoms were fever, rash, myalgia, headache, fatigue, arthralgia and conjunctivitis. In French Polynesia, some health officers have reported an sudden increase in CNS malformations [26].

General Laboratory Findings

The information regarding to laboratory findings for ZIKV is inadequate. In certain cases the complete blood count is found to be normal but blood count was abnormal and some changes may be non-specific like mild neutropenia, lymphopenia and moderate thrombocytopenia [27]. Clinical evaluation for ZIKV, CHIKV and DENV was undertaken with the one having fever, rash, myalgia or arthralgia. RT-PCR on samples was preferred as the testing method or diagnostic approach for ZIKV within 7 days of symptom onset. However serologic testing was not allowed for ZIKV [28].

Management and Prevention

Till now no specific vaccine or treatment is available for ZIKV infection. Thus, management includes rest, analgesics, antipyretics and some fluids. Some non steroidal anti-inflammatory drugs like Apirin are not allowed until dengue is excluded. Another measures are to prevent from mosquito bite include individual protection through bed-nets, repellents etc. Thus, pregnant

women are not allowed to travel the countries which are affected by ZIKV and are required to do testing in the growth weeks after every 3-4 weeks whereas the infant at the time of birth should also be tested. Males who have travel to infected countries should abstain from sexual activity as a prevention measure for ZIKV [29].

CHAPTER 4

AIM

AND

OBJECTIVES

Aim and objectives

The general objective of the research is to design a peptide based vaccine for Zika Virus using in-silico methods, thereby predicting and screening different epitopes and to study the applications of Bioinformatics techniques used for Zika Virus.

The specific objective of the research is to design and predict a suitable candidate epitope which can be used as a vaccine for Zika virus and can be used for human.

CHAPTER 5

RATIONALE

OF

STUDY

RATIONALE OF STUDY

The epidemic of Zika virus (ZIKV) infection has become a universal health concern due to it microcephaly and neurological barrier. However, there is a need to develop vaccine for treatment or prevention from infection. Through this research some bioinformatics approach was employed to predict epitopes of Zika viral proteins for development of a peptide based vaccine against it. Thus both B as well as T- cell epitopes was predicted for ZIKV NS3 and NS5 proteins. In this study, 15 epitopes of both MHC-I and II proteins were screened regarding some parameters by predicting their structures through PEP-FOLD and then were docked with MHC-I and II complexes. The stability of the resulting peptide with MHC I and II complexes was further studied by NAMD and VMD molecular dynamics simulations. The simulation results underline the limits of rigid-body docking approaches. A number of the antigenic epitopes predicted and analyzed in this work might present a preliminary set of peptides for future vaccine development against ZIKV.

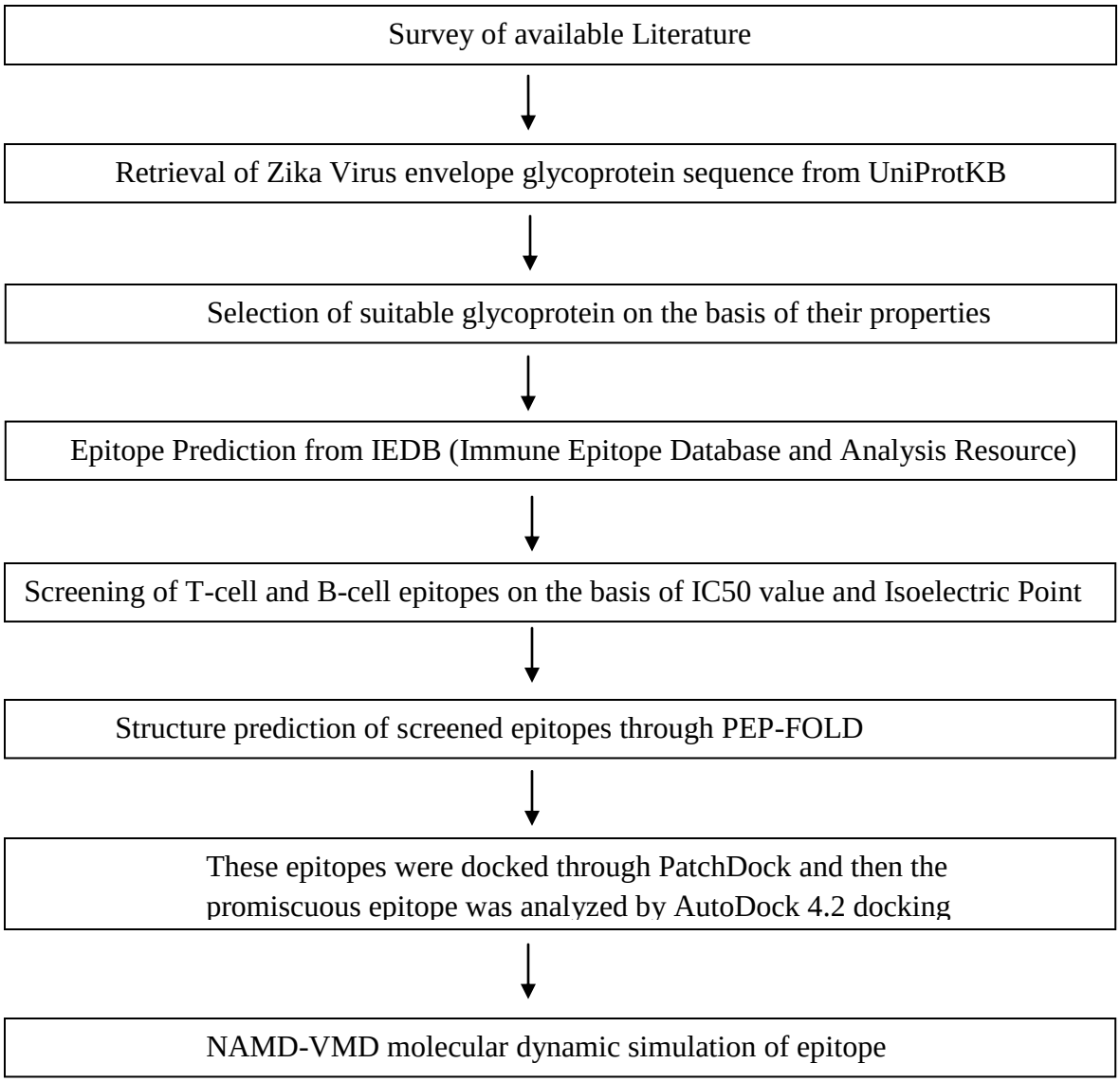
CHAPTER 6

RESEARCH METHODOLOGY

Methodology

To design an effective synthetic peptide vaccine candidate, in silico modeling and immunoinformatics strategies have been exploited which uses a variety of statistical and machine learning approaches by the help of bioinformatics software and machine learning programs. Such active vaccine candidate must have to contain minimum two antigenic epitopes; one to induce specific B cell or CTL responses while other induce specific T_H cell response.

Flowchart representing bioinformatics top down approach employed in peptide based vaccine prediction against Zika virus:



6.1. Survey of the literatures

To study the available literatures through which, one can gather knowledge about Zika Virus, its lifecycle and pathogenesis. To choose suitable method for analysis and to study various bioinformatics techniques used in vaccine designing of Zika virus.

6.2. Retrieval of Zika virus envelope glycoprotein sequence

Zika virus outer membrane glycoprotein sequence was required specifically for humans. It was derived from UniProtKB (www.uniprot.org) in FASTA format [30]. It is an online database which contains protein sequences associated with computationally generated annotation as well as large-scale functional characterization. It is considered as the central hub for collection of functional information of proteins with rich, accurate and consistent annotation. It consists of 2 sections in which one contains manually-annotated records or information derived through literature and other contains computationally analyzed results [31].

6.3. Selection of suitable glycoprotein

NS3 and NS5 are considered to be the suitable glycoprotein for analysis of epitope or epitope prediction. These two proteins promote genomic RNA cyclization. NS3 protein codes for RNA triphosphatase which lies in its helicase domain. NS5 protein is largest viral protein in which RNA-dependent RNA polymerase activity lies at the C-terminal part and thus RNA-capping is done with the help of N-terminus. It has both the two activities required for RNA cap structures formation which are N7-methyltransferase and guanylyltransferase. With respect to these capabilities and functions these two proteins are considered [32].

6.4. Epitope prediction from IEDB

The NS3 and NS5 protein sequences are considered for MHC class-I and II predictions. The epitopes were derived from IEDB (Immune Epitope Database and Analysis Resource) which provides an experimental catalog for B and T cell epitopes, as well whereas a data on Major Histocompatibility Complex (MHC) binding and elution experiments on MHC ligand. This database recognizes epitopes in humans as well as non human species [33]. Therefore, their molecular weight, half-life of dissociation and Isoelectric point was also derived through ExPASy (Bioinformatics Resource Portal).

6.5. Screening of epitopes

After B and T-cell epitopes screening for both proteins, molecules for NS3 and NS5 are screened on the basis of similarities between MHC-I and MHC-II binding regions. The parameters used to screen them were low IC50 value and isoelectric point. From both the proteins (NS3 and NS5) of MHC-I and MHC-II molecules, only 10 molecules were screened out.

6.6. Structure Prediction of the screened epitopes

PEP-FOLD is an online resource for de-novo structure prediction for peptide which aims to 3D conformation modeling of peptides within 9-25 amino acids. PEP-FOLD predicts the SA letter profiles using Hidden-Markov Model from sequence and then results the predicted fragments [33].

6.7. Docking

6.7.1 Docking through PatchDock

The docking score was determined by PatchDock software. The NS3 and NS5 epitopes were docked with the Zika virus MHC-I and MHC- II ligand and thus the 10 receptors of each protein is screened on the basis of best score.

PatchDock is an online web server designed with algorithm for molecular docking. It includes input of two molecules one ligand and other receptor of any type like drugs, proteins, peptides and DNA. The output contains a list of suitable complexes arranged by shape complementarity criteria and also displays score, area, ACE (atomic contact energy), transformation with downloadable PDB file of the complex. The algorithm of PatchDock includes 3 major stages:

- Molecular Shape Representation: this step computes the molecular surface and then apply segmentation algorithm for detecting geometric patches like concave, convex and flat surfaces.
- Surface Patch Matching: a hybrid of pose-clustering and geometric hashing matching techniques are required to compare the patches detected previously as concave ones are matched with convex and flat ones with any.

- Filtering and scoring: the suitable complex from previous step is determined while all other complexes are discarded and the remaining ones are ranked on the basis of geometric shape complementarity score [34].

6.7.2. Docking with AutoDock

AutoDock is a collection of automated docking tools. It is generally used to predict the binding capability of a molecule like substrate or drug applicant to the receptor of a recognized 3-D structure. Currently it is divided into two divisions: AutoDock 4 and AutoDock Vina.

AutoDock4 follows two main programs: *autodock* performs docking of ligand with the suite of grids which usually describes target protein; *autogrid* is used to pre-calculate the above grids.

AutoDock have several applications in lead optimization, X-ray crystallography, protein-protein docking, structure-based drug designing, chemical mechanism studies and further more. AutoDock is freely available software under GNU and Apache license for commercial as well as non-commercial purposes. AutoDock is fast software which provides high quality predictions for ligand conformations and a superior association within predicted inhibition constants and experimental ones. It is also useful for blind docking [35].

The MHC-I and II complex is then docked with this particular epitope through AutoDock by following these steps:

Converting PDB file to pdbqt:-

- The ligand molecule is in PDB format. Convert it to pdbqt format by using openBabel.
- Set INPUT FORMAT as pdb – Protein Data Bank Format.
- Select the Ligand molecule
- Set OUTPUT FORMAT as pdbqt – AutoDock PDBQT format.
- Choose the output file from ADTworkspace\filename.pdbqt
- Click on convert option. The PDB file gets converted to pdbqt format.

AutoDock:-

- Open AutoDock tools-1.5.6.
- Go to FILE and select Read molecule. Select and open the Receptor Molecule from ADTworkspace.

- Select the required chain and delete other chains. The required chain can have maximum binding of the receptor and ligand molecules.

[Now we have to delete the Water molecules and add Hydrogen]

- EDIT→Delete water.
- EDIT→Hydrogen→Add.
- EDIT→Hydrogen→Merge Non-Polar.
- EDIT→Charges→Compute Gasteiger→Total Gasteiger added →Ok.
- FILE→Save→Write PDB→filename.pdb.
- Go to Ligand→Input→Open (the ligand file).
- Ligand→Torsion tree→Choose Torsions.
- The torsion value should be less than 32. So we have to delete the route to keep the torsion value maximum 32.

[**Now we have to perform Auto grid]**

- Go to GRID→Select Macromolecule→Choose Receptor→Save as filename.pdbqt.
- GRID→Set Map Types→Choose Ligand Molecule.
- GRID→Set Map Types→Set Up Covalent Maps.
- A box of Covalent Grid Parameters, AD4 will open.
- Now, Set
  - Energy barrier height = 1000.
  - Half-Width (Angstrom) = 5.00.
- Now click on the selected chain. Select the any residue (having OH bonds).
- Now after selecting the residue go to Covalent Grid Parameters, AD4 box and click on → Use selection for attachment atoms.
- The values for X, Y and Z coordinates will be selected.
- Go to GRID →Grid Box.
- Again a Grid Options box will open.
- Set Number of points in X,Y,Z-dimensions = 100.
- Click on center grid box option.
- Set X,Y,Z-center according to the desired results.
- Go to FILE and click on Close to saving current.
- GRID →Output →Save GPF.
- Save the Grid file as filename.gpf.
- Now we have to RUN our Grid file.
- RUN→Run AutoGrid
- Program Pathname→Browse→Select file from Local Disk C: →Program files(X86) → AutoDock → 4.2.6 →AutoGrid4.exe.

- Browse the Parameter Filename
- Click on Launch.
- Now AutoGrid starts to run.

[**Now we have to perform the final docking (Run AutoDock)]**

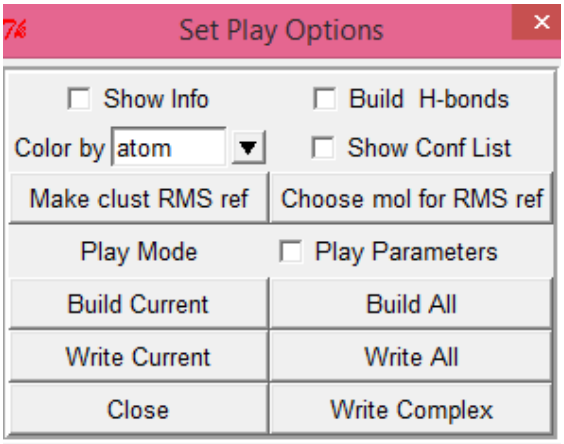
- After performing AutoGrid, go to DOCKING option.
- Docking →Macromolecule →Set rigid filename →filename.pdbqt →open.
- Select ligand from docking options
- Docking→Search Parameters→Genetic Algorithm.
- Genetic Algorithm Parameters box will appear.
- Number of GA runs = 10.
- Population Size = 150.
- Maximum numbers of evals = Medium.
- Then click on Accept.
- Docking →Output → Lamarckian GA (4.2.6) →filename.dpf (save).
- RUN→ Run AutoDock.
- Program Pathname →Browse →Select file from Local Disk C: →Program files(X86) →AutoDock → 4.2.6 → AutoDock.exe.
- Browse parameter filename
- Click on Launch.
- Now our AutoDock will start.
- After the completion of AutoDock, the docked file will save as dlg file format (dock.dlg).

Analyzing the docked results

- Select the Analyze option, open→ filename.dlg →OK
- Deselect the original ligand file by selecting it and deleting it.
- Then open Analyze → Macromolecule → Choose → filename_pregrid.pdbqt
- Then open Analyze → Conformations → Load
- In the Conformation chooser, the top 10 conformations with energy rank were displayed
- Then Analyze → Conformations → Play
- This window appears on screen ,

By changing the side bars, one can view different conformations (10 conformations).

- Press  , a new window like this appears



- In this window, several parameters can be changed according to our convenience like
  Play mode → continuously in one direction
  Select → Build H-bonds option
  Select → Show info
- By selecting show info option, a new window appears with different parameters like binding energy, ligand efficiency, intermolecular energy, electrostatic energy, desolvation energy, total energy, torsional energy, unbound energy, filename, cIRMS, refRMS, rseed1 and rseed2 with number of Hydrogen bonds formed.
- Note down the energies as per requirement
- Out of the 10 conformations, the one with the least docked energy is considered and the one with least binding energy.
- Select the "Build Current" option, after which a new ligand file will be formed
- For saving that ligand conformation, click on "Write Current" → Save
- Then click on "Write Complex" → Save [to save the complex for further analysis]

[**PyMOL used for analysis**]

- Double click on the file formed of the complex and choose PyMOL as a option to visualize the complex
- In the PyMOL viewer the receptor and ligand binding picture will appear
- On clicking the A [for action] option, a drop down list appears
- Click on find → polar contacts → to any atoms
  [A new complex will appear named "obj_01polar_contact"]
- Again click on A [action] of the original complex and in the drop down list, click Preset → ligands
- Click on S [Show] of the new complex and in the drop down list, click Labels.
- Click on the PyMOL Main window, select File → save molecule [in pdb format] → filename.pdb → OK
- This .pdb file is further used for simulation purpose.

6.8. VMD & NAMD Molecular Dynamic Simulation of Epitope

NAMD stands for Nanoscale Molecular Dynamics as this software provides molecular dynamics simulation. It was developed with a joint association of Theoretical and Computational Biophysics Group (TCB) and the Parallel Programming Laboratory (PPL) at University of Illinois at Urbana-Champaign [36].
VMD stands for Visual Molecular Dynamics, the software designed for displaying, animating and analyzing large biomolecules using 3D graphics. It was also developed by TCB (Theoretical and Computational Biophysics) Group at University of Illinois at Urbana-Champaign [37].
The steps of VMD and NAMD simulation are:

After complete installation of VMD and NAMD,

- Open VMD, 3 windows appear on the screen that are VMD 1.9 OpenGL Display, VMD Main and VMD exe
- Open VMD main → extension →TKconsole
- In TKconsole, type ' DIR', to search "namd-tutorial-files"
- Type 'CD NAMD-TUTORIAL-FILES' , to enter this folder and check its directory by typing 'DIR'
- Type 'CD 1-1-BUILD' to access this folder as this contains the input files.
- In the MAIN window
  File → New molecule → Browse →filename.pdb → Load.

[**PDF to PSF conversion**]

- Type 'SET FILENAME [ATOMSELECT TOP PROTEIN] in the TKconsole window
- Then, type '$FILENAME WRITEPDB PARAMETERFILENAME.PDB'
- In the Main window,
  Select Filename.pdb → molecule → delete molecule
- Select extension → modeling → automatic psf converter
  [Load the molecule with parameterfilename.pdb and then change the output filename according to requirement, press on I AM FEELING LUCKY to proceed further and the psf file will be formed with a display message of conversion]
- In TKconsole window , type 'DIR' to check the psf file conversion

[**Sphere and Box formation**]

- In TKconsole window, type 'SOURCE WAT_SPHERE.TCL'
- In the Main window,
  File → New molecule → select filename_ws.psf file → Load → again Browse → select filename_ws.pdb → Load → OK

- Type ‘PACKAGE REQUIRE SOLVATE’ IN TKconsole window
- Type ‘SOLVATE FILENAME.PSF FILENAME.PDB –T 5 –O FILENAME_WB’
- In the Main window,
  File → New molecule → select filename_wb.pdb → Load
- Type ‘SET EVERYONE [ATOMSELECT TOP ALL]’ in TKconsole window
- Type ‘MEASURE MINMAX $EVERYONE’
- Go to Main window → QUIT

[**Copy all files to the common folder**]

- Open command prompt
- Type ‘CD..’ , till you get into the C Drive
- Type ‘ CD VMD’ to enter the VMD folder and again ‘CD NAMD-TUTORIAL-FILES’ to enter namd-tutuorial-files folder
- Type ‘CD 1-1-BUILD’ to copy the required files to Common folder
- With commands ‘COPY FILENAME.PDB ..\COMMON’ or ‘COPY FILENAME.PSF ..\COMMON’ , copy all the box and sphere files to common folder.

Water sphere simulation with non-periodic boundary condition

- Open Wordpad → open file from 1-1-sphere folder in namd-tutorial-files → filename_ws_eq.conf → OK
- Study the opened file thoroughly with full job description.
- Run simulation
  [Download NAMD package and paste it in C Drive and in the advanced system settings of the computer edit the path by adding the NAMD folder path to run simulation]
- Now from command prompt by typing the previously used commands go into 1-2-sphere folder of namd-tutorial-files
- Type ‘FILENAME_WS_EQ.CONF > FILENAME_WS_EQ.LOG &

[Several were formed and will take around 20 minutes to complete the whole process]

Water box simulation with periodic condition

- In command prompt, by using previous commands enter into the 1-3-box folder of namd-tutorial-files
- Type ‘NAMD2 FILENAME_WB_EQ.CONF > FILENAME_WB_EQ.LOG &’ in the command prompt.

For analysis of molecular simulation, 3 things are required:

- Log file
- RMSD value

- Confirmation of protein molecule

Analysis:

- Open Wordpad → open file from 1-2-sphere folder of namd-tutorial-files → select file with name filename_ws_eq.conf → OK
- Open VMD main window → extensions → analysis → NAMD Plot
- In NAMD Plot → select file → select NAMD Log file → open filename_ws_eq from 1-2-sphere folder of namd-tutorial-files
- Select TEMP from NAMD Plot
- From NAMD Plot → select File → plot selected data [then cancel all the files open]
- In VMD Main window,
  File → New molecule → Browse → select filename_ws.psf file from common folder of namd-tutorial-files → Load → then again Browse → select filename_ws_eq.dcd file from 1-2-sphere folder of namd-tutorial-files → Load →OK
- Type 'LS' in TKconsole window , then type 'CD NAMD-TUTORIAL-FILES' and ' CD 1-2-SPHERE'
- Type ' SOURCE RMSD.TCL'
  [Side by side open an Excel sheet to plot the graph
  Open file → select rmsd.dat file from 1-2-sphere folder of namd-tutorial-files →select → cut the values that appear on the excel sheet into some another rows and columns → mark them from 0 to no. till the ending value → select all the values → go to insert option → scatter → scatter with straight lines.

Comparing the protein confirmation before and after the simulation

- Open VMD Main window, load the file named "filename_wb.psf", then
  Select Graphics → Representation → In representation window → select coloring method, drawing method for different views like new cartoon and lines.
- Type 'PROTEIN' in selected atoms of graphical representation window

[**Analyzing box simulation**]

- In the VMD Main window, set the bar according to the required frame.
- In the Main window, select the molecule named "filename_wb.psf" and select the 'save coordinates' in File menu
- In the new window, select protein from the drop down arrow of selected atoms and change the file type to pdb → click save → save the file wherever needed with name "filename_box_frames"
- Then change the frame from option of 1 to 9 and click "save" to save the required frame by providing an efficient name.
- QUIT from VMD

- Open PyMOL → file → open → "filename_box_frames"
- Select the initial and last frame while deselect the other frames
- Also deselect the last frame and from initial, select A (for all)→ preset → ligands
- Repeat the previous step for other molecules too

[Analyzing sphere simulation]

- Load the molecule named "filename_ws.psf" from common folder of namd-tutorial-files in the VMD Main window and then again load file named "filename_ws_eq.dcd".
- In the VMD Main window, set the bar according to the required frame.
- In the Main window, select the molecule named "filename_ws.psf" and select the 'save coordinates' in File menu
- In the new window, select protein from the drop down arrow of selected atoms and change the file type to pdb → click save → save the file wherever needed with name "filename_sphere_frames"
- Then change the frame from option of 1 to 9 and click "save" to save the required frame by providing an efficient name.
- QUIT from VMD
- Open PyMOL → file → open → "filename_sphere_frames"
- Select the initial and last frame while deselect the other frames
- Also deselect the last frame and from initial, select A (for all)→ preset → ligands
- Repeat the previous step for other molecules too.

CHAPTER 7

RESULTS

AND

DISCUSSION

7.1. Retrieval of envelope glycoprotein sequence

Protein Name	Protein Length	Sequence
Q32ZE1-1	3419	>sp|Q32ZE1|POLG_ZIKV Genome polyprotein OS=Zika virus (strain Mr 766) PE=1 SV=1 MKNPKEEIRRIRIVNMLKRGVARVNPLGGLKRLPAGLLLGH GPIRMVLAILAFLRFTAIKPSLGLINRWGSVGKKEAMEIIKKF KKDLAAMLRIINARKERKRRGADTSIGIIGLLLTTAMAAEIT RRGSAYYMYLDRSDAGKAISFATTLGVNKCHVQIMDLGHM CDATMSYECPMLDEGVEPDDVDCWCNTTSTWVVYGTCHH KKGEARRSRRAVTLPSHSTRKLQTRSQTWLESREYTKHLIK VENWIFRNPGFALVAVAIAWLLGSSTSQKVIYLVMILLIAPA YSIRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDK PTVDIELVTTTVSNMAEVRSYCYEASISDMASDSRCPTQGEA YLDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKF TCSKKMTGKSIQPENLEYRIMLSVHGSQHSGMIGYETDEDR AKVEVTPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYLT MNNKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALV EFKDAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKG RLFSGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKVPAETLH GTVTVEVQYAGTDGPCKIPVQMAVDMQTLTPVGRLITANP VITESTENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGS TIGKAFEATVRGAKRMAVLGDTAWDFGSVGGVFNSLGKGI HQIFGAAFKSLFGGMSWFSQILIGTLLVWLGLNTKNGSISLT CLALGGVMIFLSTAVSADVGCSVDFSKKETRCGTGVFIYND VEAWRDRYKYHPDSPRRLAAAVKQAWEEGICGISSVSRME NIMWKSVEGELNAILEENGVQLTVVVGSVKNPMWRGPQRL PVPVNELPHGWKAWGKSYFVRAAKTNNSFVVDGDTLKECP LEHRAWNSFLVEDHGFGVFHTSVWLKVREDYSLECDPAVI GTAVKGREAAHSDLGYWIESEKNDTWRLKRAHLIEMKTCE WPKSHTLWTDGVEESDLIIPKSLAGPLSHHNTREGYRTQVK GPWHSEELEIRFEECPGTKVYVEETCGTRGPSLRSTTASGRV IEEWCCRECTMPPLSFRAKDGCWYGMEIRPRKEPESNLVRS MVTAGSTDHMDHFSLGVLVILLMVQEGLKKRMTTKIIMSTS MAVLVVMILGGFSMSDLAKLVILMGATFAEMNTGGDVAH LALVAAFKVRPALLVSFIFRANWTPRESMLLALASCLLQTAI SALEGDLMVLINGFALAWLAIRAMAVPRTDNIALPILAALTP LARGTLLVAWRAGLATCGGIMLLSLKGKGSVKKNLPFVMA LGLTAVRVVDPINVVGLLLLTRSGKRSWPPSEVLTAVGLICA LAGGFAKADIEMAGPMAAVGLLIVSYVVSGKSVDMYIERA GDITWEKDAEVTGNSPRLDVALDESGDFSLVEEDGPPMREII LKVVLMAICGMNPIAIPFAAGAWYVYVKTGKRSGALWDVP APKEVKKGETTDGVYRVMTRRLLGSTQVGVGVMQEGVFH

		TMWHVTKGAALRSGEGRLDPYWGDVKQDLVSYCGPWKL DAAWDGLSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAV ALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQG KREEETPVECFEPSMLKKKQLTVLDLHPGAGKTRRVLPEIV REAIKKRLRTVILAPTRVVAAEMEEALRGLPVRYMTTAVNV THSGTEIVDLMCHATFTSRLLQPIRVPNYNLNIMDEAHFTDP SSIAARGYISTRVEMGEAAAIFMTATPPGTRDAFPDSNSPIM DTEVEVPERAWSSGFDWVTDHSGKTVWFVPSVRNGNEIAA CLTKAGKRVIQLSRKTFETEFQKTKNQEWDFVITTDISEMGA NFKADRVIDSRRCLKPVILDGERVILAGPMPVTHASAAQRR GRIGRNPNKPGDEYMYGGGCAETDEGHAHWLEARMLLDNI YLQDGLIASLYRPEADKVAAIEGEFKLRTEQRKTFVELMKR GDLPVWLAYQVASAGITYTDRRWCFDGTTNNTIMEDSVPA EVWTKYGEKRVLKPRWMDARVCSDHAALKSFKEFAAGKR GAALGVMEALGTLPGHMTERFQEAIDNLAVLMRAETGSRP YKAAAAQLPETLETIMLLGLLGTVSLGIFFVLMRNKGIGKM GFGMVTLGASAWLMWLSEIEPARIACVLIVVFLLLVVLIPEP EKQRSPQDNQMAIIIMVAVGLLGLITANELGWLERTKNDIA HLMGRREEGATMGFSMDIDLRPASAWAIYAALTTLITPAVQ HAVTTSYNNYSLMAMATQAGVLFGMGKGMPFMHGDLGV PLLMMGCYSQLTPLTLIVAIILLVAHYMYLIPGLQAAAARA AQKRTAAGIMKNPVVDGIVVTDIDTMTIDPQVEKKMGQVL LIAVAISSAVLLRTAWGWGEAGALITAATSTLWEGSPNKYW NSSTATSLCNIFRGSYLAGASLIYTVTRNAGLVKRRGGGTGE TLGEKWKARLNQMSALEFYSYKKSGITEVCREEARRALKD GVATGGHAVSRGSAKIRWLEERGYLQPYGKVVDLGCGRG GWSYYAATIRKVQEVRGYTKGGPGHEEPMLVQSYGWNIVR LKSGVDVFHMAAEPCDTLLCDIGESSSSPEVEETRTLRVLSM VGDWLEKRPGAFCIKVLCPYTSTMMETMERLQRRHGGGLV RVPLCRNSTHEMYWVSGAKSNIIKSVSTTSQLLLGRMDGPR RPVKYEEDVNLGSGTRAVASCAEAPNMKIIGRRIERIRNEHA ETWFLDENHPYRTWAYHGSYEAPTQGSASSLVNGVVRLLS KPWDVVTGVTGIAMTDTTPYGQQRVFKEKVDTRVPDPQEG TRQVMNIVSSWLWKELGKRKRPRVCTKEEFINKVRSNAAL GAIFEEEKEWKTAVEAVNDPRFWALVDREREHHLRGECHS CVYNMMGKREKKQGEFGKAKGSRAIWYMWLGARFLEFEA LGFLNEDHWMGRENSGGGVEGLGLQRLGYILEEMNRAPGG KMYADDTAGWDTRISKFDLENEALITNQMEEGHRTLALAVI KYTYQNKVVKVLRPAEGGKTVMDIISRQDQRGSGQVVTYA LNTFTNLVVQLIRNMEAEEVLEMQDLWLLRKPEKVTRWLQ SNGWDRLKRMAVSGDDCVVKPIDDRFAHALRFLNDMGKV RKDTQEWKPSTGWSNWEEVPFCSHHFNKLYLKDGRSIVVP CRHQDELIGRARVSPGAGWSIRETACLAKSYAQMWQLLYF HRRDLRLMANAICSAVPVDWVPTGRTTWSIHGKGEWMTTE DMLMVWNRVWIEENDHMEDKTPVTKWTDIPYLGKREDLW CGSLIGHRPRTTWAENIKDTVNMVRRIIGDEEKYMDYLSTQ

		VRYLGEEGSTPGVL
prM	123-290	>sp|Q32ZE1|123-290 AEITRRGSAYYMYLDRSDAGKAISFATTLGVNKCHVQIMDL GHMCDATMSYECPMLDEGVEPDDVDCWCNTTSTWVVYGT CHHKKGEARRSRRAVTLPSHSTRKLQTRSQTWLESREYTKH LIKVENWIFRNPGFALVAVAIAWLLGSSTSQKVIYLVMILLI APAYS
Envelope protein	291-790	>sp|Q32ZE1|291-790 IRCIGVSNRDFVEGMSGGTWVDVVLEHGGCVTVMAQDKPT VDIELVTTTVSNMAEVRSYCYEASISDMASDSRCPTQGEAY LDKQSDTQYVCKRTLVDRGWGNGCGLFGKGSLVTCAKFTC SKKMTGKSIQPENLEYRIMLSVHGSQHSGMIGYETDEDRAK VEVTPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLYYLTMN NKHWLVHKEWFHDIPLPWHAGADTGTPHWNNKEALVEFK DAHAKRQTVVVLGSQEGAVHTALAGALEAEMDGAKGRLF SGHLKCRLKMDKLRLKGVSYSLCTAAFTFTKVPAETLHGTV TVEVQYAGTDGPCKIPVQMAVDMQTLTPVGRLITANPVITE STENSKMMLELDPPFGDSYIVIGVGDKKITHHWHRSGSTIGK AFEATVRGAKRMAVLGDTAWDFGSVGGVFNSLGKGIHQIF GAAFKSLFGGMSWFSQILIGTLLVWLGLNTKNGSISLTCLAL GGVMIFLSTAVSA
NS1	791-1142	>sp|Q32ZE1|791-1142 DVGCSVDFSKKETRCGTGVFIYNDVEAWRDRYKYHPDSPR RLAAAVKQAWEEGICGISSVSRMENIMWKSVEGELNAILEE NGVQLTVVVGSVKNPMWRGPQRLPVPVNELPHGWKAWG KSYFVRAAKTNNSFVVDGDTLKECPLEHRAWNSFLVEDHG FGVFHTSVWLKVREDYSLECDPAVIGTAVKGREAAHSDLG YWIESEKNDTWRLKRAHLIEMKTCEWPKSHTLWTDGVEES DLIIPKSLAGPLSHHNTREGYRTQVKGPWHSEELEIRFEECPG TKVYVEETCGTRGPSLRSTTASGRVIEEWCCRECTMPPLSFR AKDGCWYGMEIRPRKEPESNLVRSMVTA
NS2A	1143-1368	>sp|Q32ZE1|1143-1368 GSTDHMDHFSLGVLVILLMVQEGLKKRMTTKIIMSTSMAVL VVMILGGFSMSDLAKLVILMGATFAEMNTGGDVAHLALVA AFKVRPALLVSFIFRANWTPRESMLLALASCLLQTAISALEG

		DLMVLINGFALAWLAIRAMAVPRTDNIALPILAALTPLARG TLLVAWRAGLATCGGIMLLSLKGKGSVKKNLPFVMALGLT AVRVVDPINVVGLLLLTRSGKR
NS2B	1369-1498	>sp|Q32ZE1|1369-1498 SWPPSEVLTAVGLICALAGGFAKADIEMAGPMAAVGLLIVS YVVSGKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDES GDFSLVEEDGPPMREIILKVVLMAICGMNPIAIPFAAGAWYV YVKTGKR
NS3	1499-2115	>sp|Q32ZE1|1499-2115 SGALWDVPAPKEVKKGETTDGVYRVMTRRLLGSTQVGVG VMQEGVFHTMWHVTKGAALRSGEGRLDPYWGDVKQDLV SYCGPWKLDAAWDGLSEVQLLAVPPGERARNIQTLPGIFKT KDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNG SYVSAITQGKREEETPVECFEPSMLKKKQLTVLDLHPGAGK TRRVLPEIVREAIKKRLRTVILAPTRVVAAEMEEALRGLPVR YMTTAVNVTHSGTEIVDLMCHATFTSRLLQPIRVPNYNLNI MDEAHFTDPSSIAARGYISTRVEMGEAAAIFMTATPPGTRD AFPDSNSPIMDTEVEVPERAWSSGFDWVTDHSGKTVWFVPS VRNGNEIAACLTKAGKRVIQLSRKTFETEFQKTKNQEWDFV ITTDISEMGANFKADRVIDSRRCLKPVILDGERVILAGPMPV THASAAQRRGRIGRNPNKPGDEYMYGGGCAETDEGHAHW LEARMLLDNIYLQDGLIASLYRPEADKVAAIEGEFKLRTEQR KTFVELMKRGDLPVWLAYQVASAGITYTDRRWCFDGTTNN TIMEDSVPAEVWTKYGEKRVLKPRWMDARVCSDHAALKS FKEFAAGKR
NS4A	2116-2242	>sp|Q32ZE1|2116-2242 GAALGVMEALGTLPGHMTERFQEAIDNLAVLMRAETGSRP YKAAAAQLPETLETIMLLGLLGTVSLGIFFVLMRNKGIGKM GFGMVTLGASAWLMWLSEIEPARIACVLIVVFLLLVVLIPEP EKQR
2K	2243-2265	>sp|Q32ZE1|2243-2265 SPQDNQMAIIIMVAVGLLGLITA

NS4B	2266-2516	>sp|Q32ZE1|2266-2516 NELGWLERTKNDIAHLMGRREEGATMGFSMDIDLRPASAW AIYAALTTLITPAVQHAVTTSYNNYSLMAMATQAGVLFGM GKGMPFMHGDLGVPLLMMGCYSQLTPLTLIVAIILLVAHY MYLIPGLQAAAARAAQKRTAAGIMKNPVVDGIVVTDIDTM TIDPQVEKKMGQVLLIAVAISSAVLLRTAWGWGEAGALITA ATSTLWEGSPNKYWNSSTATSLCNIFRGSYLAGASLIYTVTR NAGLVKRR
NS5	2517-3419	>sp|Q32ZE1|2517-3419 GGGTGETLGEKWKARLNQMSALEFYSYKKSGITEVCREEA RRALKDGVATGGHAVSRGSAKIRWLEERGYLQPYGKVVDL GCGRGGWSYYAATIRKVQEVRGYTKGGPGHEEPMLVQSY GWNIVRLKSGVDVFHMAAEPCDTLLCDIGESSSSPEVEETRT LRVLSMVGDWLEKRPGAFCIKVLCPYTSTMMETMERLQRR HGGGLVRVPLCRNSTHEMYWVSGAKSNIIKSVSTTSQLLLG RMDGPRRPVKYEEDVNLGSGTRAVASCAEAPNMKIIGRRIE RIRNEHAETWFLDENHPYRTWAYHGSYEAPTQGSASSLVN GVVRLLSKPWDVVTGVTGIAMTDTTPYGQQRVFKEKVDTR VPDPQEGTRQVMNIVSSWLWKELGKRKRPRVCTKEEFINK VRSNAALGAIFEEEKEWKTAVEAVNDPRFWALVDRREREHH LRGECHSCVYNMMGKREKKQGEFGKAKGSRAIWYMWLG ARFLEFEALGFLNEDHWMGRENSGGGVEGLGLQRLGYILEE MNRAPGGKMYADDTAGWDTRISKFDLENEALITNQMEEGH RTLALAVIKYTYQNKVVKVLRPAEGGKTVMDIISRQDQRGS GQVVTYALNTFTNLVVQLIRNMEAEEVLEMQDLWLLRKPE KVTRWLQSNGWDRLKRMAVSGDDCVVKPIDDRFAHALRF LNDMGKVRKDTQEWKPSTGWSNWEEVPFCSHHFNKLYLK DGRSIVVPCRHQDELIGRARVSPGAGWSIRETACLAKSYAQ MWQLLYFHRRDLRLMANAICSAVPVDWVPTGRTTWSIHGK GEWMTTEDMLMVWNRVWIEENDHMEDKTPVTKWTDIPYL GKREDLWCGSLIGHRPRTTWAENIKDTVNMVRRIIGDEEKY MDYLSTQVRYLGEEGSTPGVL

Table_2: Protein length and sequence of ZIKV

7.2. Epitopes predicted from IEDB

Both T-cell and B-cell epitope prediction was done in which the epitopes derived were as follows:

MHC-I: For NS3, 28650 epitopes were derived and for NS5 it amounts to 48303.

MHC-II: For NS3 it amounts to 16281 whereas for NS5 it was 5013 molecules.

All these molecules were screened on the basis of IC50 value. It is the half maximal inhibitory concentration which measures the effectiveness of a substance while inhibiting a particular biological or biochemical function. The IC50 value is predicted in nM and a lower number of it indicates higher affinity thus the value < 50 nM is considered as high, < 500nM as intermediate and < 5000 nM as low. The epitopes may have high, intermediate or low affinity but cannot have more than 5000 nM value. As the output predictions are quantitative so there can be systemic deviations from these experimental values. With respect to IC 50 value of every peptide, a percentile rank is obtained which is generated by comparing the IC50 value against a random set of peptides. The low percentile rank of epitope is considered as low affinity [38].

Thus with these considerations, 15 epitopes from both NS3 and NS5 proteins were screened of MHC-I and II molecules with their molecular weight, half-life of dissociation and Isoelectric point which was derived through ExPASy (Bioinformatics Resource Portal).

NS3 MHC-I Screened epitopes (Table_3):

S. No.	Epitope	Mol. Wt.	Isoelectric Pt.	IC50 value	Half-life of dissociation (sec)
1	SVPAEVWTKY	1179.34	5.72	0.1	1.45
2	AETDEGHAHW	1152.14	4.63	0.15	1.468
3	APTRVVAAEM	1044.23	6.05	0.15	0.908
4	GPMPVTHASA	967.11	6.74	0.15	0.758
5	YQVASAGITY	1072.18	5.52	0.15	1.074
6	QEGVFHTMW	1134.27	5.24	0.15	1.612
7	AEMEEALRG	1005.11	4.25	0.15	1.645
8	VEMGEAAAI	890.02	3.79	0.15	1.082
9	RAWSSGFDW	1111.18	5.84	0.15	2.5

10	GEGRLDPYW	1092.18	4.37	0.15	1.225
11	DVKQDLVSY	1066.18	4.21	0.2	1.322
12	GPWKLDAAW	1043.19	5.84	0.2	2.385
13	FPDSNSPIM	1007.13	3.8	0.2	2.614
14	AEMEEALRGL	1118.27	4.25	0.2	1.765
15	MEDSVPAEVW	1162.41	5.074	0.2	2.576

NS5 MHC-I Screened epitopes (Table_4):

S. No.	Epitope	Mol. Wt.	Isoelectric Pt.	IC50 value	Half-life of dissociation (sec)
1	QEWKPSTGW	1118.21	6	0.1	2.382
2	DENHPYRTW	1217.26	5.32	0.1	1.27
3	RPAEGGKTVM	1045.22	8.75	0.1	1.64
4	RPRVCTKEEF	1264.47	8.22	0.1	1.1
5	EPMLVQSYGW	1209.38	4	0.1	1.36
6	RETACLAKSY	1141.31	8.2	0.15	1.226
7	GEEGSTPGVL	944.99	3.79	0.15	2.598
8	REDLWCGSL	1078.21	4.37	0.15	1.452
9	MSALEFYSYK	1238.42	5.75	0.15	0.231
10	RTWAYHGSY	1140.22	8.6	0.15	1.455
11	ETACLAKSY	985.12	6.09	0.15	1.392
12	WYMWLGARF	1229.46	8.75	0.15	0.796
13	LMANAICSAV	992.22	5.52	0.15	0.908
14	YAQMWQLLY	1215.43	5.52	0.15	1.421
15	YAQMWQLLYF	1362.61	5.52	0.15	1.445

NS3 MHC-II Screened epitopes (Table_5):

S.no.	Epitope	Mol.wt.	Isoelectric point	Percentile_rank	Half-life of dissociation (sec)
1	TDGVYRVMTRRLLGS	1724.23	6.618	0.01	0.654
2	DGVYRVMTRRLLGST	1724.23	6.618	0.01	0.558
3	GVYRVMTRRLLGSTQ	1737.28	6.81	0.01	0.428
4	VYRVMTRRLLGSTQV	1779.36	6.809	0.02	0.323
5	HWLEARMLLDNIYLQ	1915.48	5.909	0.03	0.363
6	WLEARMLLDNIYLQD	1893.42	5.587	0.03	0.363

7	LEARMLLDNIYLQDG	1764.26	5.593	0.03	0.363
8	EARMLLDNIYLQDGL	1764.26	5.593	0.03	0.904
9	ARMLLDNIYLQDGLI	1748.31	5.779	0.03	0.862
10	TTDGVYRVMTRRLLG	1738.26	6.617	0.04	0.513
11	YRVMTRRLLGSTQVG	1737.28	6.81	0.04	0.807
12	PNYNLNIMDEAHFTD	1794.14	5.295	0.07	2.273
13	NYNLNIMDEAHFTDP	1794.14	5.295	0.11	1.851
14	YNLNIMDEAHFTDPS	1767.11	5.313	0.11	1.934
15	GANFKADRVIDSRRC	1708.17	6.609	0.11	1.673

NS5 MHC-II Screened epitopes(Table_6):

S.no.	Epitope	Mol.wt.	Isoelectric point	Percentile_rank	Half-life of dissociation (sec)
1	NEHAETWFLDENHPY	1902.18	5.293	0.01	1.505
2	EHAETWFLDENHPYR	1944.26	5.65	0.01	1.558
3	HAETWFLDENHPYRT	1916.25	5.813	0.01	1.423
4	AETWFLDENHPYRTW	1965.32	5.699	0.01	1.532
5	ETWFLDENHPYRTWA	1965.32	5.699	0.01	1.324
6	TWFLDENHPYRTWAY	1999.38	5.862	0.01	1.324
7	WFLDENHPYRTWAYH	2035.42	5.991	0.01	1.316
8	SGQVVTYALNTFTNL	1628.04	5.741	0.02	1.498
9	GQVVTYALNTFTNLV	1640.1	5.76	0.02	0.921
10	QVVTYALNTFTNLVV	1682.18	5.759	0.02	0.9
11	VVTYALNTFTNLVVQ	1682.18	5.759	0.02	1.371
12	VTYALNTFTNLVVQL	1696.21	5.761	0.02	1.42
13	AENIKDTVNMVRRII	1772.32	6.363	0.03	0.688
14	NEHAETWFLDENHPY	1902.18	5.923	0.03	1.505
15	KVRSNAALGAIFEEE	1634.03	5.911	0.03	0.623

7.3. Screened epitopes

Out of all the epitopes screened for NS3 and NS5 only 10 in total were selected on the basis of low IC50 value and Isolectric point.

IC50 value: It is the half-maximal inhibitory concentration which measures the effectiveness of a molecule while inhibiting any biological or biochemical function. It can be use to compare the potency of two antagonist. The lower number of value defines higher affinity [38].

Isoelectric point: This term is defined as the pH of peptide where the net charge of it becomes zero. At the pH values above pI, the peptide carries net negative charge and below pI carries net positive charge [39].

Half life of dissociation: It is the biological half-life or terminal half-life of a substance and is denoted by abbreviation $t^{1/2}$.

Screened NS3 protein Epitopes(Table_7):

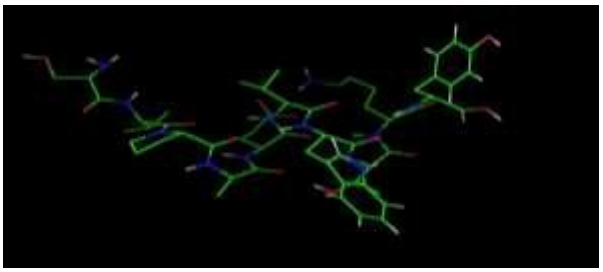
S.no.	Epitope	Mol.wt.	Isoelectric Pt.	IC50 value	Half-life of dissociation (sec)	MHC-I/ II
1	SVPAEVWTKY	1179.34	5.72	0.1	1.45	MHC-I
2	MEDSVPAEVW	1162.41	5.074	0.2	2.576	MHC-I
3	AEMEEALRG	1005.11	4.25	0.15	1.645	MHC-I
4	VEMGEAAAI	890.02	3.79	0.15	1.082	MHC-I
5	DGVYRVMTRRLLGST	1724.23	6.618	0.01	0.558	MHC-II
6	GVYRVMTRRLLGSTQ	1737.28	6.81	0.01	0.428	MHC-II
7	VYRVMTRRLLGSTQV	1779.36	6.809	0.02	0.323	MHC-II
8	LEARMLLDNIYLQDG	1764.26	5.593	0.03	0.363	MHC-II
9	ARMLLDNIYLQDGLI	1748.31	5.779	0.03	0.862	MHC-II
10	PNYNLNIMDEAHFTD	1794.14	5.295	0.07	2.273	MHC-II

Screened NS5 Epitopes(Table_8):

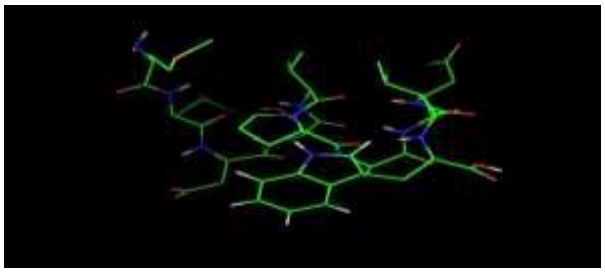
S.no.	Epitope	Mol.wt.	Isoelectric Pt.	IC50 value	Half-Life of dissociation (sec)	MHC-I/II
1	YAQMWQLLY	1215.43	5.52	0.15	1.421	MHC-I
2	RETACLAKSY	1141.31	8.2	0.15	1.226	MHC-I
3	ETACLAKSY	985.12	6.09	0.15	1.392	MHC-I
4	NEHAETWFLDENHPY	1902.18	5.293	0.01	1.505	MHC-II
5	ETWFLDENHPYRTWA	1965.32	5.699	0.01	1.324	MHC-II
6	TWFLDENHPYRTWAY	1999.38	5.862	0.01	1.324	MHC-II
7	WFLDENHPYRTWAYH	2035.42	5.991	0.01	1.316	MHC-II
8	SGQVVTYALNTFTNL	1628.04	5.741	0.02	1.498	MHC-II
9	GQVVTYALNTFTNLV	1640.1	5.76	0.02	0.921	MHC-II
10	QVVTYALNTFTNLVV	1682.18	5.759	0.02	0.9	MHC-II

7.4. PEP-FOLD structure prediction

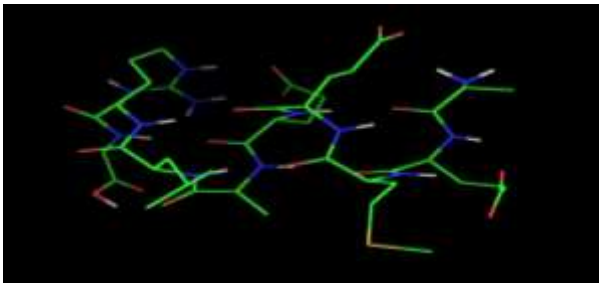
Using PEP-FOLD database, NS3 and NS5 protein epitopes structure are predicted which are as follows:



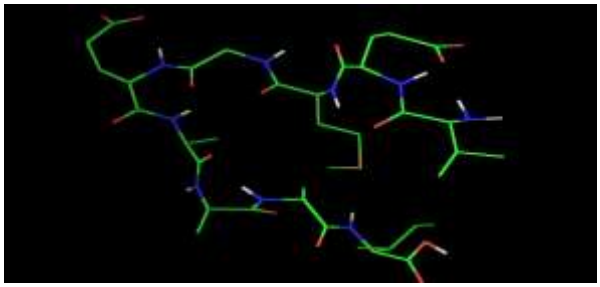
NS3 Epitope 1: SVPAEVWTKY



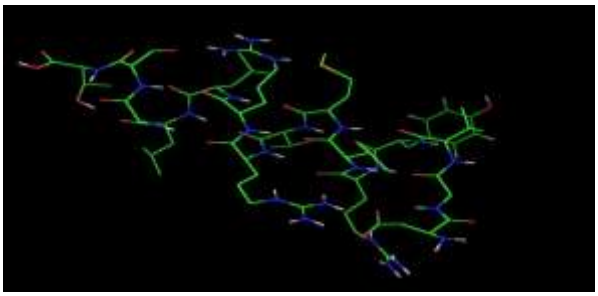
NS3 Epitope 2: MEDSVPAEVW



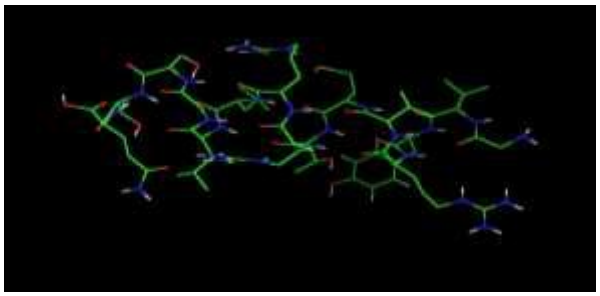
NS3 Epitope 3: AEMEEALRG



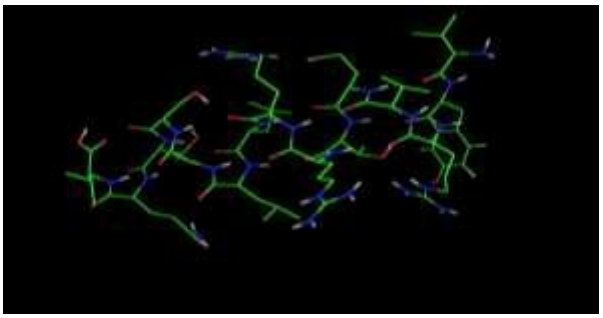
NS3 Epitope 4: VEMGEAAAI



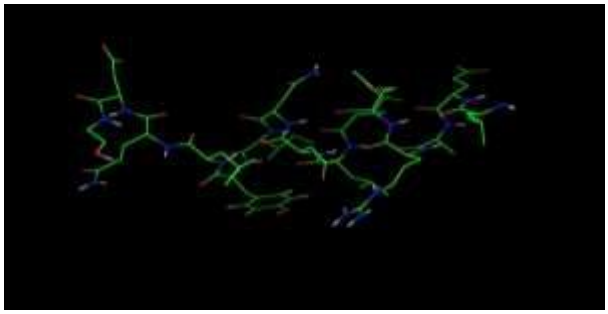
NS3 Epitope 5: DGVYRVMTRRLLGST



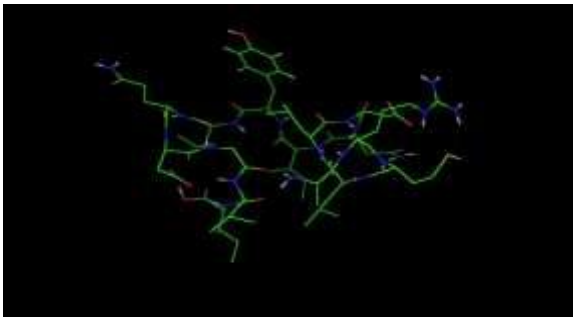
NS3 Epitope 6: GVYRVMTRRLLGSTQ



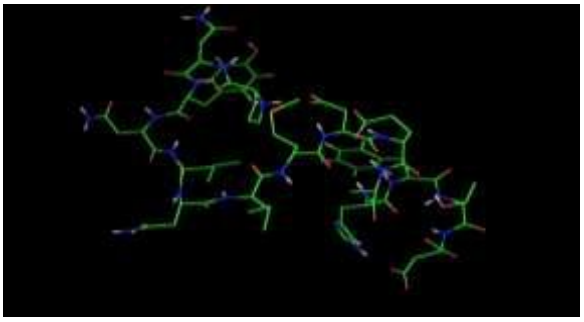
NS3 Epitope 7: VYRVMTRRLLGSTQV



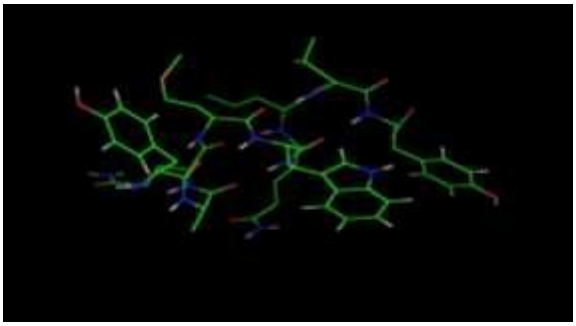
NS3 Epitope 8: LEARMLLDNIYLQDG



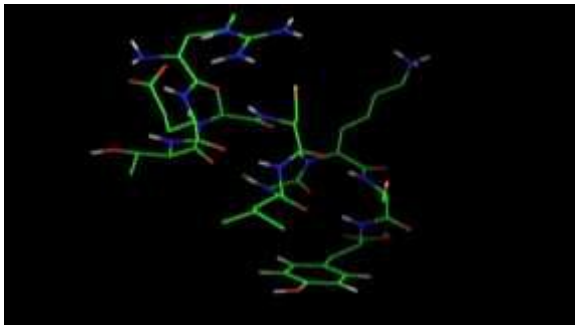
NS3 Epitope 9: ARMLLDNIYLQDGLI



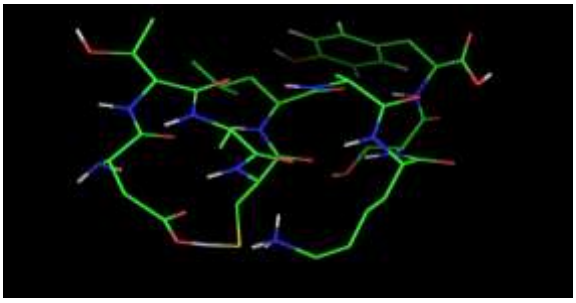
NS3 Epitope 10: PNYNLNIMDEAHFTD



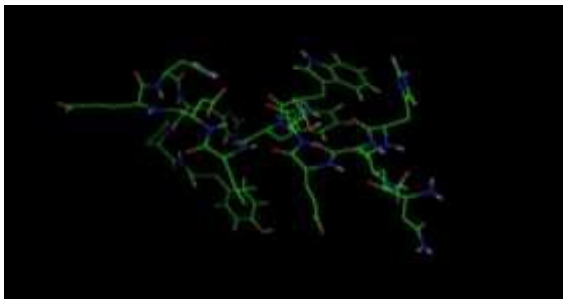
NS5 Epitope 1: YAQMWQLLY



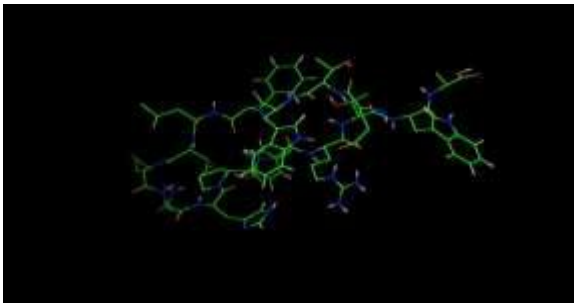
NS5 Epitope 2: RETACLAKSY



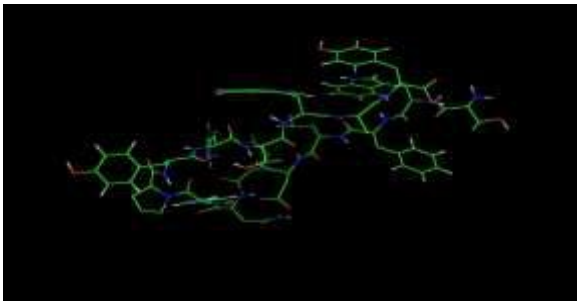
NS5 Epitope 3: ETACLAKSY



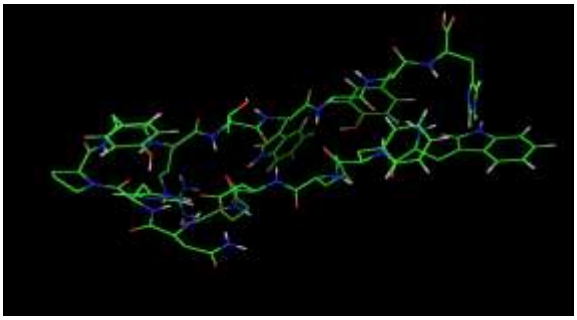
NS5 Epitope 4: NEHAETWFLDENHPY



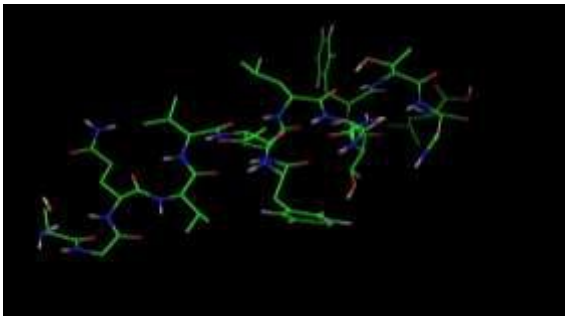
NS5 Epitope 5: ETWFLDENHPYRTWA



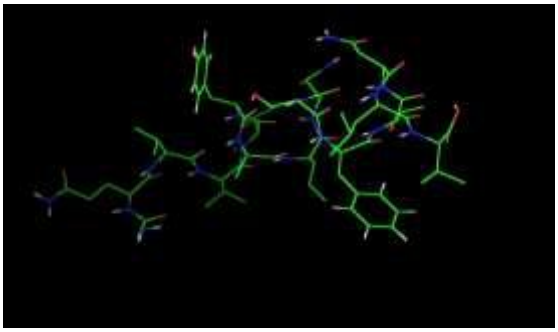
NS5 Epitope 6: TWFLDENHPYRTWAY



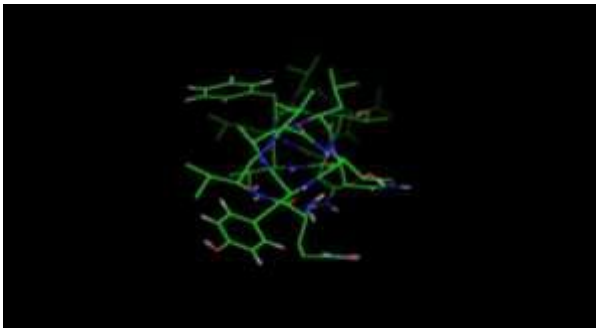
NS5 Epitope 7: WFLDENHPYRTWAYH



NS5 Epitope 8: SGQVVTYALNTFTNL



NS5 Epitope 9: GQVVTYALNTFTNLV



NS5 Epitope 10: QVVTYALNTFTNLVV

7.5. Docking Results

7.5.1. PatchDock Results

Through PatchDock, a docking score was determined which further helps in screening of the suitable epitope. Thus, the ACE obtained provides negative binding energy which helps us to determine the consistent candidate.

Docking score of NS3 peptides with MHC-I and MHC-II are:

S.no	Epitope	Mol.wt.	Isoelectric Pt.	IC50 value	Half-life of dissociation (sec)	MHC-I/ II	MHC-I score	MHC-II score
1	SVPAEVWTKY	1179.34	5.72	0.1	1.45	MHC-I	10230	8212
2	MEDSVPAEVW	1162.41	5.074	0.2	2.576	MHC-I	9992	8522
3	AEMEEALRG	1005.11	4.25	0.15	1.645	MHC-I	8338	7628
4	VEMGEAAAI	890.02	3.79	0.15	1.082	MHC-I	8312	8312
5	DGVYRVMTRRLLGST	1724.23	6.618	0.01	0.558	MHC-II	10316	10516
6	GVYRVMTRRLLGSTQ	1737.28	6.81	0.01	0.428	MHC-II	10446	10654
7	VYRVMTRRLLGSTQV	1779.36	6.809	0.02	0.323	MHC-II	10772	10796
8	LEARMLLDNIYLQDG	1764.26	5.593	0.03	0.363	MHC-II	12080	11602

9	ARMLLDNIYLQDGLI	1748.31	5.779	0.03	0.862	MHC-II	9926	11544
10	PNYNLNIMDEAHFTD	1794.14	5.295	0.07	2.273	MHC-II	10710	10992

Docking score of NS5 peptides with MHC-I and MHC-II are:

S.no.	Epitope	Mol.wt.	Isoelectric Pt.	IC50 value	Half-Life of dissociation (sec)	MHC-I/II	MHC-I score	MHC-II score
1	YAQMWQLLY	1215.43	5.52	0.15	1.421	MHC-I	9482	8832
2	RETACLAKSY	1141.31	8.2	0.15	1.226	MHC-I	9672	7980
3	ETACLAKSY	985.12	6.09	0.15	1.392	MHC-I	8204	7780
4	NEHAETWFLDENHPY	1902.18	5.293	0.01	1.505	MHC-II	11230	11020
5	ETWFLDENHPYRTWA	1965.32	5.699	0.01	1.324	MHC-II	10626	11512
6	TWFLDENHPYRTWAY	1999.38	5.862	0.01	1.324	MHC-II	11594	10722
7	WFLDENHPYRTWAYH	2035.42	5.991	0.01	1.316	MHC-II	11502	11474
8	SGQVVTYALNTFTNL	1628.04	5.741	0.02	1.498	MHC-II	11660	10520
9	GQVVTYALNTFTNLV	1640.1	5.76	0.02	0.921	MHC-II	10728	10070
10	QVVTYALNTFTNLVV	1682.18	5.759	0.02	0.9	MHC-II	11284	11298

7.5.2. AutoDock Results

From the results of PatchDock, docking results of NS3 and NS5 epitopes with both MHC-I and MHC-II are screened on the basis of high docking score, IC50 value and half life of dissociation. From both NS3 and NS5 epitopes, the best screened epitope was of NS3 i.e. **VEMGEAAAI** with molecular weight of 890.02. The results obtained were more refined which provides results with more parameters like binding energy, ligand efficiency, intermol energy, vdW + H-bond + desolv Energy, Electrostatic energy, Total Internal Energy, Torsional energy, Unbound energy, refRMS and no. of H-bonds formed.

Binding Energy: is defined as the sum of the intermolecular energy and the torsional free-energy penalty.

refRMS: is defined as the rms difference within the present conformation coordinates and present reference structure.

clRMS: is also defined as the rms difference within the present conformation and the lowest energy conformation in the molecule cluster.

torsional_energy: is defined as the number of active torsions * .3113 [it is defined as autodock parameter].

rseed1 and rseed2: are defined as the precise arbitrary amount of seeds used for existing conformation's docking run.

Docking Energy: is defined as the sum of the ligand's internal and intermolecular energies [40].

Ligand Efficiency: it measures the molecular properties like lipophilicity, size of tiny molecules that gains binding affinity to target a drug. Its application is in the assortment and optimization of fragments [41].

AutoDock Results with MHC-I of the ligand with 10 different conformations:

S. no.	Binding_energy	ligand_efficiency	intermol_energy	desolv_energy	electrostatic_energy	total_internal	torsional_energy	unbound_energy	refRMS	H-bonds
1	6.43	0.11	-2.82	-3.81	0.99	-17.59	9.25	-17.59	33.07	1
2	2.51	0.04	-6.74	-6.83	0.09	-14.74	9.25	-14.74	30.54	2
3	3.87	0.06	-5.38	-4.88	-0.5	-12.68	9.25	-12.68	29.82	1
4	3.52	0.06	-5.73	-6.15	0.42	-14.43	9.25	-14.43	27.61	3
5	2.76	0.05	-6.49	-6.89	0.4	-13	9.25	-13	28.9	2
6	4.5	0.07	-4.74	-3.47	-1.28	-14.19	9.25	-14.19	25.69	2
7	3.87	0.06	-5.37	-4.17	-1.2	-13.16	9.25	-13.16	27.82	1
8	3.72	0.06	-5.53	-4.12	-1.41	-13.32	9.25	-13.32	31.62	1
9	0.23	0	-9.02	-9.56	0.54	-12.77	9.25	-12.77	31.86	2
10	0.92	0.02	-8.33	-7.48	-0.85	-13.15	9.25	-13.15	31.9	2

Autodock Results with MHC-II of the ligand with 10 different conformations:

S. no.	Binding_energy	ligand_efficiency	intermol_energy	desolv_energy	electrostatic_energy	total_internal	torsional_energy	unbound_energy	refRMS	H-bond
1	1.77	0.03	-7.47	-6.55	-0.93	-13.35	9.25	-13.35	38.18	4
2	0.1	0	-9.14	-7.56	-1.58	-11.36	9.25	-11.36	32.25	2
3	2.83	0.05	-6.42	-4.91	-1.51	-13.72	9.25	-13.72	38.74	2

4	3.43	0.06	-5.82	-5.48	-0.34	-13.18	9.25	-13.18	33.91	1
5	4.1	0.07	-5.14	-4.95	-0.2	-16.35	9.25	-16.35	30.56	no
6	2.47	0.04	-6.77	-5.15	-1.62	-13.37	9.25	-13.37	33.03	2
7	2.01	0.03	-7.23	-6.47	-0.76	-13.2	9.25	-13.2	30.84	3
8	3.75	0.06	-5.49	-5.08	-0.42	-13.45	9.25	-13.45	33.17	no
9	0.22	0	-9.03	-7.35	-1.68	-9.57	9.25	-9.57	34.02	3
10	2.08	0.03	-7.16	-5.69	-1.47	-14.95	9.25	-14.95	31.34	2

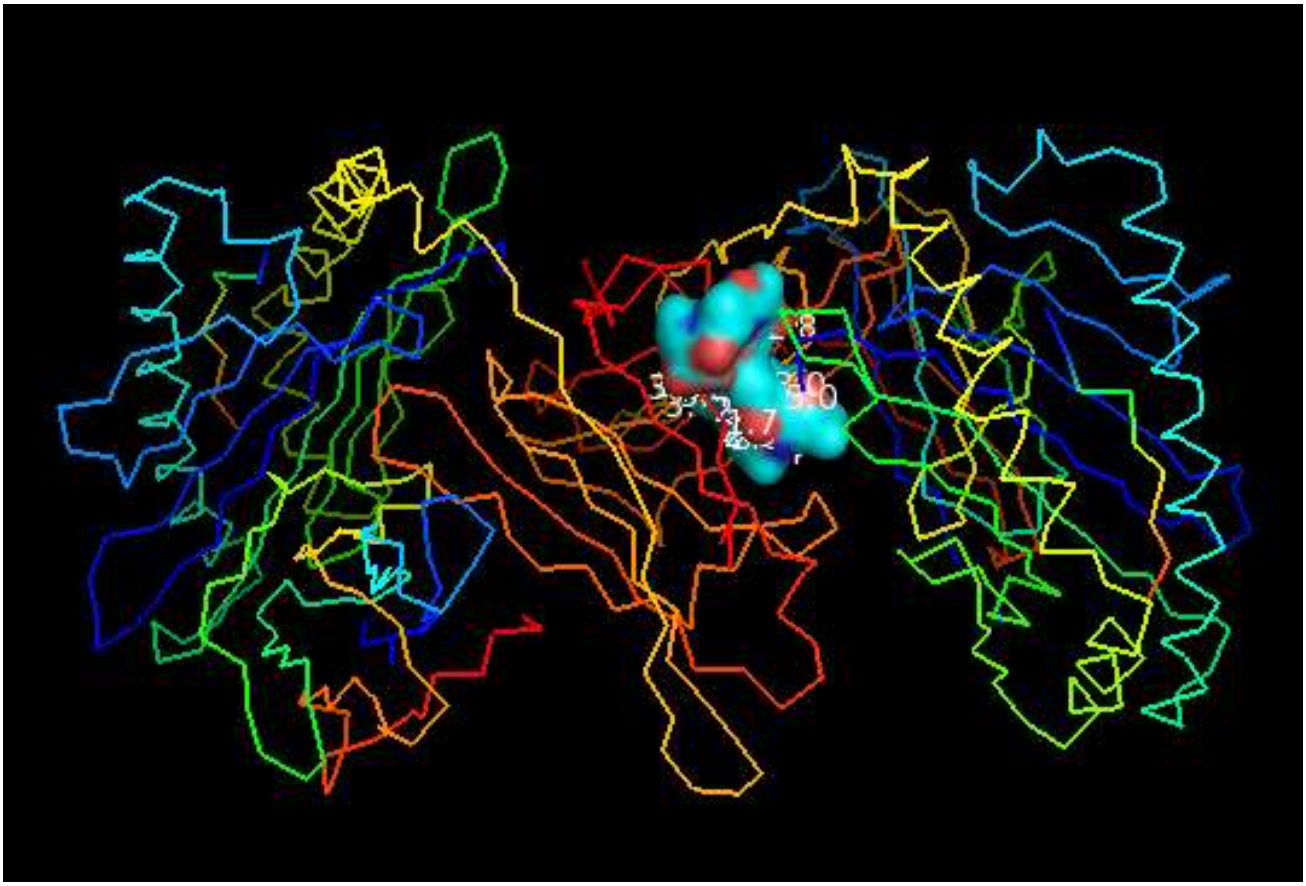


Image representing docking through AutoDock of VEMGEAAAI with MHC-I complex

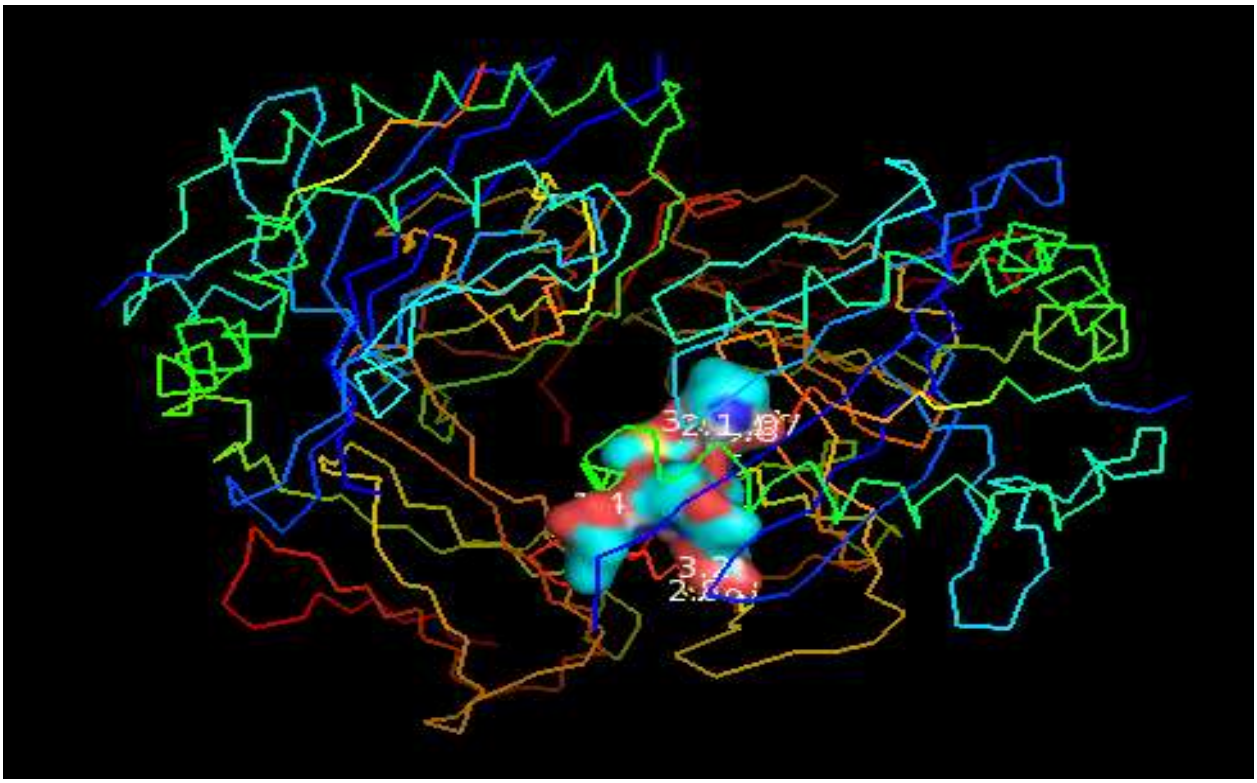
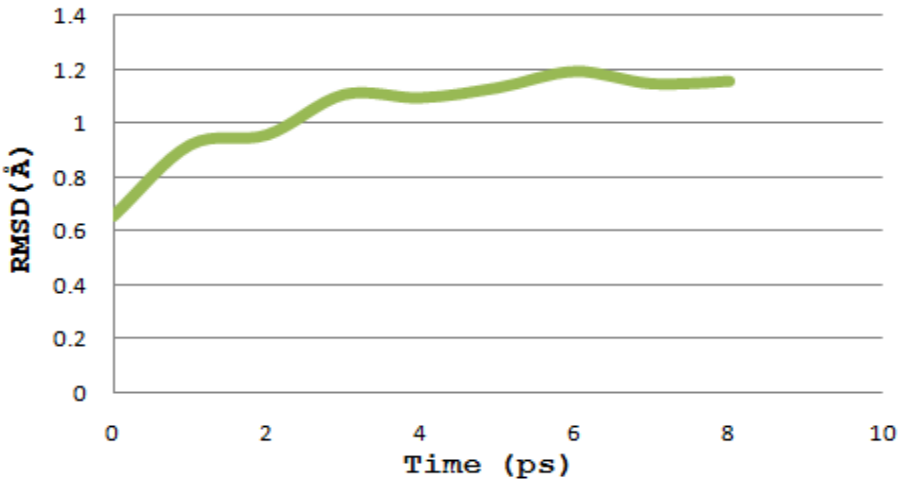


Image representing docking through AutoDock of VEMGEAAAI with MHC-II complex

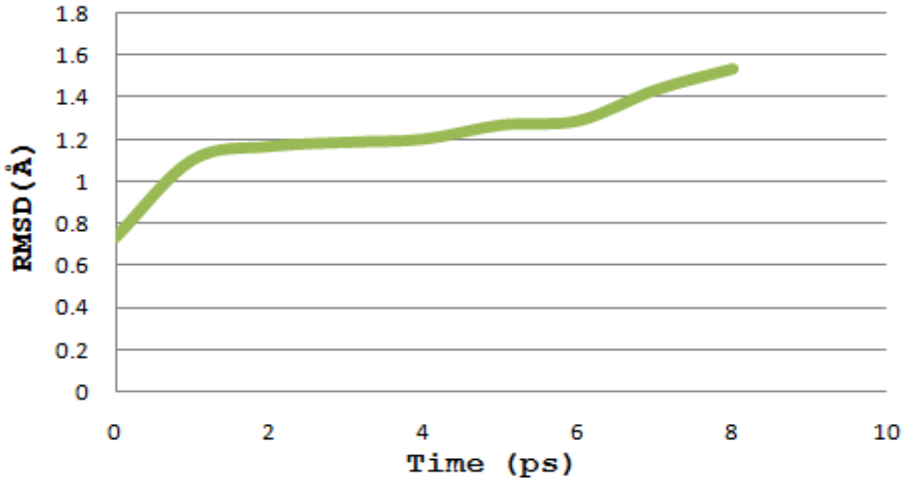
7.6. Molecular dynamics simulation results:

Molecular dynamics simulations were done using the NAMD (Nanoscale Molecular Dynamics program; v2.7) graphical interface module [36] incorporated visual molecular dynamics (VMD 1.9.2) [37]**.** The protein-ligand complex was immersed in the center of a 50 Å box of water molecules where all water molecule atoms (H-O-H) were closer than 1.5 Å and a CHARMM (Chemistry at Harvard Macromolecular Mechanics) 22 parameter file for proteins and lipids; phi and psi cross-term map correction were used in the force field for complexes. For the minimization and equilibration of complex in the water box, we assumed force-field parameters excluding scaling of 1.0 Å and a cutoff of Coulomb forces with a switching function starting at 12 Å, reaching zero at a distance of 10 Å, ending at 14 Å with a margin of 3.0 Å, and all atoms, including those of hydrogen, were illustrated explicitly. A protein structure file (psf) stores structural information of the protein, such as various types of bonding interactions. The psf was created from the initial pdb and topology files using psfgen package of VMD. After running psfgen, two new files were generated protein pdb and protein psf and by accessing PSF and PDB

files; NAMD generated the trajectory DCD file. After the simulations, the results were analyzed in VMD by calculating the Root mean square deviation (RMSD) of the complex using rmsd tcl source file from the Tk console and finally rmsd.dat was saved and accessed in Microsoft office excel 2007.



Graph displaying root mean square deviation (RMSD) in relation to time (5 ps) at 310 K for NAMD-VMD simulation of VEMGEAAAI and MHC-I complex, resulted in highest peak at 1.15 Å.



Graph displaying root mean square deviation (RMSD) in relation to time (5 ps) at 310 K for NAMD-VMD simulation of VEMGEAAAI and MHC-II complex, resulted in highest peak at 1.54 Å.

CHAPTER 7

CONCLUSION

OF

RESEARCH

Zika virus is a mosquito-borne flavivirus which is now-a-days focus of an ongoing epidemic and public health. ZIKV causes asymptomatic infection as well as presents meek symptoms in infected patients. Though, vaccination for it is a civic health concern due to its congenital and neuropathological abnormalities. Thus, the general objective of the research is to design a peptide based vaccine for Zika Virus using in-silico methods, thereby predicting and screening different epitopes and to study the applications of Bioinformatics techniques used for Zika Virus. Both B as well as T- cell epitopes was predicted for ZIKV NS3 and NS5 proteins. However, 15 epitopes of both MHC-I and II proteins were screened regarding some parameters by predicting their structures through PEP-FOLD and then were docked with MHC-I and II complexes. The stability of the resulting peptide with MHC I and II complexes was further studied by NAMD and VMD molecular dynamics simulations. The simulation results underline the limits of rigid-body docking approaches. A number of the antigenic epitopes predicted and screened in this research. In contrast to these epitopes, VEMGEAAAI was found to be the high affinity binding epitopes for MHC Class I and MHC Class II complexes. Graph displaying root mean square deviation in relation to time (5 ps) at 310 K for NAMD-VMD simulation of VEMGEAAAI with MHC-I and MHC-II complex, resulted in highest peak at 1.15 Å and 1.54 Å respectively. These findings conclude that the designed epitope can be used in wet laboratory formulations of vaccines against ZIKV.

CHAPTER 8

BIBLIOGRAPHY

References

1. Buathong R, Hermann L, Thaisomboonsuk B, Rutvisuttinunt W, Klungthong C, Chinnawirotpisan P, Manasatienkij W, Nisalak A, Fernandez S, Yoon IK, Akrasewi P, Plipat T(2012-2014). Detection of Zika Virus Infection in Thailand, Am. J. Trop. Med. Hyg., 93(2), 2015, pp. 380–383
2. Tappe D, Nachtigall S, Kapaun A,Schnitzler P, Günther S, Schmidt-Chanasit J (September,2014). Acute Zika Virus Infection after Travel to Malaysian Borneo, DOI: http://dx.doi.org/10.3201/eid2105.141960.
3. Alera MT, Hermann L, Tac-An IA, Klungthong C, Rutvisuttinunt W, Manasatienkij W, Villa D, Thaisomboonsuk B, Velasco JM, Chinnawirotpisan P, Lago CB, Roque VG, Jr., Macareo LR, Srikiatkhachorn A, Fernandez S, Yoon IK (2012). Zika Virus Infection, Philippines, DOI: http://dx.doi.org/10.3201/eid2104.141707.
4. Diallo D, Sall AA, Diagne CT, Faye O, Faye O, Ba Y, Hanley KA, Buenemann M, Weaver SC, Diallo M (2011). Zika Virus Emergence in Mosquitoes in Southeastern Senegal, doi:10.1371/journal.pone.0109442.
5. Gourinat AC, Connor OO, Calvez E, Goarant C, Rouzeyrol MD (2015) Detection of Zika Virus in Urine, DOI: http://dx.doi.org/10.3201/eid2101.140894.
6. 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, Yap State, Micronesia, 2007. Infect Dis. 2008; 14:1232–9. DOI: http://dx.doi.org/10.3201/eid1408.080287
7. Kuno G, Chang GJ. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses.Arch Virol.2007; 152:687–96. DOI: http://dx.doi.org/10.1007/s00705-006-0903-z
8. Gatherer D, Kohl A. Zika virus: a previously slow pandemic spreads rapidly through the Americas. J Gen Virol. 2016;97:269– 73. http://dx.doi.org/10.1099/jgv.0.000381
9. 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. http://dx.doi.org/ 10.1371/journal.pntd.0002636

10. 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. http://dx.doi.org/10.1128/ JVI.00354-15
11. 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. http://dx.doi.org/10.1099/ 0022-1317-69-8-1913
12. http://www.zikavirusnet.com/zika-virus.html
13. https://microbewiki.kenyon.edu/index.php/Zika_virus
14. Henderson BR, Saeedi BJ, Campagnola G, Geiss BJ. "Analysis of RNA binding by the Dengue virus NS5 RNA capping enzyme". In Jeang, K.T. PLoS ONE 2011; 6 (10).
15. Gaunt M.W., Sall A.A., de Lamballerie X., Falconar A.K.I., Dzhivanian T.I., Gould E.A. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. J. Gen. Virol. 2001; 82:1867-1876.
16. Kuno G, Chang GJ. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Arch Virol, 2007; 152: 687–696.
17. http://viralzone.expasy.org/all_by_species/43.html
18. Dick GW, Kitchen SF, Haddow AJ. Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg. 1952;46:509– 20. http://dx.doi.org/10.1016/0035-9203(52)90042-4
19. Diallo D, Sall AA, Diagne CT, Faye O, Faye O, Ba Y, et al. Zika virus emergence in mosquitoes in southeastern Senegal, 2011. PLoS One. 2014;9:e109442. http://dx.doi.org/10.1371/ journal.pone.0109442
20. Ledermann JP, Guillaumot L, Yug L, Saweyog SC, Tided M, Machieng P, et al. Aedes hensilli as a potential vector of Chikungunya and Zika viruses. PLoS Negl Trop Dis. 2014;8:e3188. http://dx.doi.org/10.1371/journal.pntd.0003188
21. Ioos S, Mallet HP, Leparc Goffart I, Gauthier V, Cardoso T, Herida M. Current Zika virus epidemiology and recent epidemics. Med Mal Infect. 2014;44:302–7. http://dx.doi.org/ 10.1016/j.medmal.2014.04.008
22. Oliveira Melo AS, Malinger G, Ximenes R, Szejnfeld PO, Alves Sampaio S, Bispo de Filippis AM. Zika virus intrauterine infection causes fetal brain abnormality and

microcephaly: tip of the iceberg? Ultrasound Obstet Gynecol. 2016;47:6–7. http://dx.doi.org/10.1002/uog.1583135.

23. Simpson DI. Zika virus infection in man. Trans R Soc Trop Med Hyg. 1964;58:335–8. http://dx.doi.org/10.1016/ 0035-9203(64)90201-9
24. Besnard M, Lastere S, Teissier A, Cao-Lormeau V, Musso D. Evidence of perinatal transmission of Zika virus, French Polynesia, December 2013 and February 2014. Euro Surveill. 2014;19:20751. http://dx.doi. org/10.2807/1560-7917.ES2014.19.13.20751
25. 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. http://dx.doi.org/10.1056/NEJMoa0805715
26. Oehler E, Watrin L, Larre P, Leparc-Goffart I, Lastere S, Valour F, et al. Zika virus infection complicated by Guillain-Barre syndrome—case report, French Polynesia, December 2013. Euro Surveill. 2014;19:20720. http://dx.doi.org/10.2807/1560-7917. ES2014.19.9.20720
27. Fauci AS, Morens DM. Zika virus in the Americas—yet another arbovirus threat. N Engl J Med. 2016;374:601–4. http://dx.doi.org/ 10.1056/NEJMp1600297
28. Dupont-Rouzeyrol M, O'Connor O, Calvez E, Daurès M, John M, Grangeon JP, et al. Co-infection with Zika and dengue viruses in 2 patients, New Caledonia, 2014. Emerg Infect Dis. 2015;21:381–2. http://dx.doi.org/10.3201/eid2102.141553
29. Centers for Disease Control and Prevention. Zika virus disease in the United States, 2015–2016. 2016 Feb 10 [cited 2016 Feb 17]. http://www.cdc.gov/zika/geo/united-states.html
30. Bollati M., Alvarez K., Assenberg R., et.al. Structure and functionality in flavivirus NS-proteins: Perspectives for drug design, Antiviral Research 87(2010)125-148.DOI: dx.doi.org/10.1016/j.antiviral.2009.11.009.
31. www.uniprot.org/help/uniprotkb
32. Vita R.,Zarebski L., Greenbaum J.A., Emami H., Hoof I.,Salimi N.,Damle R.,Sette A.,Peters B., The Immune Epitope Database 2.0. D854–D862 Nucleic Acids Research, 2010, Vol. 38, Database issue doi:10.1093/nar/gkp1004.

33. Maupetit J.,Derreumaux P.,Tuffery P., PEP-FOLD: an online resource for de novo peptide structure prediction, W498–W503 Nucleic Acids Research, 2009, Vol. 37, Web Server issue.doi:10.1093/nar/gkp323
34. https://bioinfo3d.cs.tau.ac.il/PatchDock/patchdock.html
35. http://autodock.scripps.edu/
36. Phillips JC.,Braun R.,Wang W.,Gumbart J et.al., Scalable Molecular Dynamics with NAMD, J Comput Chem. 2005 December ; 26(16): 1781–1802.
37. Humphrey W.,Dalke A.,Schulten K., VMD: visual molecular dynamics, J Mol Graph. 1996 Feb;14(1):33-8, 27-8, PMID:8744570
38. http://tools.iedb.org/mhci/help/#Method
39. http://www.biosyn.com/faq/what-is-isoelectric-point.aspx
40. http://autodock.scripps.edu/faqs-help/tutorial/using-autodock-with-autodocktools/UsingAutoDockWithADT_v2e.pdf
41. http://real.mtak.hu/12783/1/LE_NRDD%20final.pdf