Computational analysis in designing HLA restricted T-cell epitopes enriched peptides of dengue envelope protein

A Dissertation

Submitted in the partial fulfillment of the requirement for

the award of the degree of

Master of Science

In

Biochemistry

Submitted By:

Neha Kaushal (301707015)

School of chemistry and biochemistry

Submitted To:

Dr. Manoj Baranwal

Department of Biotechnology

Thapar institute of engineering and technology, Patiala (Punjab) -147004

JULY 2019

Candidate declaration

I, hearby declare that the work presented in the thesis entitled **"Computational analysis in designing HLA restricted T-cell epitopes enriched peptides of dengue envelope protein"** in partial fulfillment of the award of the degree of Master of Science in Biochemistry, School of chemistry and Biochemistry , Thapar institute of engineering and technology, Patiala. This is an authentic record of my work during the period of six months from Jan, 2019 to July, 2019 under the guidance of Dr. Manoj Baranwal, Associate Professor, Department of Biotechnology, Thapar institute of engineering and technology, Patiala. I have not submitted the matter embodied in the thesis for award of any other degree or diploma.

Place: Patiala

Date: 15/7/19

Neha Kaushal

(301707015)

Certificate

This is to certify that the thesis entitled **"Computational analysis in designing HLA restricted T-cell epitopes enriched peptides of dengue envelope protein"** submitted by Neha Kaushal in partial fulfillment of the award of the degree of Master of Science in Biochemistry, School of chemistry and Biochemistry, Thapar institute of engineering and technology, Patiala in record of student's own work carried out by her. The report has not 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, TIET

Patiala

Acknowledgement

Working in the Department of Biotechnology at Thapar institute of engineering and technology has been a golden opportunity for me. I am so grateful for having a chance to met wonderful people and professionals who led me through this six month period. It is my radiant sentiment to place on record my best regards, deepest sense of gratitude to **Dr. Manoj Baranwal** for his careful and precious guidance which were extremely valuable for my study both theoretically and practically.

I perceive as this opportunity as a big milestone in my career development. I will strive to use gained skills and knowledge in the best possible way and I will continue to work on their improvement, in order to attain desired career objectives. Hope to continue cooperation with all of you in the future.

I also express my heartiest gratitude to **Dr. Amjad Ali** (Head and Professor, School of Chemistry and Biochemistry) for his support throughout the period and all the members of School of Chemistry and Biochemistry for their help and suggestions at different stages of this work.

I invariably fall short of words to express my heart-felt gratitude and profound thanks to research scholars **Mr. Sahil Jain**, **Mrs. Yogita Gupta Kautish** and **Mrs. Neha Srivastava** for their valuable suggestions, patience and inspiring discussions. I wish to express my thanks to them for their cooperation and help during the course of my research work.

I would like to thank my lab mates **Ms. Baneet Chawala**, **Ms. Khushpreet Kaur** and **Ms. Neha Garg** for their moral support, love and kindness for timely help.

I would like to express my love to my sister, **Ms. Sonali Puri**, who inspires me everyday to be strong and keep me going with her constant support and love.

I would like to express my deepest gratitude to my beloved parents, **Mr. Rajesh Kaushal** and **Mrs. Suman Kaushal** for providing me the best education and inspiring me so that I can accomplish my dreams.

Table of Contents

Abstract

Dengue virus (DENV) envelope protein is considered as an important target for vaccine development as it exhibits receptor binding sites responsible for viral cell attachment and entry. In this study, immunoinformatics approach was applied to design a vaccine peptide candidate containing multiple HLA restricted T cell epitopes against dengue virus envelope protein of serotype II (most infectious serotype). Six T cell prediction tools were used to find peptides having $CD8^+$ and $CD4^+$ T cell epitopes in nine conserved fragments (≥ 70 %) obtained after conservancy analysis, followed by identification overlapping $CD8^+$ and $CD4^+$ T cell epitopes. The peptides containing both $CD8^+$ and $CD4^+$ T cells were screened for screening allergic, autoimmune and toxic response. Molecular docking by molecular virtual docker and CABS dock revealed that the identified peptides have strong binding affinities and RMSD values with most of the HLA class I and II molecules. IEDB population coverage was done across various geographical regions. Four out of five peptides have shown an average population coverage > 92% across various regions. Hence, it is suggested that these peptides may be considered as potential candidates for synthetic peptide vaccine against DENV.

Keywords: Dengue virus, HLA, Molecular virtual docker, CABS-dock, Binding affinities, RMSD.

List of tables

List of figures

1. Introduction

Dengue virus (DENV), member of *Flavivirus* genus, has an enveloped, single stranded positive sense RNA (Diamond et al., 2015; Wahala et al, 2011). DENV has four serotypes (DENV-1, DENV-2, DENV-3 and DENV-4), which differ by 25-40% in their amino acid sequence (Diamond et al., 2015). All the four DENV serotypes infect humans (Shrivastava et al., 2018). The DENV-2 is most infectious serotype as it is chiefly responsible for causing dengue hemorrhage fever (DHF) in majority of outbreaks (Fried et al., 2010).

DENV infection is an arthropod borne, acute systemic infection transmitted by *Aedes Aegypti* [primary vector] and *Aedes albopictus* [secondary vector] (Guo et al., 2017; Diamond et al., 2015; Bhatt et al, 2013; Rajpakse et al., 2012). It is a global health concern predominantly in tropical and sub-tropical regions with 390 million infections per year (El Sahili et al., 2017; Diamond et al., 2015). Because of spread of mosquito vector, urbanization and global air travel, it is proliferating at an alarming rate (Govindrajan et al., 2015; Urcuqui-Urcuqui-Inchima et al., 2010). DENV infection ranges from dengue fever to severe dengue fever. Dengue fever (DF) is manifested by headache, severe pain in joints and rashes which usually persist for 7-14 days and severe dengue is manifested by capillary leakage, thrombocytopenia and mild to severe liver injury (Diamond et al., 2015).

Predominantly DENV infection occurs annually in America, Asia and Australia (Diamond et al., 2015). In 1953, the first outbreak of dengue fever was reported in Southeast Asia, Africa and North America. DENV infection was transmitted from Southeast Asia to Malaysia, subtropical and tropical asian countries, Philippines, the Indian subcontinent, Sri Lanka, southern China, northeastern Australia, southern Taiwan and pacific islands. In India, DENV infection is more prevalent in Punjab, Haryana, Rajasthan, Gujarat, Dadra and Nagar Haveli, Kerala and Puducherry with an average incidence rate of > 50 per million populations (Mutheneni et al., 2017). Only Europe and Antarctica continents do not experience dengue fever outbreaks (WHO, 2019).

The DENV has three structural proteins: Capsid (C), Membrane (prM/M) and Envelope (E) and seven non-structural proteins: NS1, NS2A, NS2B, NS3, NS4, NS4B and NS5 (Junglen et al., 2016). E protein is responsible for response and virulence and is the primary target of neutralizing and amplifying antibodies (Hsieh et al., 2010; Mukhopadhyay et al., 2005; Urcuqui-Inchima et al., 2010). It forms an outer shell of virus and has two transmembrane helices, a receptor binding site, a fusion peptide. Upon virus invasion, antibodies are produced against the envelope protein which elicits an immune response against virus (Urcuqui-Inchima et al., 2010; Lee et al., 2010). It is

responsible for the icosahedral assembly and integration with the host cell membrane for entry (Klein et al., 2013).

The prime factors responsible for the unchecked spread of DENV since decades include heterogeneity and virus breathing; former states the antigenic variability among the four DENV serotypes and latter refers to the changes in conformation of DENV at high temperature conditions (Rey et al., 2018; Fibriansah et al., 2013). The transition rates are 10 times the transversion rates in DENV (Nandy et al., 2016). Consequently the development of vaccine against Flaviviruses has become more demanding. So far, only preventive measures have been followed to restrain the spread of the virus (Collins et al., 2017). A successful vaccine entails both effective immunogen and an optimized protocol for its *in-vivo* delivery (Collins et al., 2017). The vaccine should be safe, potent, cost-effective, genetically stable, long lasting immunity, easy storage and transportation (Torresi et al., 2017; Ghosh et al., 2015). Though many experiments have been carried out for finding an apt vaccine candidate against DENV, no successful vaccine has been developed (Reginald et al., 2018). A live attenuated vaccine, Dengvaxia® (CYD-TDV) is the first licensed vaccine against DENV in 20 countries but it is not effective in children below the age of 9 years (Shim et al., 2019; Swaminathan et al., 2019). Other potent vaccine candidates are in various clinical trial stages (Ghosh et al., 2015).

In addition to the traditional method of vaccine development, a strategy of peptide based vaccine has been introduced for eliminating the constraints of traditional vaccines such as antibody dependent enhancement, chances of reversion, low yield and high cost. Immunoinformatics is a branch of bioinformatics, termed as immunomics which deals with study of various mechanisms of immune system (Tomar et al., 2010). The potential epitopes have been identified against H1N1 influenza virus and Ebola virus by using immunoinformatics (Jain et al., 2019; Lohia et al., 2015). In this study, immunoinformatics approach has been adopted for identification of potential peptides containing T cell epitopes of envelope protein of dengue serotype II which is most infectious among four serotypes and screening peptides for undesired responses such as allergenicity, toxicity and autoimmunity. Further, the identified peptides were analyzed *in silico* for their binding efficacy with different HLA alleles via molecular docking studies conducted with the help of Molecular virtual docker and CABS-dock tools. Population coverage analysis was carried out to assess the efficacy of these peptides in a global population while conservancy analysis of peptides in other members of

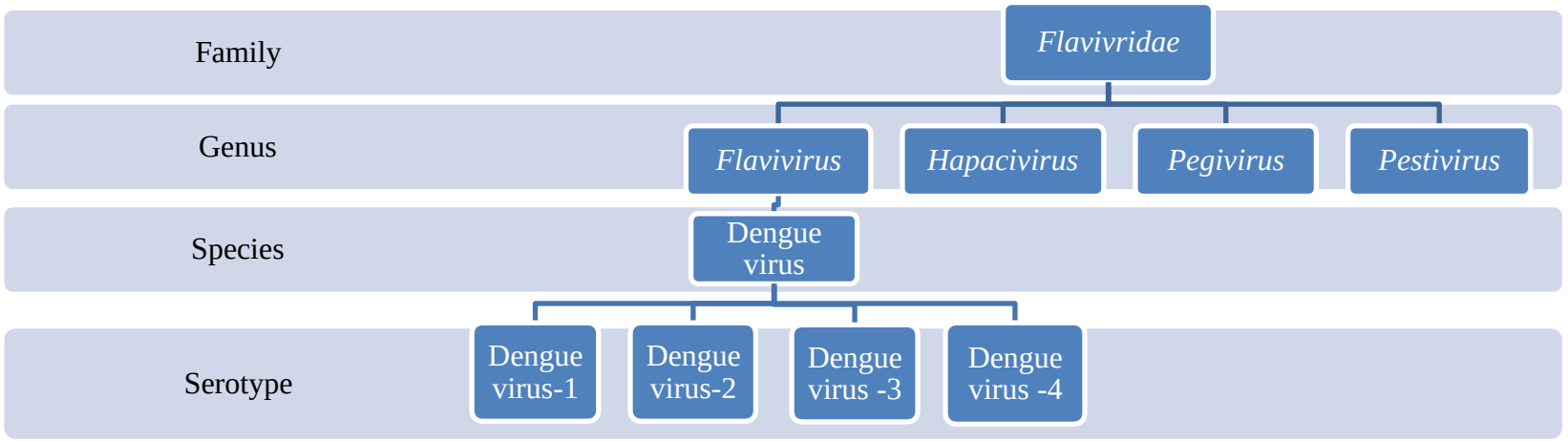
Flaviviruses was also carried out to find a vaccine candidate which may provide the cross protective immunity among the members of *Flaviviridae* family.

2. Review of Literature

2.1. Dengue virus: Taxonomy:

DENV is belongs to the *Flaviviridae* family and the genus *Flavivirus* (Table 1); (Salles et al., 2018; Gubler et al., 2006).

Table 1: Flowchart of DENV taxonomy (ICTV, 2019)



2.2 DENV Structure

DENV consists of an enveloped, positive sense, single stranded RNA genome with an icosahedral symmetry. It has an approximate length of 9.2-11 kb and is 50 nm in diameter (Laureti et al., 2018; Salles et al., 2018; El Sahili et al., 2017; Simmonds et al., 2017). Its genome consists of a single open reading frame (ORF) flanked by 5'and 3' untranslated region (UTR) having 100 and 350 to 1,200 nucleotides respectively (El Sahili et al., 2017; Simmonds et al., 2017; Junglen et al., 2016; Perera et al., 2008). The genome has 5' cap but lacks Poly A tail (Urcuqui-Inchima et al., 2010).

A large single polyprotein present in ORF gets cleaved by co- and posttranslational modification and thereby releases three structural and seven non structural viral proteins (Figure 1); (Junglen et al., 2016).

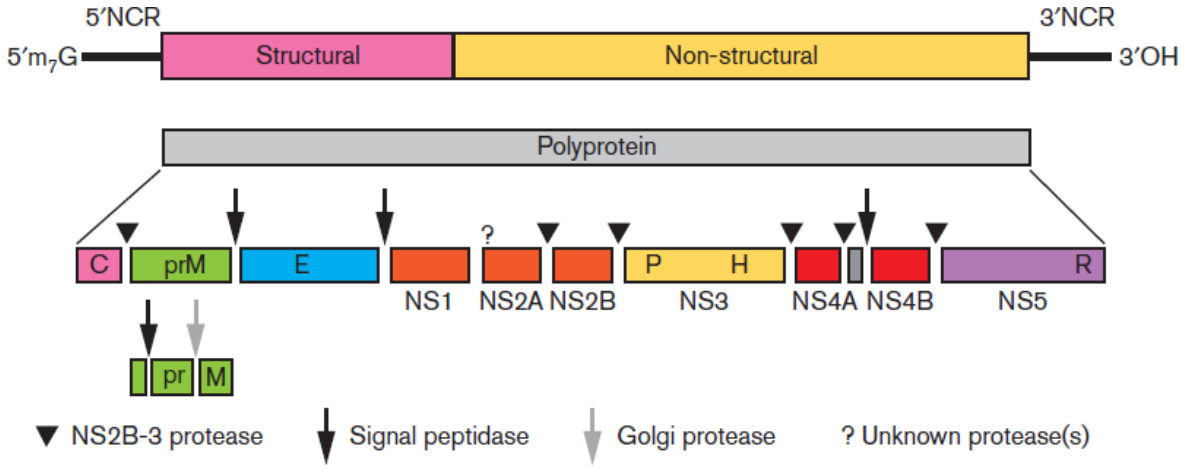


Figure 1: Genome organization and polyprotein processing of members of the genus *Flavivirus* (Simmonds et al., 2017).

The structural proteins constitute the virus particle while the non structural proteins participate in virus replication. The capsid protein (C), participates in genome packaging and formation of nucleocapsid core. The membrane protein (prM or M), acts as a shield for folding and assemblage of the E protein. The envelope protein (E) has receptor binding site and integration peptide (Mukhopadhyay et al., 2005). A summary of the various DENV proteins and their functions has been presented in table 2.

Table 2: Functions of various proteins of Dengue virus

Protein	Functions
Capsid (C)	Assembly of virus
Envelope (E)	Icosahedral assembly and host invasion
Membrane (prM/M)	Folding and secretion of E protein
	Virus entry
NS1	Virion replication and immune response activation
NS2A	Virus assembly and inhibition of host immune response
NS2B	Serves as a cofactor for NS3
NS3	Viral assembly
NS4A	Serves as a cofactor for NS3
NS4B	Viral replication and anti-host response
NS5	Dengue fever (DF) pathogenesis

2.3 Structural proteins

DENV consists of three structural proteins namely core or capsid (C), envelope (E) and membrane (prM/M) proteins. They form the components of the virion and participate in packaging, export and subsequent viral entry (Urcuqui-Inchima et al., 2010; Perera et al., 2008; Pang et al, 2001).

2.3.1 Capsid protein (C)

Capsid protein is dimeric, basic with 100 amino acids and 11KDa size (Cruz-Oliveira et al., 2015). It is found in cytoplasm and nucleus and has nuclear localization signal. C proteins bind with intracellular membranes via conserved hydrophobic domains and their N and C termini have charged residues (Nemésio et al., 2011). The hydrophobic sequence of C – terminus acts as a signal peptide for translocation to endoplasmic reticulum (ER). They form nucleocapsid with viral RNA, which activates its dimerization required for virus assembly (Nemésio et al., 2011; Urcuqui-Inchima et al., 2010). C proteins bind with DAXX protein and human Sec 3 (a

transcription and translation repressor), which thereby triggers apoptosis and retards infection respectively (Urcuqui-Inchima et al., 2010; Bhuvanakantham et al., 2010).

2.3.2 Membrane protein (prM)

Membrane glycoprotein has 167 amino acids and 21 kDa size (Cruz-Oliveira et al., 2015). It is cleaved into pr and M with 13 kDa and 8 kDa size respectively by an endo-protease, furin which is present in Trans Golgi network. The cleavage is essential for causing infection. It has His 39 which is necessary for the virus morphology, release and invasion into the cell. E protein and prM proteins together form spikes on the virus surface and facilitates folding and release of envelope (E) protein. Membrane protein binds with vacuolar ATPase (V-ATPase) for release of virus. After exposure to neutral pH and secretion of particles, pr leaves virions. It interacts with E protein and prevents fusion of E protein with the cell membrane at acidic pH. For virus entry, it binds with the claudin-1 tight junction membrane protein (Urcuqui-Inchima et al., 2010).

2.3.3 Envelope protein (E)

Envelope glycoprotein has 495 amino acids and 50 kDa size. It is responsible for response and virulence and is the primary target of neutralizing and amplifying antibodies (Hsieh et al., 2010; Mukhopadhyay et al., 2005; Urcuqui-Inchima et al., 2010). It forms an outer shell of virus and has two transmembrane helices, a receptor binding site, a fusion peptide. Upon virus invasion, antibodies are produced against the envelope protein which elicits an immune response against virus (Urcuqui-Inchima et al., 2010; Lee et al., 2010). It is responsible for the icosahedral assembly and integration with the host cell membrane for entry (Klein et al., 2013). It has three domain structures: central domain (D1) which is flanked by domain II (Elongated dimerization domain) and domain III (Immunoglobulin- like domain) (Figure 2 and Table 4). DIII is a continuous peptide, which forms immunoglobulin like fold of seven anti- parallel β sheets, whereas DI and DII are discontinuous peptides, joined together by EDI/EDII hinge, which contain epitopes determinants of serotype specificity and facilitates antibody dependent enhancement in serotype specific manner (Messer et al., 2014; Waterson et al., 2012). The hinge region plays pivotal role in conformational changes in the E protein, thereon fusion with endosome and then virus entry into the host (Messer et al., 2014). In the absence of Asn-153 glycosylation site in the E protein DENV infection can be decreased. The two chaperons (Calnexin and Calreticulin) are vital for folding, assembly and production of virion particles (Limjindaporn et al., 2009). The mature and immature virion exists in dimeric and trimeric conformation respectively. The E protein is arranged in the anti-parallel orientation whereas E

protein is oriented away from the virus surface and thereby forms spikes (Figure 4) (De la Guardia et al., 2014).

Table 3: Location of different regions in the Dengue envelope protein (Nandy et al., 2016).

Regions	Location
Domain I	1-52, 132-191 and 278-294
Domain II	53-131 and 192-277
Domain III	295-392
Stem segment	394-449
Transmembrane anchor	449-495

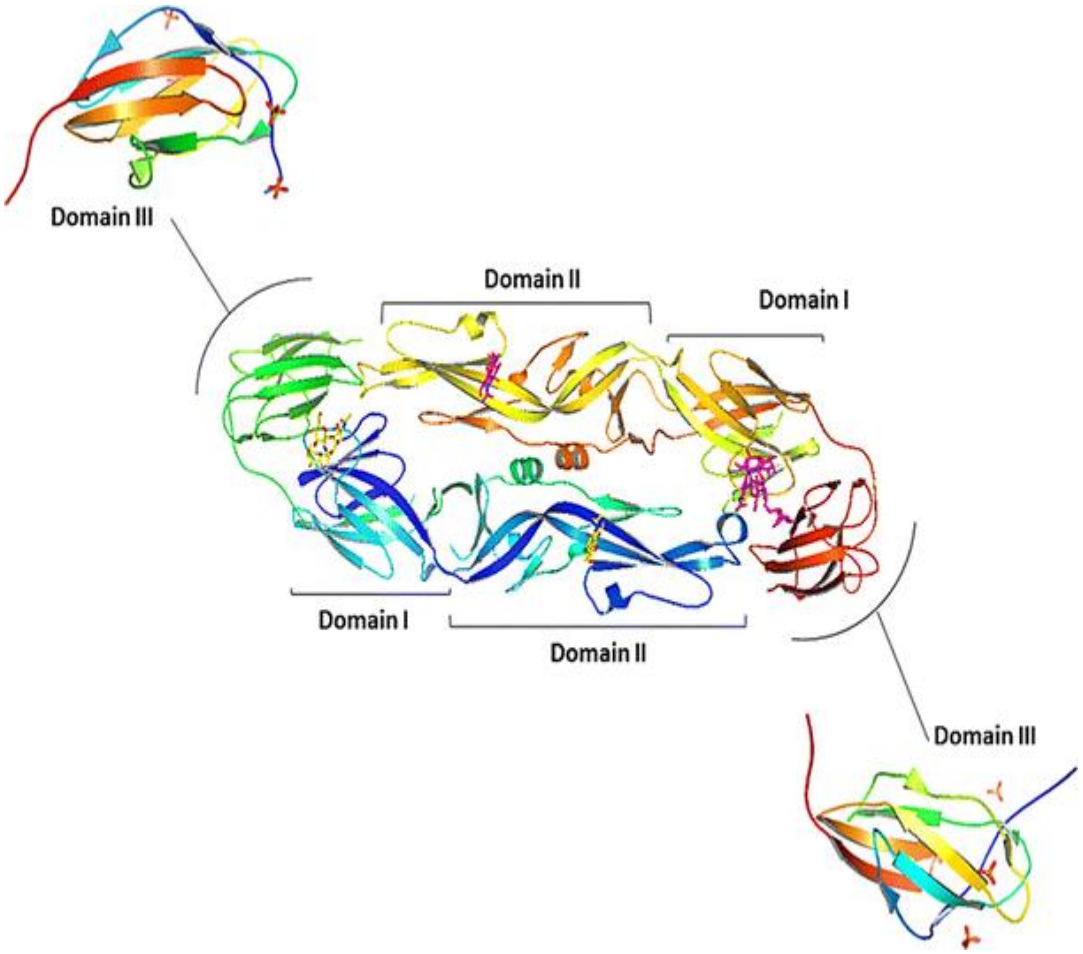


Figure 2: Crystal structure of dimeric Dengue envelope protein (Fahimi et al., 2018)

DENV E protein binds with 467-530 amino acid residues of binding protein (BiP) (Figure 3). Its transmembrane domain facilitates retention and assembly of E protein on ER (Hsieh et al. (2010). E protein is intracellularly bound to prM protein, which protects E protein from premature acidification and E protein undergo conformational changes upon cleavage of prM and in low pH of Golgi apparatus. The low pH of E protein induces formation fusion complexes (Figure 4). It interacts with V-ATPase and natural killer cell activating receptor (NKp44), former leads to entry and egress of virus and latter causes NK cells activation and destruction of virally infected cells (Lee et al., 2010; Limjindaporn et al., 2009; Duan et al., 2008).

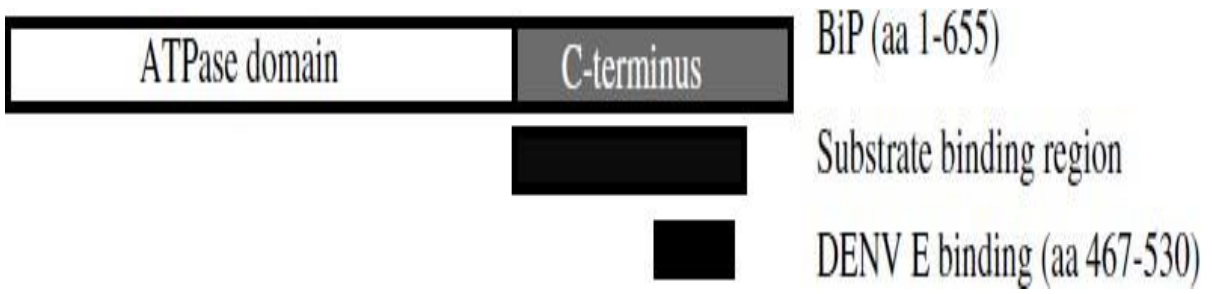


Figure 3: Binding of Dengue envelope protein with Binding protein (Limjindaporn et al., 2009)

DII of DENV E protein has two N-linked glycans: Asn-153 and Asn-67 and latter is responsible for binding of virus with C-type lectin and DC-SIGN expressed on dendritic cells and macrophages; thereby promotes virus pathogenesis (Lee et al., 2010). Glycosaminoglycans, is a ubiquitous protein, which plays pivotal role in, motility, adhesion proliferation and integration with host cell surface to invade the host cell (Waterman et al.,2012).

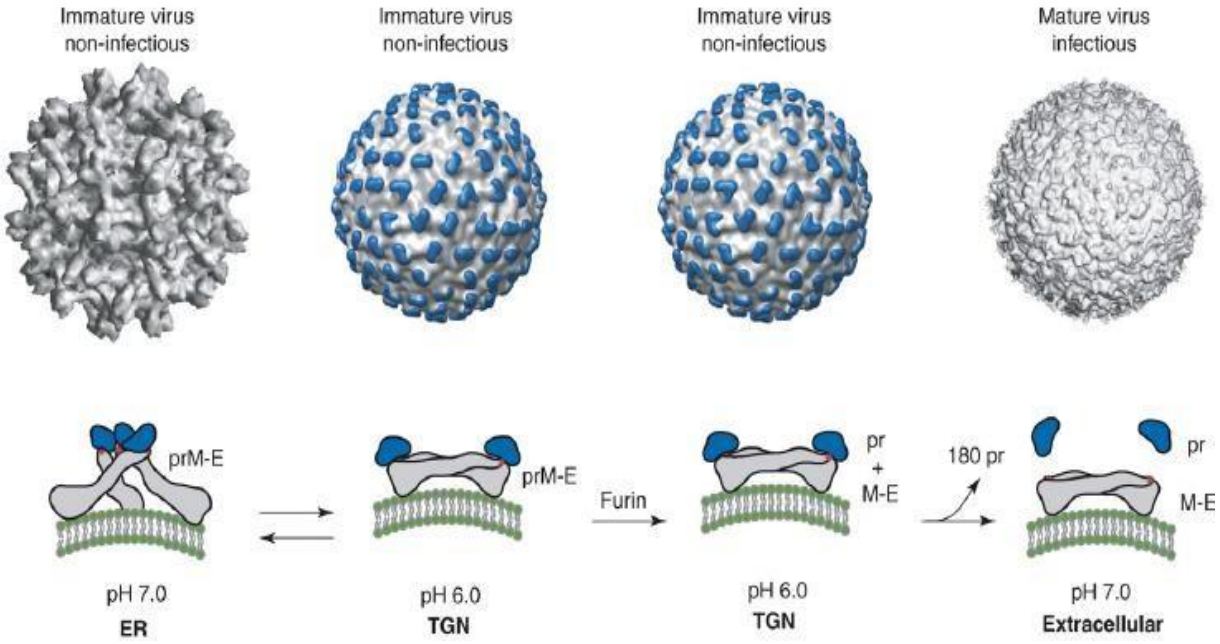


Figure 4: Conformational changes in dengue envelope protein (Perera et al., 2008)

2.4 Non-structural (NS) proteins

There are seven non-structural proteins of Flaviviruses: NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. Main roles of NS proteins are organization of cell, maturation of polyprotein and replication of viral particle (El Sahili et al., 2017).

2.4.1 NS1

NS1 is pervasive, conserved glycoprotein with 352 amino acids and 46 and 54 kDa molecular weight on the basis of N- glycosylation sites at positions 130 and 207 respectively (Alcala et al., 2018; Rastogi et al., 2016; Edeling et al., 2014). Its 225-245 amino acids are highly conserved in all four serotypes of DENV (Omokoko et al., 2014). It exists in three forms: monomer, dimer (membrane bound) and hexamer (secreted protein); the intracellular NS1 is responsible for viral replication, although membrane and secreted NS1 activates the immune response (Rastogi et al., 2016; Edeling et al., 2014). NS1 attached to the lumen of ER membrane is responsible for anchoring viral replication complex (El Sahili et al., 2017). Infected *Aedes aegypti* mosquito secretes high level of NS1 protein as it is present in the saliva of mosquito (Alcala et al., 2018). Its glycosylation is necessary for effective secretion and viral replication and also facilitates activation and inhibition of toll-like receptors (TLRs) and complement system respectively in DENV infection (Rastogi et al., 2016). Antibodies generated against NS1 protein cross react with the antigens on the surface of endothelial cells, thereby causing plasma leakage, dehydration and hypovolemic shock (Zeidler et al., 2018).

2.4.2 NS2

NS2 is a transmembrane protein with 217 amino acids and cleaves into mature proteins, NS2A and NS2B (Chen et al., 2017; Perera et al., 2008). NS2A has 22 kDa molecular weight (Xie et al., 2013). It engages in virus assembly and inhibition of host immune response, apoptosis, cell growth and cell cycle regulation (Chen et al., 2017; Xie et al., 2013). Its N-terminus 68 amino acids reside in ER, 69-209 amino acids form five transmembrane segments and C-terminal 210-218 amino acids reside in the cytosol. Arg-84 residue is essential for viral RNA synthesis, virion assembly and maturation (Xie et al., 2013).

2.4.3 NS3

NS3 is a multifunctional protein with 618 amino acids (El Sahili et al., 2017). It has serine protease (Chymotrypsin) at C –terminus and RNA helicase and RNA triphosphatase at N-terminus (Chen et al., 2017; Xie et al., 2013; Bollati et al., 2010; Perera et al., 2008). During mRNA capping, 5’ triphosphate mRNA is converted into diphosphate by RNA triphosphatase. RNA helicase facilitates virion assembly (Bollati et al., 2010). NS3 requires NS2B as cofactor for its activity, NS2B3 as a complex inhibits innate immunity and ruptures the C protein (Chen et al., 2017). During infection, NS3 binds with fatty acids synthase, thereby increases

synthesis of fatty acids as the energy requirement increases during infection (Zeidler et al., 2017). It causes apoptosis of infected cells (Perera et al., 2008).

2.4.4 NS4

It is a highly hydrophobic protein, which exists in two forms, NS4A and NS4B joined by a transmembrane protein, 2K which is separated upon NS4A maturation. NS4A is 16KDa protein which functions as a cofactor of NS3 protein and elicits autophagy (Zeidler et al., 2017; Nemésio et al., 2012; Perera et al., 2008).NS4B is a 27KDa protein, which forms convoluted membrane in which viral protein gets processed and also inhibits mitochondrial fission process, thereby enhancing the oxidative metabolism (Zeidler et al., 2017; Perera et al., 2008). Both of them engages in the replication of virus and evades host immune response (Nemésio et al., 2012).

2.4.5 NS5

NS5 is the largest, highly conserved protein with 102 kDa molecular weight (El Sahili et al., 2017). NS5 N-terminus and C-terminus has methyltransferase (MTase) domain and RNA dependent RNA polymerase (RdRp) respectively (El Sahili et al., 2017; Bollati el al., 2010; Perera et al., 2008). It exhibits 70% conservancy among all the four serotypes of DENV (El Sahili et al., 2017). It participates in DF pathogenesis by enhancing the expression of cytokines (Zeidler et al., 2017). MTase domain is responsible for 5’ capping of viral RNA and RdRp for viral replication. RdRp can be used as a drug target, since host lacks RdRp activity (El Sahili et al., 2017). NS5 impedes the interferon signaling process (Chen et al., 2017).

2.5. DENV: Life cycle

Flavivirus follow type II fusion process for the viral host membrane fusion.

Virus entry into the host cell is facilitated by receptor -mediated endocytosis and antibody-dependent enhancement (ADE) into monocytes, macrophages and dendritic cells (Figure 5 & 6) (Laureti et al., 2018; Tuiskunen Bäck et al., 2013; Nemésio et al., 2011). Multiple receptors (such as DC-SIGN, mannose receptors, CD14, heparin sulfate and C-type lectin receptors) promote internalization and attachment of the virus into an endosomal compartment, where acidic pH promotes integration of viral envelope in the endosome by rearrangement of capsid proteins, which facilitates invasion of virus into the host cell.

During viral entry, the virus attaches to glycosaminoglycans (heparin sulfate), thereafter binding with receptor commences the development of clathrin-coated particles throughout the virus-receptor complex. Pre-existing clathrin- coated pits engulfs the newly formed particles. The clathrin coated pits fuse with the plasma membrane and move out of the plasma membrane into the cytosol by dynamin GTPase (Medhigeshi et al., 2011).

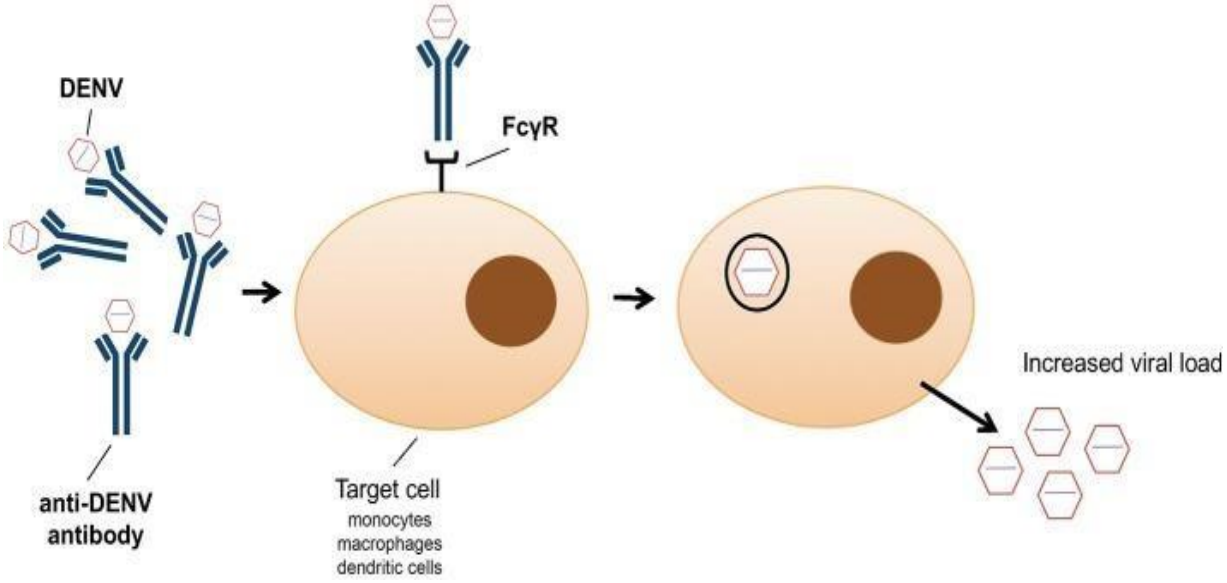


Figure 5: Dengue entry by antibody dependent enhancement mechanism (Laureti et al., 2018).

Virus replication first occurs in Langerhans cells, consequently in lymph nodes and other tissues, which eventually leads to proliferation of infection in target sites (Medhigeshi et al., 2011).

Released viral RNA into the host cytoplasm enters the endoplasmic reticulum (ER), where it undergoes two different processes: First, translation of positive sense RNA produces a polyprotein, which cleaves into three structural (Membrane, prM/M, Envelope, E and Capsid, C) and seven non- structural (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) proteins by post -translation modification (Figure 2); (Junglen et al., 2016; Aruna et al., 2014) and second, viral NS5 RNA dependent RNA polymerases converts genetic material of virus into a negative sense RNA to form positive- stranded RNA copies, followed by the formation of nucleoprotein by packaging of viral genome in the cytoplasm by protein C while the prM and E proteins get heterodimerized in ER and commence viral proliferation (Rey et al., 2018; Daep et al., 2014; Fernandez-Garcia et al., 2009). The nascent virion particles travel into the Golgi apparatus from the ER lumen via a secretory pathway. Separation of prM/E heterodimer is activated by

changes in pH in Trans-Golgi network by activation of cellular furin (Endo-protease). Activated protease acts at an Arg-X-(Lys/Arg)-Arg recognition sequence and cleaves prM protein to produce protein M and the peptide pr, which leads to the formation of fully mature infectious virion (Aruna et al., 2014; Snyder et al., 2014; Daep et al., 2014).

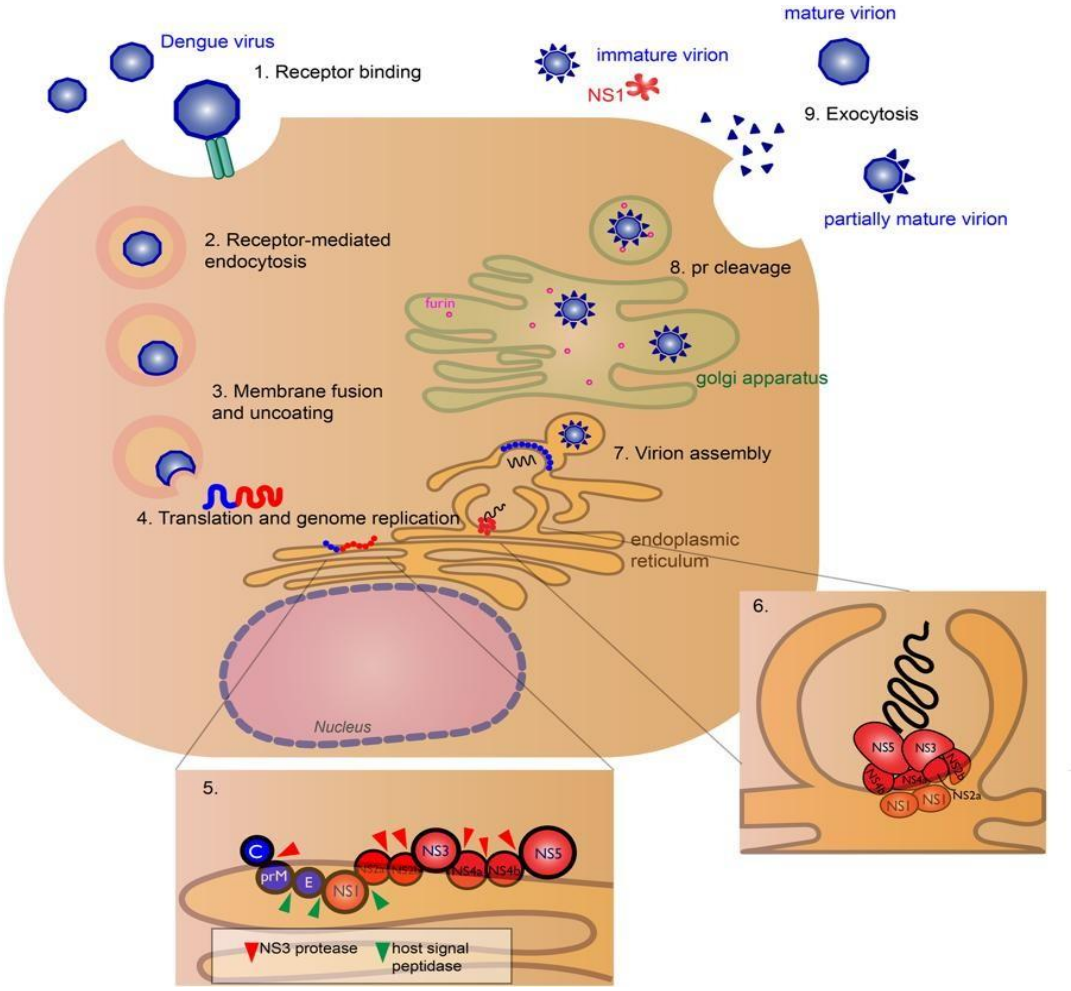


Figure 6: Life cycle of flavivirus infection (Uno et al., 2018)

2.6 History

Dengue fever (DF) was known as “water poison”, related with flying insect from the Jin Dynasty. Benjamin Rush gave the term break bone fever.

2.6.1 First evidences of human infection:

In 1789, first dengue epidemic was reported in Asia, Africa and North America simultaneously (Gupta et al., 2012; Gubler et al., 2006). In India, the first dengue epidemic was reported in

Chennai in 1780 (Gupta et al., 2012). In 1780, the first clinically proven epidemic caused by dengue virus was reported in the United States by Benjamin Rush (Dick et al., 2012).

2.6.2 First isolation:

In 1943, the Dengue virus was isolated from mice inoculated by serum of infected patient in Japan and in 1944 from serum samples of US soldiers in Chennai (Gupta et al., 2012).

2.6.3 Spread of dengue (1789 to till date):

DENV infection is a global health concern predominantly in tropical and sub-tropical regions of the world with 390 million infections per year (Diamond et al., 2015; Christofferson et al, 2015). Before 1970, the dengue fever epidemic was faced by only nine countries. DF is endemic and occurs in more than 100 countries every year (Diamond et al., 2015). Key factors responsible for the spread of dengue fever are urbanization and international travel (Messina et al., 2014). In 2017, 53% reduction in dengue cases was observed (Table 4).

Table 4: Dengue outbreak reports

Year	Area	Cases	Reference
1780	United States	No report	WHO, 2019
1789	Asia, Africa and North America	No report	
1963-1964	Calcutta and Eastern coast of India	No report	
Before 1970	Only one nine countries	No report	
2012	Portugal	~ 2000 cases	
2015	America	2.35 million cases	
	Delhi	15,000 cases	
	Island of Hawaii	181 cases	
2016 (worldwide)	America	<1.5 million cases	
	Brazil	1,032 deaths	
	Philippines	1,76,411cases	
	Malaysia	1,00,028 cases	
	Solomon Islands	>7000 cases	
	Africa	1061 cases	
2018	America	4,46,000	CDTR,2018
	Brazil	2,18,000	
	Mexico	62,000	
	Nicaragua	46,000	
	Colombia	32,000	
	Paraguay	30,000	
2019	Cambodia	693cases	WHO,2019
	China	52 cases	
	Loa PDR	368 cases	
	Malaysia	1,622 cases	
	Philippines	786 cases	
	Singapore	161 cases	
	Viet Nam	1,513 cases	
	Australia	429 cases	
	French Polynesia	42 cases	
	New Caledonia	3,081 cases	

2.7 Homology among four serotypes of DENV

Out of 495 amino acids of DENV E protein, 156 amino acids are conserved among the four serotypes (Nandy et al., 2016). His-282 and His-317 residues in E protein are highly conserved among all the Flaviviruses. All the four distinct serotypes of DENV exhibit 69-78% amino acids and 67-73% nucleotide identity, specifically from residue 320-368 and differ by 25-40% at amino acid level and elicits cross reactive antibodies (Diamond et al., 2015; Liu et al, 2014; Nemésio et al., 2011; Perera et al., 2008; Pang et al., 2001). Among the four serotypes of DENV, serotype II is most infectious, severe and occurs frequently (Rico-Hesse et al., 2003). In all DENV serotypes rate of T-C transition is greater than A-G transitions and purine to pyrimidine transversion speed is less than pyrimidine to purine transversion. In the comparative study of codon usage by the E protein of DENV four serotypes, it has been reported that usage of arginine is greater among the four serotypes, use of AGA codon by DENV-3 and DENV-4 is greater and CGA codon by DENV-1. The utility of threonine and leucine is higher in DENV-1 and the utility of valine and glycine is immense in DENV-4. The percentage of hydrophobic residues is higher in DENV-3 whereas the percentage of hydrophilic residues is higher in DENV-1 and DENV-4 (Nandy et al., 2016).

2.8 DENV host immune evasion

The mRNA of DENV resembles the host mRNA at 5' cap and 2'O methylation. These post-transcriptional modifications executed by non structural, NS5 protein of DENV, thereby evades the host immune response. Upon DENV infection, innate immunity serves as first line of defense and activates pathogen recognition receptors (PRRs): Melanoma differentiation associated protein-5 (MDA-5), which perceives long dsRNA and retinoic acid inducible gene – I (RIG-I), which perceives short dsRNA. Both MDA-5 and RIG-1 shares same adaptor , mitochondrial antiviral signaling (MAVS) and elicits IFN production by inducing tank binding kinase-1 (TBK-1) and IκB kinase ϵ (IKKϵ) (Uno et al., 2018). DENV NS2 protein is a DNA protease and targets stimulator of interferon (IFN) genes (STINGS), which binds with MVAS and RIG-I, thereby escalates antiviral response. During DENV infection a cytosolic DNA sensor, cyclic GMP-AMP synthase (cGAS), gets activated upon delivery of mtDNA in the cytosol, which consequently leads to NS2B mediated degradation of cGAS via autophagy lysosome pathway and prevents IFN signaling. DENV NS2A and NS4A proteins inhibit IFN-3 phosphorylation and DENV-1 inhibits TBK-1 for evading host immune response (Kao et al., 2018). DENV evades RNA silencing pathway by transforming elements of RNA interference (RNAi), thereby facilitates viral RNA replication. DENV evades autophagy by inhibiting

binding of autophagosome with lysosome and eliciting phosphotidylinositol 3 kinase 3/Akt pathway which consequently promotes viral replication and maturation.

In the absence of target cells, deactivation of flavivruses occurs at pH 6.0-6.5 as they are sensitive to slightly acidic pH but the mature virus exhibits virus breathing property, which prompts host invasion in the absence of receptor by interaction of exposed antibodies. The prime factors responsible for the antigenicity of flaviviruses are heterogeneity and virus breathing (Rey et al., 2018). The binding affinity of DENV E protein with DC-SIGN is enhanced by glycosylation of Asn167, which consequently enhances dendritic cell infection. The interaction occurs via formation of hydrogen bond between mannose bound to Asn67 and Asn272 of DC-SIGN (De La Guardia et al., 2014). The viral replication can reduce 100 times, if the valine-251 residue in E protein is mutated by alanine. Asp-215, Pro-217 and His-244 residues of DENV-2 E protein are pivotal for viral replication as their replacement with alanine residue reduces the viral replication and Asp-215 and His-244 are conserved among flaviviruses (Kroschewski et al., 2009). The transition rates are 10 times the transversion rates in DENV (Nandy et al., 2016).

Cross reactivity is the prime cause of mutation in DENV, consequently over time the available anti viral therapy will not be able to produce immune response against DENV. It has been reported that after DENV epidemic in 1981 and 1997, secondary infection has become more severe (Rodriguez-Roche et al., 2016). Conserved regions of virus genome or proteome are mainly considered during drug development, even though viruses develop resistance against drugs. RNA-dependent RNA polymerase and ability of RNA viruses to adapt quickly are mainly responsible for development of resistance against anti-viral therapies over time (Sessions et al., 2015). By preventing the interaction between DENV and dendritic cells in the beginning harmful immune response can be curbed (De La Guardia et al., 2014).

2.9 Dengue fever and symptoms

Flavivirus infection causes mild to severe illness which includes fever, headache, body aches, vomiting, diarrhea , joint pains, enhancement endothelial permeability, plasma leakage, spread of intravascular coagulation hemorrhage and hypovolemic shock (Zeidler et al., 2017; Rocha et al., 2014; Urcuqui-Inchima et al., 2010). Among Flaviviruses, in severe cases DENV leads to hemorrhagic fever (Perera-Lecoin et al., 2014).

Generally, symptoms such as high fever, chills, headache, nausea and vomiting appear after 3-6 days of incubation. Clinical investigation displays increased levels of liver enzymes, leucopenia, thrombocytopenia and clotting abnormalities (Gould et al., 2008).

2.10 Diagnosis of DENV

Principle markers for diagnosis of DENV infection are virus isolation, detection of NS1 protein and measurement of immunoglobulin (Ig) M and IgG levels (Muller et al, 2017; Tuiskunen Bäck et al., 2013).

2.10.1 Virus isolation

It is traditional method of diagnosis. It involves culturing of patients sample in cell lines of either mosquito such as C6/36, AP-61 or mammals such as BHK-21 cells. It gives good results after five days of appearance of primary infection, whereas isolation process becomes cumbersome in case of secondary infection because of the expeditious generation of cross-reactive antibodies. It is a time consuming method and gives report after few days or weeks (Muller et al., 2017).

2.10.2 RT-PCR (Molecular method)

It is highly sensitive, fast, specific and less cumbersome than virus isolation method. It gives good results after the appearance of infection and patients gets report either on same or next day of diagnosis (Muller et al., 2017).

2.10.3 NS1 capture

NS1 can be detected in the blood after 9 days of primary infection by using antigen capture enzyme linked immune sorbet assay (ELISA) method (Figure 8 and Table 2). It is an epitome diagnostic method, since it is released by infected cells only. More than 600 ng/mL level of NS1 manifests succession to DSS/DHF. It permits early diagnosis of disease progression (Muller et al., 2017).

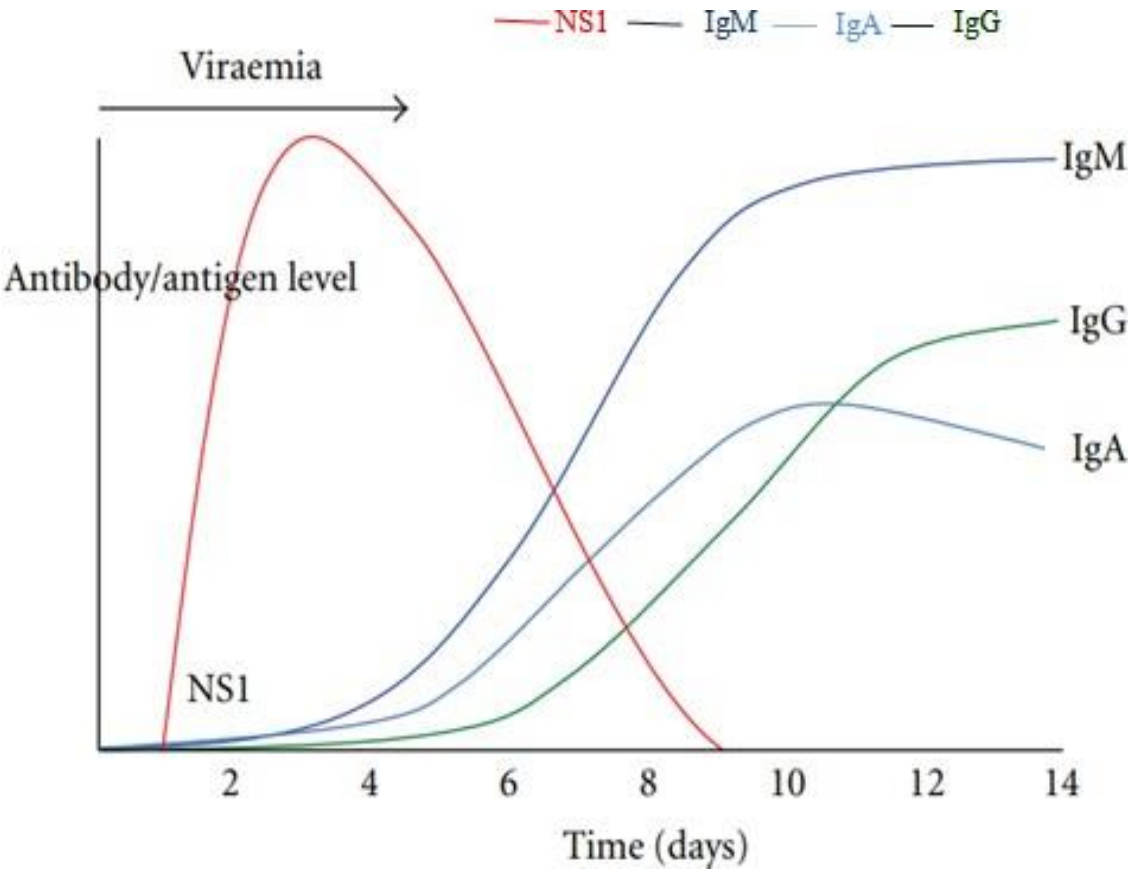


Figure 7: Peak levels of biomarkers during primary dengue infection (Blacksell et al., 2012)

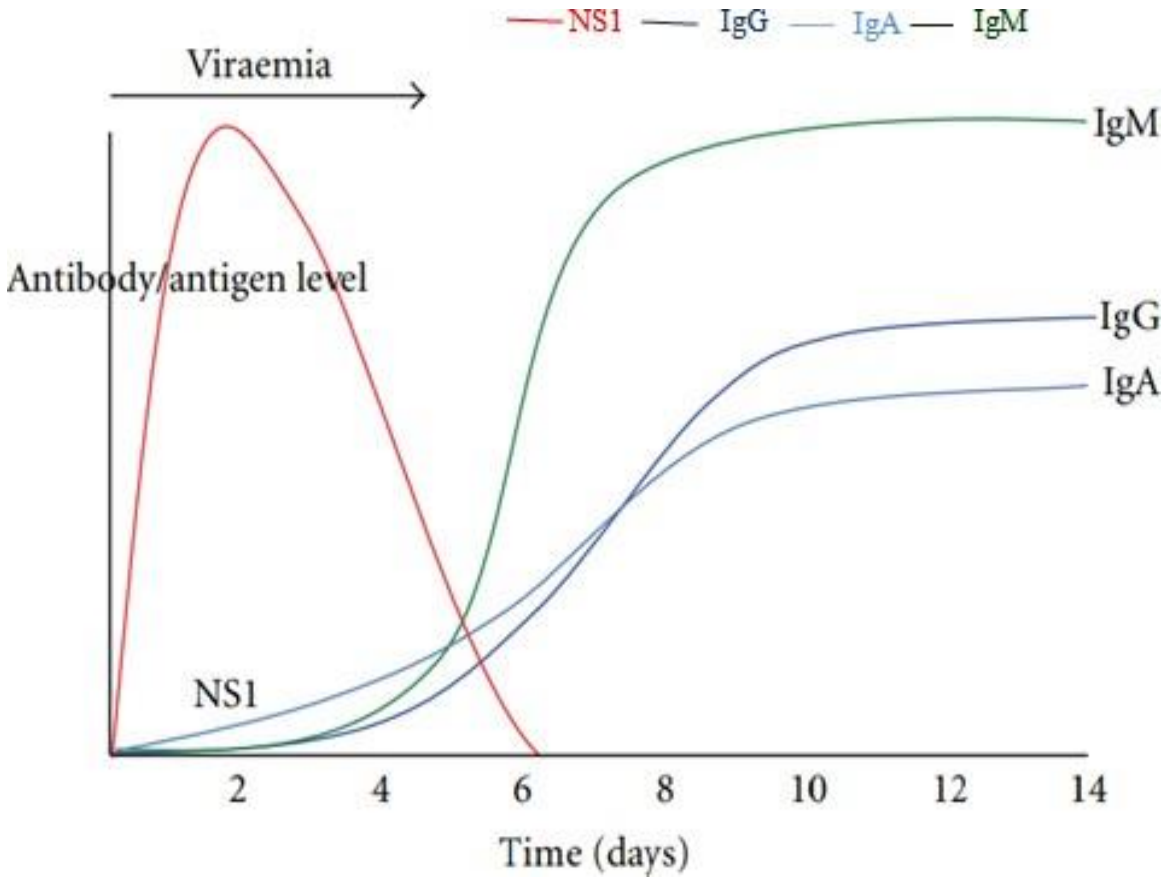


Figure 8: Peak levels of biomarkers during secondary dengue infection (Blacksell et al., 2012).

2.10.4 Serological diagnosis

Hemagglutination inhibition (HI) assay and IgM and IgG antibody capture ELISA assays are mainly employed for serological diagnosis. IgM can be detected in patient after 3-5 days of primary infection and remains for several months, whereas IgG can be detected within 3 days in case of secondary infection (Figure 8 and Table 2). Due to the formation of cross-reactive antibodies with other Flaviviruses, it might give inaccurate results, but this problem can be solved by performing IgM-IgG antibody capture ELISA along with NS1 capture method. In laboratories, usually IgM-IgG antibody capture ELISA assay and immune-chromatographic strip tests are employed, since they remain unaffected by the formation of cross reactive antibodies (Muller et al., 2017; Tuiskunen Bäck et al., 2013).

Table 5: Diagnostic approaches employed for DENV infection in laboratories (Muller et al., 2017)

	Diagnostic approach	Methodology	Time
Virus detection	Virus isolation	Mosquito cell culture	1 weak or more
	Nucleic acid detection	RT-PCR, real time RT-PCR	1-2days
	Antigen detection	NS1 Ag rapid test	Few minutes
		NS1 Ag capture ELISA	1 day
Serological response	IgG or IgM	ELISA	1-2 days
	IgM detection	MAC*-ELISA	1-2 days
		IgM rapid tests	Few minutes
	IgG detection	IgG ELISA	1-2 days
		IgG rapid tests	Few minutes

MAC*: IgM antibody capture

Some diagnostic methods for DENV infection are under development such as micro or paper fluidics, *in vivo* micro-patches, isothermal PCR, electrochemical and pizochemical detection method. Currently, IgM and IgG antibody capture along with NS1 detection method is most efficient method and exhibits 63% to 94% sensitivity on basis of DENV serotype and whether the infection is primary or secondary (Muller et al., 2017; Tuiskunen Bäck et al., 2013). The primary and secondary infection can be classified based upon IgM/IgG ratio, ratio more than 1: 78 manifests primary infection whereas lower than 1:78 manifests secondary infection (Dutra et al., 2009).

2.11 Treatment and Prevention

No propitious treatment has been yet developed for curing infection caused by Flaviviruses, only supporting treatment is available. The major challenge in the development of vaccine is chances of antibody dependent enhancement (ADE), since long term protection can be generated against a serotype but if the person gets infected with the other serotype, it may lead to severe illness by drastically increasing the rate of replication (Shim et al., 2019; Govindrajan et al., 2015). By the development of vaccine, which can elicit an immune response against majority of the flaviviruses, can prevent the risk of ADE. Prime targets for the vaccine development are inhibition of virus entry, replication, adhesion, ADE. (WHO, 2019).Preventive measures must be taken by people for halting flavivirus infection. Some of the preventive measures are halting the breeding of mosquito by not allowing stagnant water around your surroundings, use mosquito spray, repellents, use mosquito net, keep water containers and tanks tight. Patient should take rest and drink a plenty of water. Do not take aspirin or ibuprofen, as it can enhance possibility of bleeding (CDC, 2019). Spread of infection can be restrained by vaccinating individuals travelling in and out of endemic areas.

2.12 Vaccines against DENV

Vectors are the genetically engineered bacteria, viruses or plasmid, which lacks the ability to replicate itself inside the host. They can facilitate targeted delivery of desired antigens and can carry multiple genes or proteins of interest for eliciting specific immune response against more than one pathogen (Liu et al., 2010). Therefore for development of an effective and safe vaccine, vectors can be employed for targeted delivery and limiting the chances of cross reactivity among the DENV serotypes or other flaviviruses.

A major issue of ADE can be evaded by development of a vaccine with multiple strains of DENV to induce high affinity, neutralizing antibodies against all serotypes of DENV (Pang et al., 2001). An immune response is developed for about 18 months against a similar strain after first exposure and cross protection against different strain for 2-9 months, but in severe cases it may cause DFF/DSS (Inchima wt al., 2010). In addition to ADE, other challenges in the development of vaccine are absence of animal models for disease and suitable makers responsible for immunity (Ghosh et al., 2015). A live attenuated vaccine, Dengvaxia® (CYD-TDV) is the first licensed vaccine against DENV in 20 countries, (Torresi et al., 2016; Collins et al., 2017; Shim et al., 2019). CYD-TDV stands for chimeric yellow fever virus-DENV tetravalent dengue vaccine, which constitutes the prM and E sequence of all four DENV serotypes (Collins et al., 2017; Guy et al., 2016). Vaccines in clinical trial are mentioned in Table 6.

Table 6: Licensed and clinical trial vaccines against Dengue (Gromowski et al., 2018; Diaz et al., 2018)

Type of vaccine	Developer/ Manufacture	Phase of clinical trial
Tetravalent vaccine (Live attenuated vaccine)	Walter Reed Army Institute of Research (WRAIR), Glaxo Smith Kline (GSK)Biologicals	II
Tetravalent vaccine	National Institutes of Health (NIH) and St. Louis University Health Science Centre, Sanofi Pasteur	Licensed (Collins et al., 2017)
Tetravalent Vaccine	Inviragen Inc.	II
Tetravalent vaccine	National Institute of Allergy and Infectious Disease (NIAID), National Institutes of Health (NIH), Butantan Institute	I
Tetravalent vaccine (purified Inactivated vaccine)	Walter Reed Army Institute of Research (WRAIR), GlaxoSmithKline Biologicals, Oswaldo Cruz Foundation	I/II
Monovalent vaccine	Merck and Co.	I
Monovalent vaccine	Naval Medical Research Centre, Walter Reed Army Institute of Research (WRAIR)	I
Tetravalent vaccine	Panacea Biotech Ltd	I/II
Tetravalent vaccine	Butantan	III

2.13 Peptide based vaccines

Prime requirement for producing immunization against a disease is generation of immune response, in which T cells and B cells plays an important role. T cells are of two types, T helper cell, T_H and cytotoxic T Cell, T_C, which contains $CD4^+$ and $CD8^+$ markers on the surface respectively (Marintcheva et al., 2018). $CD4^+$ and $CD8^+$ present peptide to class I and class II major histocompatibility complexes (MHCs) respectively. T cell memory is mainly responsible for eliminating and destroying foreign pathogens upon exposure after vaccination (Haung et al., 2017).

Traditional vaccines including live attenuated vaccine, inactivated whole vaccine, recombinant subunit vaccine and DNA vaccine have been developed for treating various diseases, but exhibit some drawbacks such as ADE, chances of reversion, low yield and high cost. These drawbacks can be overcome by the development of peptide based vaccine (Figure 10).

Activation of B cells, cytotoxic T cells, interferon γ, and interleukins is elicited by T_H cells, thereby T_H or $CD4^+$ T cell epitopes are vital component of peptide based vaccine (Skwarczynski et al., 2016).

Peptide based vaccines are usually safe and cost effective. Peptide based vaccines are soluble in water, can be freeze dried for storage. A peptide vaccine may contain more than one epitope, can serve as vaccines against more than one disease (Skwarczynski et al., 2016). Even though no peptide based vaccine has been licensed as yet, many potent candidates are in different stages of clinical trials (Skwarczynski et al., 2016). Peptide vaccines for human immunodeficiency virus (HIV), hepatitis C virus (HCV), malaria, foot and mouth disease, swine fever, influenza, anthrax and human papilloma virus (HPV) are under clinical trials (Figure 9); (Li et al., 2014). Among flavivirus, peptide based vaccine for ZIKV is under clinical trials (Wilder-Smith et al., 2018).

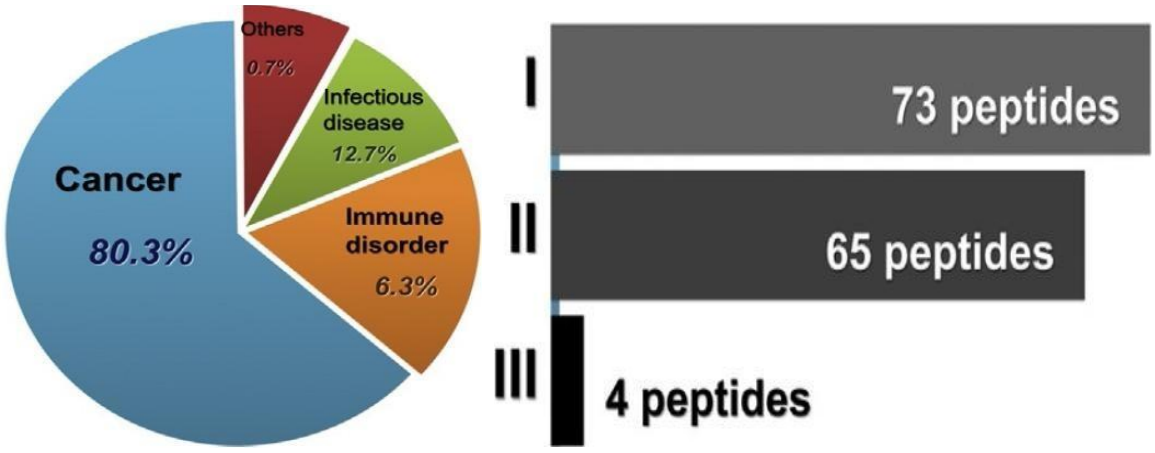


Figure 9: Peptide vaccines in different phases of clinical trial for various diseases (Yang et al., 2015)

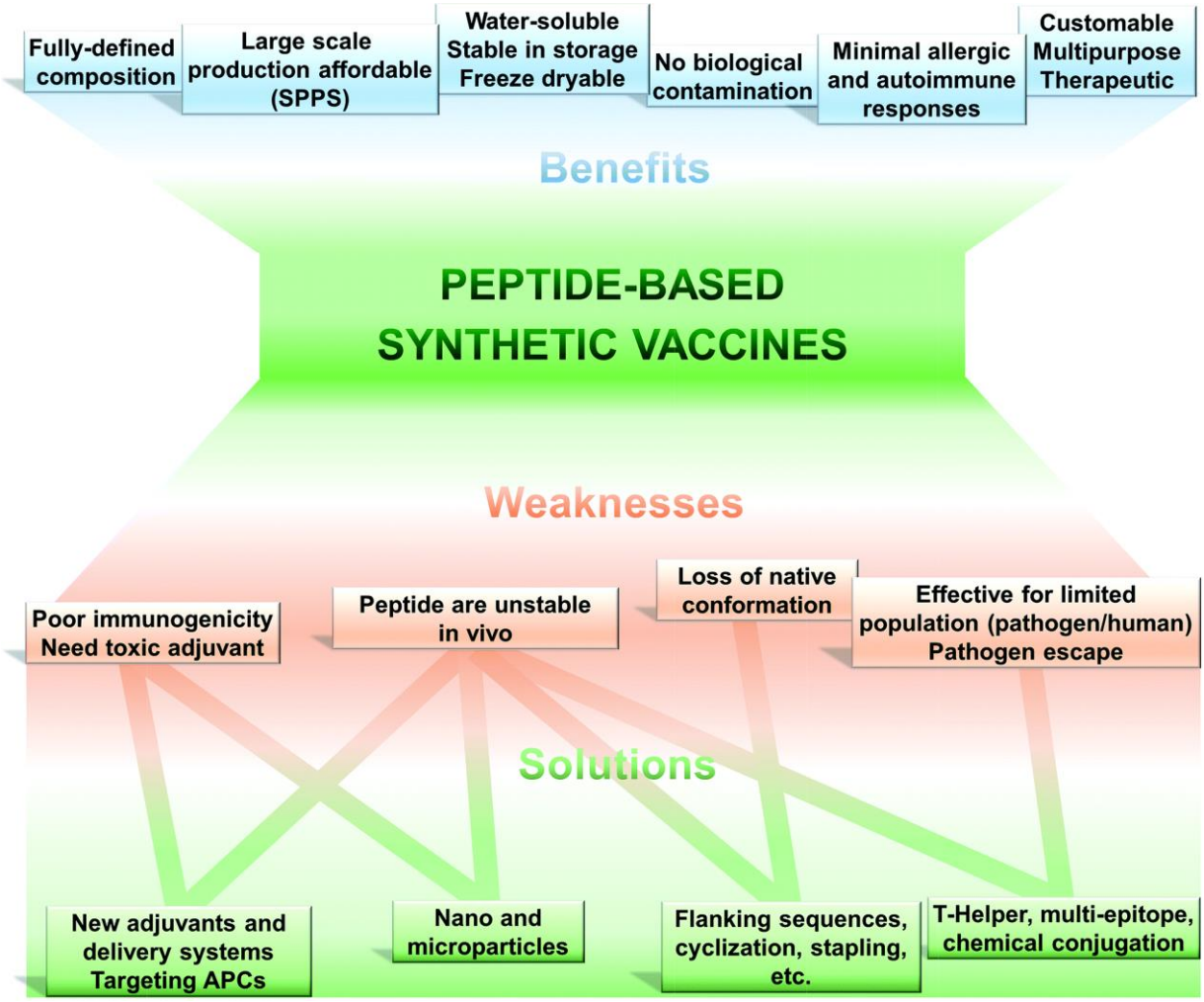


Figure 10: Pros, Cons and solutions of peptide based vaccines (Skwarczynski et al., 2016).

The molecular weight of a molecule should be more than 5,000 Daltons for eliciting an immune response, since peptide contains 9-30 amino acid fragment and low molecular weight, thereby they are unable to elicit an effective immune response but this problem can be solved by conjugating peptide with adjuvants and an optimized in vivo delivery method (De Brito et al., 2018; Li et al., 2014).

2.14 Immunoinformatics

Immunoinformatics is a branch of bioinformatics, termed as immunomics which deals with study of various mechanisms of immune system (Tomar et al., 2010). Its first prediction tool was for predicting HLA binding peptides and is widely used for identification of T cell and B cell epitopes of various diseases for the development of a drug and vaccine (Backert et al., 2015; Yan et al., 2010). The method of development of vaccines using immunoinformatics is termed as reverse vaccinology (Tomar et al., 2010). It reduces experimentation time and cost of treatment, since traditional methods of identification of immunogenic peptides involves synthesis of all possible epitopes and then evaluating immune response of each peptide *in vitro* and *in vivo*, which is a cumbersome process. It plays vital role in the development personalized medicines, which are designed on the basis of immune system of an individual (Yan et al., 2010). Various applications of immunoinformatics have been mentioned in Figure 11

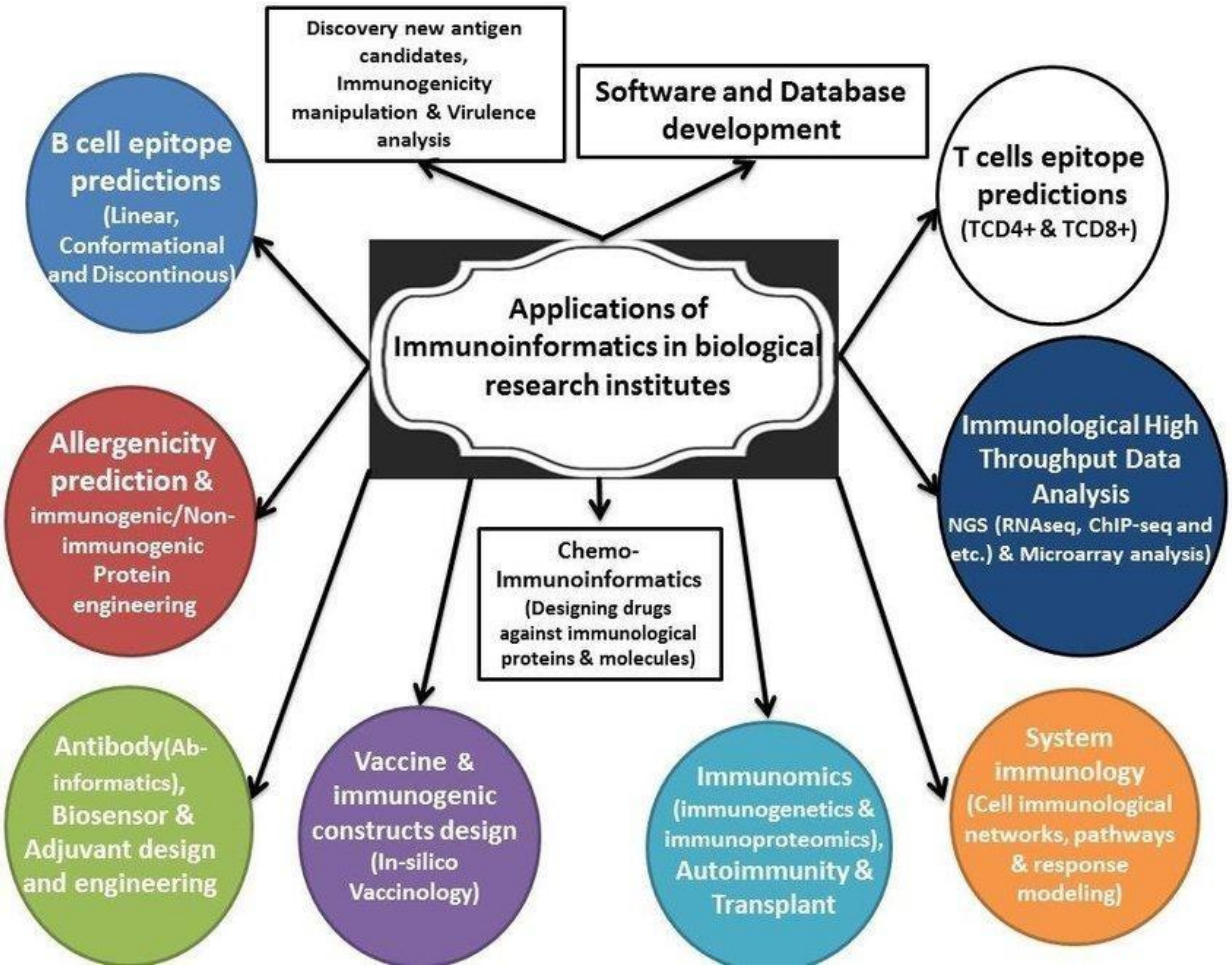


Figure 11: Various applications of Immunoinformatics (Ranjbar et al., 2019).

2.15 HLA polymorphism and Vaccine design

Human leukocyte antigens (HLA) present on chromosome 6, are subdivided into class I, class II and class III region. HLA molecules are also known as major histocompatibility complexes (MHCs). 40% of HLA gene loci participate in various functions of immune system, (Sasazuki et al., 2016). They recognize foreign antigens and elicit an immune response; consequently leads to commencement of surplus events such as inflammation, production of antibodies and cellular cytotoxicity (Posch et al., 2011). They play an important role in transplantation rejection reactions, development of placenta during pregnancy. Cancerous cells endure some modifications to circumvent recognition by immune cells (Monas et al., 2019). Class I and class II regions are further categorized as classical and non-classical HLA molecules. Class III regions facilitates stress response and complement cascade (Petersdorf et al., 2019). HLA class

I molecules consist of three classical HLA (i.e. HLA-A, B and C) and three non classical HLA (i.e. HLA-E, F, G) whereas HLA class II molecules consist of three classical HLA (HLA-DP, DQ and DR) and two non-classical HLA (i.e. HLA-DM and DO). (Monas et al, 2019; Sasazuki et al, 2016; Dos et al,2015; Pratheek et al, 2014). Classical HLA molecules are highly polymorphic whereas non-classical HLA molecules are less polymorphic among diverse ethnic groups. Classical HLA molecules are composed of a heavy chain and β 2- microglobulin (β2m).

Neoplastic cells present very high levels of HLA class I molecule whereas moderate level of classical HLA class I molecules are presented by most of the tissues (Pratheek et al, 2014; Posch et al., 2011). HLA Class II molecules are presented by finite cells only such as antigen presenting cells (APCs) and T lymphocytes and expression level of HLA-DR is highest on antigen presenting cells (APCs), followed by HLA-DQ and eventually HLA-DP (stern et al., 2009).

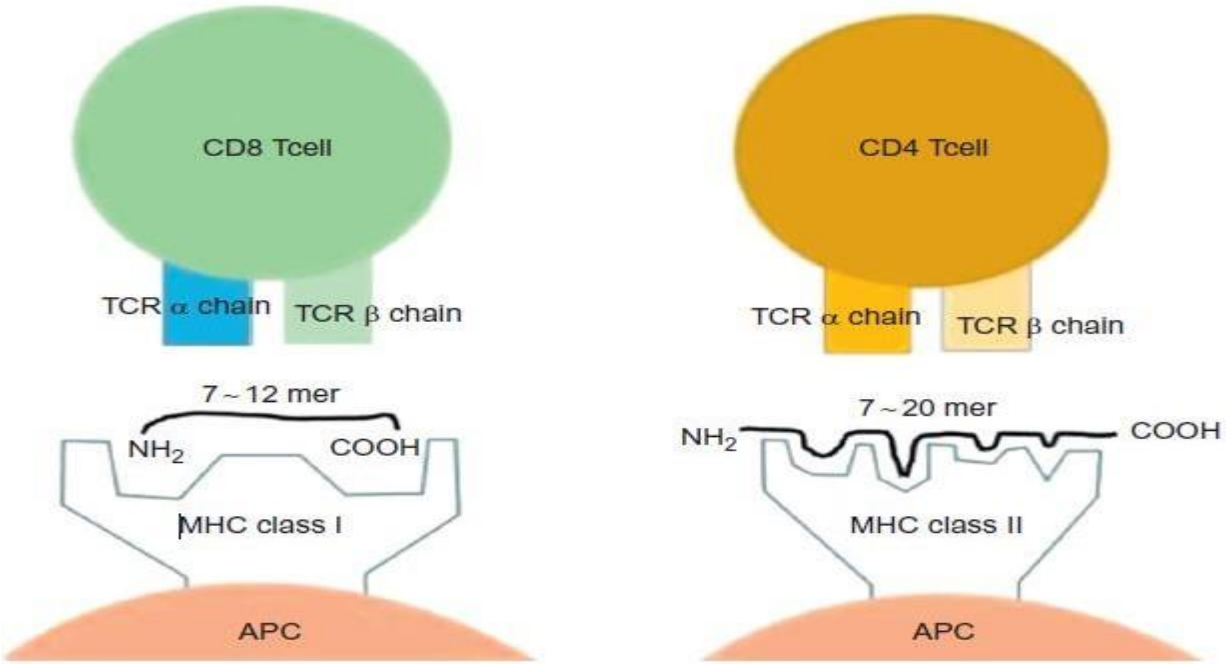


Figure 12: presentation of processed peptides by class I and class II MHCs to T_H and T_C cells (sasazuki et al., 2016)

HLA molecules are traversed at telomeric region, centromeric by HLA class I and class II regions respectively (Figure 13). Class I region can accommodate 7-12 amino acid peptide in its groove, whereas class II can accommodate 7-20 amino acid peptide (Figure 12 and Figure

14). Antigen binding groove of class I molecules is closed at both ends whereas class II molecule groove is open at both ends (Sasazuki et al., 2016).

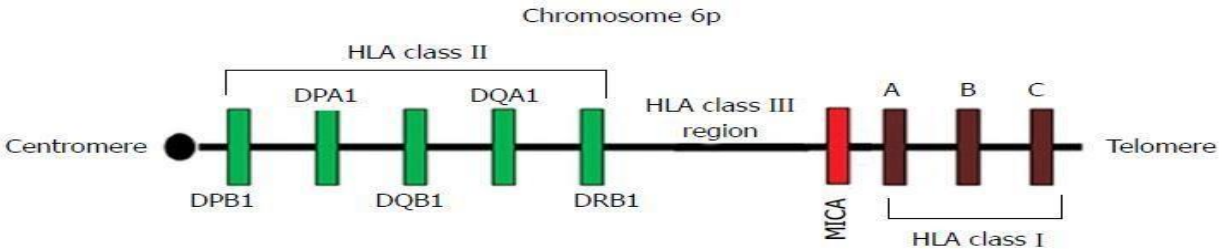


Figure 13: Structural organization of MHC regions on human MHC genome (Edinur et al., 2016).

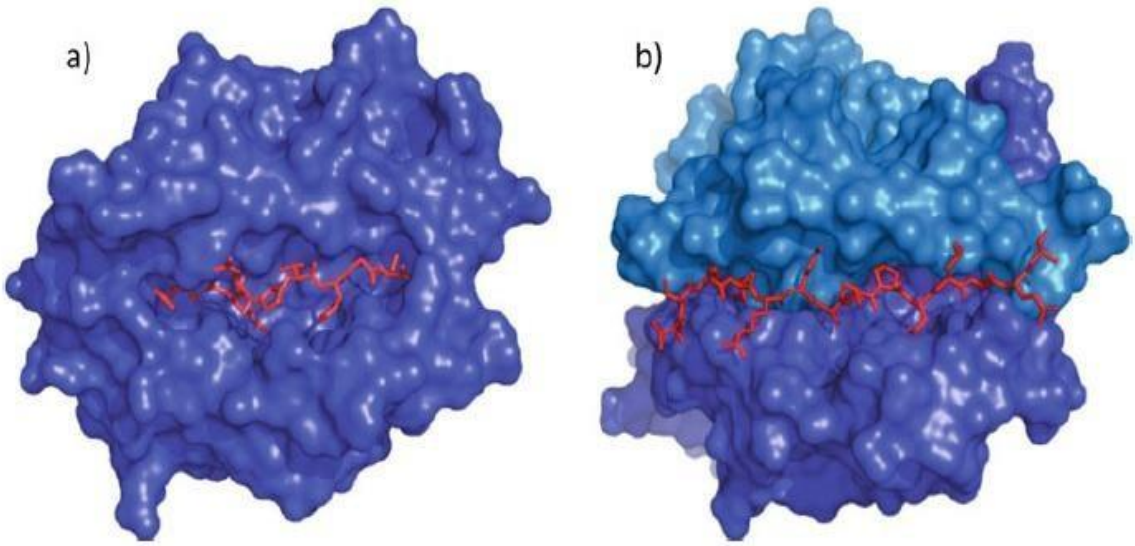


Figure 14: (a) Binding groove of HLA class I molecule accommodates 9-mer amino acid peptide and is closed at both ends (b) Binding groove of class HLA II molecules can accommodate larger peptide (i.e. between 9-20 amino acids) and is open at both ends (Andreatta et al., 2018).

According to April, 2019 report, 22,528 HLA alleles have been reported in IPD/IMGT database (EMBL-EBI, 2019).

Various factors such as age, gender and genetic factors trigger immune response upon vaccination. HLA polymorphism is the major genetic factor for immunization by presenting processed foreign peptide to T cells; thereby study of interaction of immunogenic peptide against a virus with HLA molecule is pivotal for vaccine development (Stern et al., 2009). Classical HLA molecules present processed peptides to T cells whereas non-classical HLA molecules either prompts activation or deactivation of natural killer (NK) cells (Halenius et al.,

2014). Therefore in this study, only classical HLA molecules were considered for evaluating HLA-peptide interaction for identification of potential vaccine candidates.

2.16 Epitope prediction algorithm

Several prediction methods have been developed over time for identification of immunogenic epitopes, among them hidden Markov model, position specific matrices are the foremost methods, followed by machine learning methods such as artificial neural network (ANN) and support vector machines (SVM). With expanding knowledge about process of presentation of peptide by MHCs to T cells, predictions for transport associated with antigen presentation (TAP) binding and proteasomal cleavage has been introduced to class I MHC prediction model (Lundegaard et al.,2010).

In the present study T cell epitope prediction was carried out by multiple algorithms for ensuring the specificity and selectivity of identified epitopes.

2.16.1 $CD8^+$ T cell epitope prediction tools

2.16.1.1 SYFPEITHI

SYFPEITHI constitute more than 7000 sequences and is primarily used for identification of class I and II MHC molecules, MHC motifs, T cell epitopes, MHC alleles etc. It is freely accessible for public use. It operates on motif based approach and the peptides are scored on the basis of whether they are anchor, auxiliary or preferred residue (Rammensee et al., 1999). Ideal, unusual, auxiliary and preferred anchors are scored 10, 6-8, 4-6 and 1-4 points respectively. It is 80% authentic for predicting HLA class I molecules and 50% for class II MHC molecules (www.syfpeithi.de).

2.16.1.2 NetCTL1.2

NetCTL1.2 speculates class I epitopes on the basis of proteasomal bifurcation, TAP transport regulation and binding affinity of 12 HLA class I supertypes with peptides by employing artificial neural network and weight matrix based method (Larsen et al., 2007). Its prognosis execution is higher than EpiJen, MAPPP and MHC-pathway tools.

2.16.1.3 IEDB

The immune epitope database (IEDB) employs artificial neural network (ANN) for HLA class I epitope prognosis for various infectious diseases, allergy, transplant etc. in humans and primates. It is a freely accessible tool engaged in the evolution of diagnostic therapeutics (Dhanda et al., 2019). It constitutes eight prediction methods such as consensus, netMHC pan 4.0, ANN 4.0, stabilized matrix method (SMM), netMHCcons, PickPocket, NetMHCstabpan (tools.iedb.org) and peptides having 8-`14 amino acids are accepted in FASTA format. It can predict peptides for eight host species: Human, cow, pig, mouse, gorilla, rat, and chimpanzee and macaque and constitutes 3,600 alleles (Fleri et al., 2017).

2.16.2 $CD4^+$ T cell epitope prediction tools

2.16.2.1 MHC2Pred

MHC2pred is a machine learning SVM tool which operates input sequence in binary form classifies nine amino acid or more peptides as binder (+1) or non-binder (-1) and speculates 42 class II MHC alleles (Lata et al., 2007). The method has been authenticated by five-fold cross-validation (crdd.osdd.net) Matrix optimization techniques (MOTs) have been employed for recognition of core peptide. It is ~86% accurate, ~85% sensitive and ~87% specific (Bhasin et al., 2003).

2.16.2.2 ProPred

Propred employs quantitative matrix based prediction method in which peptide is highlighted with the red colour of first amino acid followed by the blue colour of rest 8 amino acids. This graphics based tool evaluates antigenic peptides for 51 HLA-DR alleles and reads input sequence by using ReadSeq program (Singh et al., 2001). The binding scores are presented in tabular form. The sequence can be uploaded in FASTA, EMBL, and PIR format. Percentage threshold parameter permits to choose stringency levels for attuning the prediction results; as lower levels refer to a high stringency and higher levels refer to lower stringency (webs.iiitd.edu.in)Threshold value higher more than 3% should not be taken for prediction, since the stringency of epitope prediction will decrease drastically and might give false results. The results can be obtained in three formats: HTML view I, HTML view II and graphical view.

2.16.2.3 IEDB

IEDB has six algorithms for prediction of class II peptides viz., consensus, sturniolo, NetMHCIIpan, combinatorial library, NN-align and SMM-align for two host species (human and mouse). It constitutes 17, 128, 28, 104 and 256 alleles for DPA, DPB, DQA, DQB and DR respectively (Fleri et al., 2017).

3. Objectives

1. Design of peptides to contain multiple T cell epitopes based on consensus prediction algorithm.

2. Computational analysis of interaction between peptides and diverse HLA molecules.

4. Material and Methods

4.1 Sequence retrieval

DENV-2 envelope protein sequences were retrieved from virus pathogen resource (ViPRbrc) database. Total number of DENV E protein sequences till September 2018 were 1237, but among them 342 unique sequences were downloaded in FASTA format by chosing the option "Remove duplicate genome sequences".

4.2 Identification of conserved regions

Multiple sequence comparison by Log – expectation (MUSCLE) was employed for multiple sequence alignment (MSA). It is a reliable, accurate and fast tool incomparison to other MSA tools such as T- Coffee, MAFFT and CLUSTALW and can align upto 500 sequences with upto 500 amino acids. MUSCLE operates on two distance methods, a *k*mer distance and Kimura distance method for aligned and unaligned sequences respectively (Edgar et al., 2004). It can give results in various formats (FASTA, Clustal W, HTML, and Phylip) In this study results were downloaded in FASTA format (EBI, 2019) and used as input for AVANA during identification of conserved regions.

Conserved regions were identified in E protein of DENV with the help of antigenic variability analyzer (AVANA) tool. Conserved regions prompt stability, less variation and promising treatment. Since four serotypes of DENV are immunogenitically distinct; hence 70% conserved regions were chosen for further analysis by using AVANA which works on ethereum algorithm. It calculates entropy, frequency at each position and differences in peptides (Moitto et al., 2008). The output result of MUSCLE was loaded for analysis and it can read input in either .FASTA format or .TALN format. Range of peptide length was set to 9-25 amino acids. Since the core peptide which binds with groove of MHC is nine amino acids in length. The output result was analyzed for finding conserved fragments by melding overlapping regions.

4.3 Prediction of peptides containing $CD8^+$ T cell epitopes

In the present study, SYPEITHI, NetCTL1.2 and IEDB tools were used for epitope prediction for identification of epitopes by multiple algorithms, thereby ensuring the specificity and selectivity of identified epitopes.

The threshold values for prediction of $CD8^+$ T cells in SYPEITHI, NetCTL1.2 and IEDB tools were 20 score, 0.75 threshold value and IC_{50} value upto 500 respectively(Jain et al., 2019; Lohia et al., 2015). The epitopes commonly predicted by the three tools were selected. The overlapping epitopes were merged to obtain peptide fragments containing multiple $CD8^+$ T-cell epitopes.

4.4. Prediction of peptides containing $CD4^+$ T cell epitopes

Three tools viz., MHC2Pred, ProPred and IEDB were used for epitope prediction.

The threshold values for prediction of $CD4^+$ T cells in MHC2Pred was 3% whereas in both ProPred and IEDB tools threshold values were IC_{50} value less than 500.

The epitopes commonly predicted by the three tools were selected. The overlapping epitopes were merged to obtain peptide fragments containing multiple $CD4^+$ T-cell epitopes. Further, the regions common to peptides containing $CD8^+$ and $CD4^+$ epitopes were identified so as to obtain peptide fragments containing both types of epitopes.

4.5 Peptide screening for allergic, autoimmune and toxic

The allergenicity of predicted peptides was identified by AlgPred tool, which employs six approaches V.IZ., SVM method based on amino acid composition and dipeptide composition, Blast search and hybrid approach (webs.iiitd.edu.in).In this study mapping of IgE epitope and PID and MEME/MAST motif options were selected for analysis. Former is 93.94% sensitive and 33.34% specific and latter is 17.47% sensitive and 98.14% specific. If the peptide has similarity among at least six consecutive amino acids is considered as allergic (Saha et al., 2006).

The toxicity of peptides was analyzed by using ToxinPred, which calculates toxicity by employing quantitative matrix based method (crdd.osdd.net).

Basic Local Alignment Search Tool (BLAST) is extensively used for computing differences and similarity in protein or nucleotide sequence. It is a heuristic program which takes shortcuts

for executing fast searches (Madden et al., 2013).The predicted peptides containing both class I and class II epitopes were rejected if similarity was found in seven consecutive amino acids with respect to humans proteome (Jain et al., 2019).

4.6 Mapping of peptide fragments

The three dimensional structure of dengue virus envelope protein (PDBid: 3J2P) was downloaded from protein data bank (PDB). Discovery visual studio v3.5 tools was used for mapping of identified peptides on the downloaded structure.

4.7 Population coverage analysis

HLA polymorphism is very intricate and gives different frequencies among diverse ethnic groups. Population coverage analysis IEDB tool was employed to estimate the global HLA coverage of the predicted peptides as it provides allelic frequency for 115 countries and 21 ethnic groups (Adhikari et al., 2018). It is crucial for development of an effective peptide based vaccine among diverse ethnic groups (Bui et al., 2006).

4.8 HLA-Peptide interaction analysis via Docking

Docking is significant for evaluating interaction between peptide and HLA molecules via in-silico approach for development of vaccine**.** Various methods have been developed for docking which provide binding affinity scores by using either force-field, empirical or knowledge based method (Ferreira et al., 2015). Lower energy scores corresponds to good protein-ligand binding (Thomsan et al., 2006).

In this study, CABS-dock and Molecular virtual docker tools were employed for performing HLA-Peptide docking for analyzing the root mean square deviation value (RMSD) and binding energies of HLA-Peptide complex respectively. The HLA class I and class II molecules selected for HLA-peptide interactions via docking has been mentioned in Table 7.

Table7: PDBid, allele and resolution of HLA molecules

Class	PDBid	HLA allele	Resolution (A°)
Class I	3BO8	A1	1.8
	3MRE	A2	1.1
	1X7Q	A*1101	1.45
	3WL9	A24	1.66
	3SPV	B8	1.3
	3C9N	B*1501	1.87
	4XXC	B*1801	1.426
	1K5N	B*2709	1.09
	3BWA	B*3508	1.3
	2HJL	B5703	1.5
Class II	1D5M	DR4	1.72
	5NIG	DRB1*04:01	1.35
	3C5J	DR52c	1.8
	1S9V	DQ2	2.22
	6DIG	DQ6	2
	4P5M	DP	1.7

4.8.1 CABS-dock

CABS-dock is freely accessible online public server for protein-peptide docking. Only the receptor molecule and peptide sequence is uploaded for docking, unlike other methods, it defines binding groove by itself. It operates upon coarse-grained CABS model for docking. The RMSD value <3A°, between 3A° to 5.5A° and more than 5.5A° corresponds to high, medium and low quality prediction (Blaszczyk et al., 2015; Kurcinski et al., 2015). The procedure for performing docking in CABS-dock is mentioned in Figure 15.

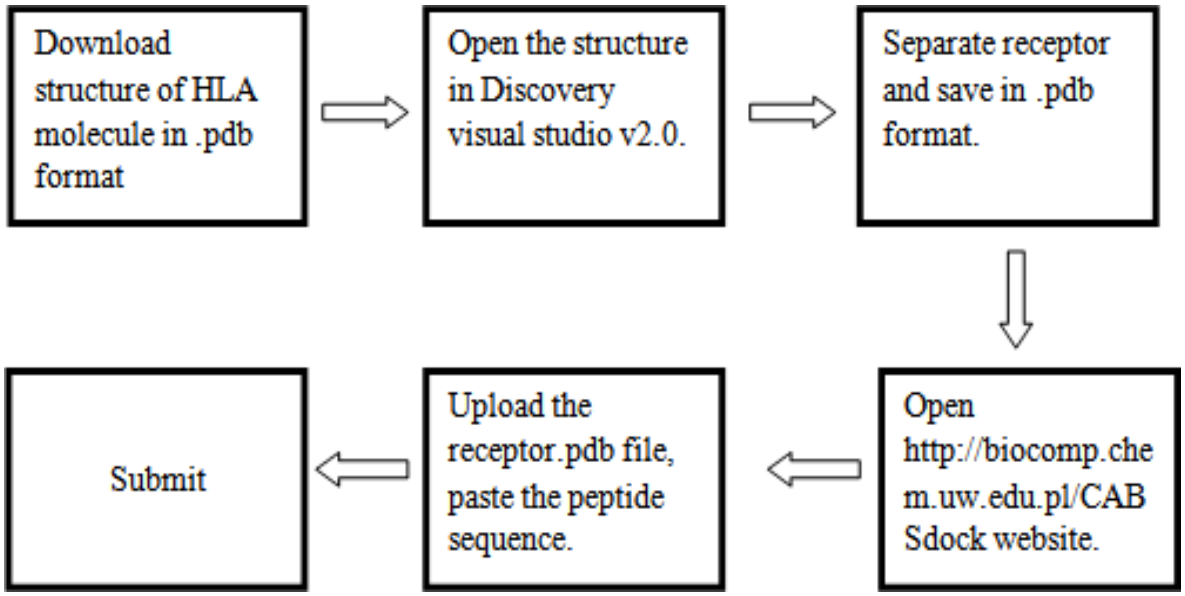


Figure 15: Procedure for performing docking by using CABS-dock.

4.8.2 Molecular virtual docker

Molecular virtual docker (MVD) or MolDOCK a stochastic method employs guided differential evolution and force field based scoring function and exhibits 87% accuracy (Kusumaningrum et al., 2014). Former is a combination of differential evolution and cavity detection algorithm. The speed and efficiency of docking is enhanced by the cavity detection algorithm (Thomsan et al., 2006). The binding site or groove is automatically recognized by cavity detection algorithm.

4.8.2.1 Generation of peptide structures

The three dimensional structure of identified peptide was generated by using PEP-FOLD which operates on greedy algorithm and a coarse grained approach (PEP-FOLD, 2018).

4.8.2.2 Docking using Molecular virtual docker

The procedure for performing docking is shown in Figure 16 a), 16 b) and 16 c).

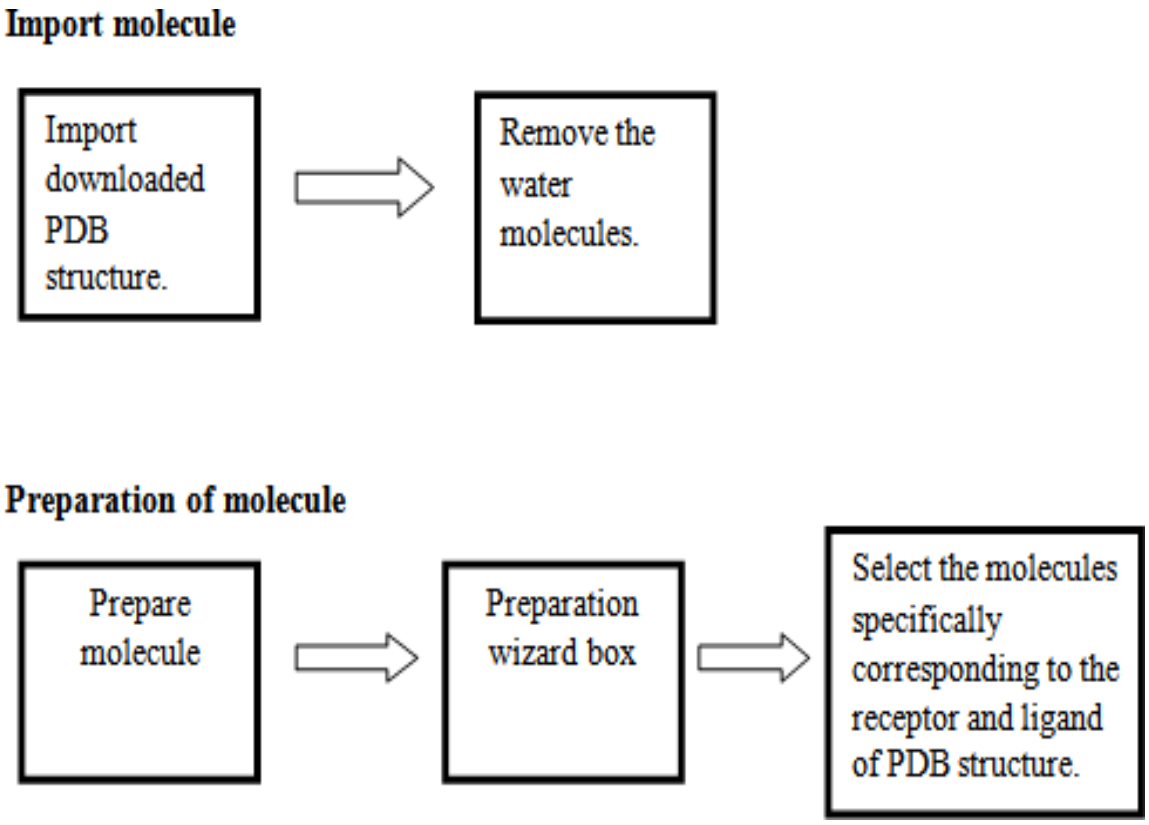


Figure 16 a) Procedure for import and preparation of molecule

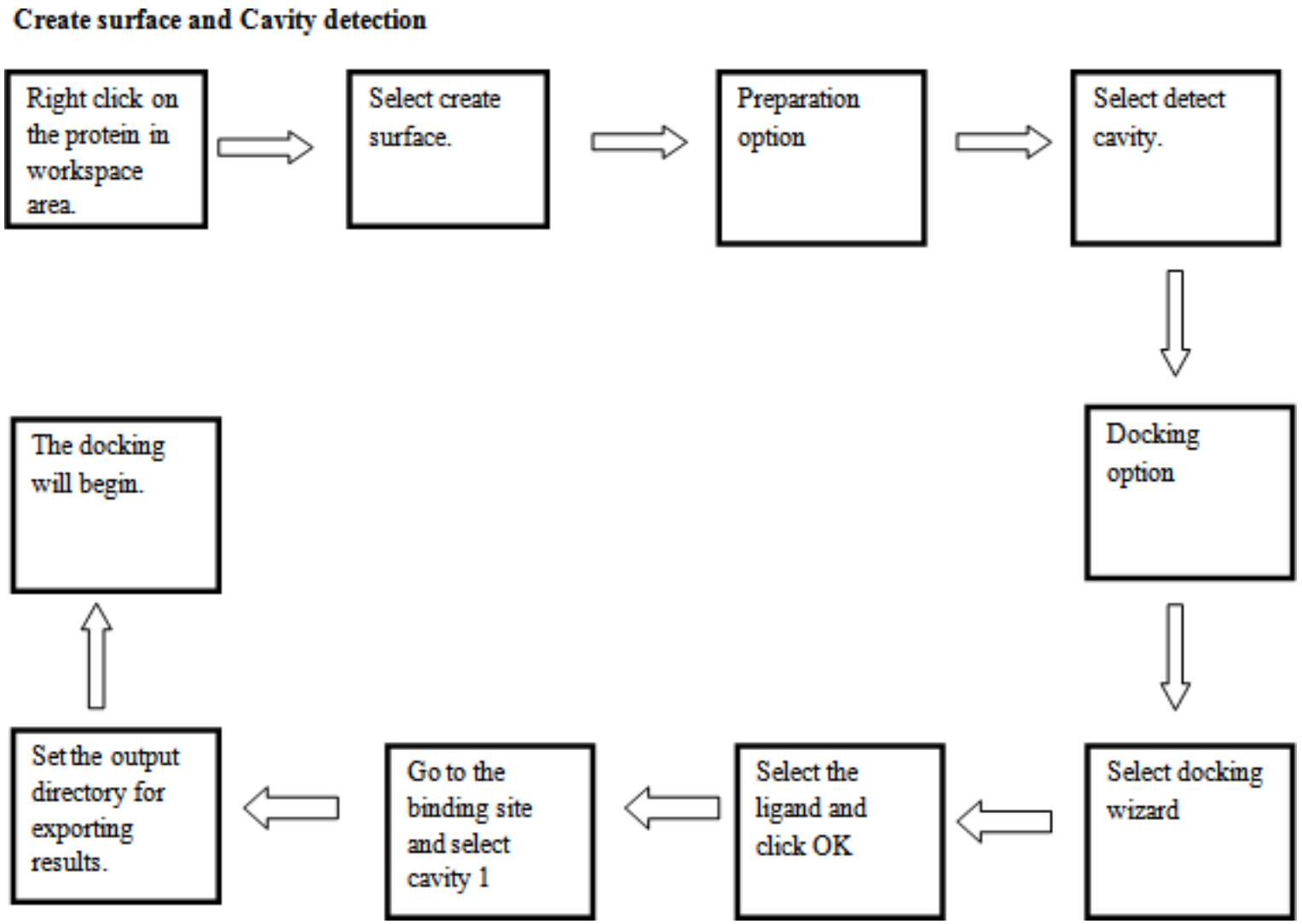


Figure 16 b): Procedure for creating surface and cavity detection

Result analysis and export pose

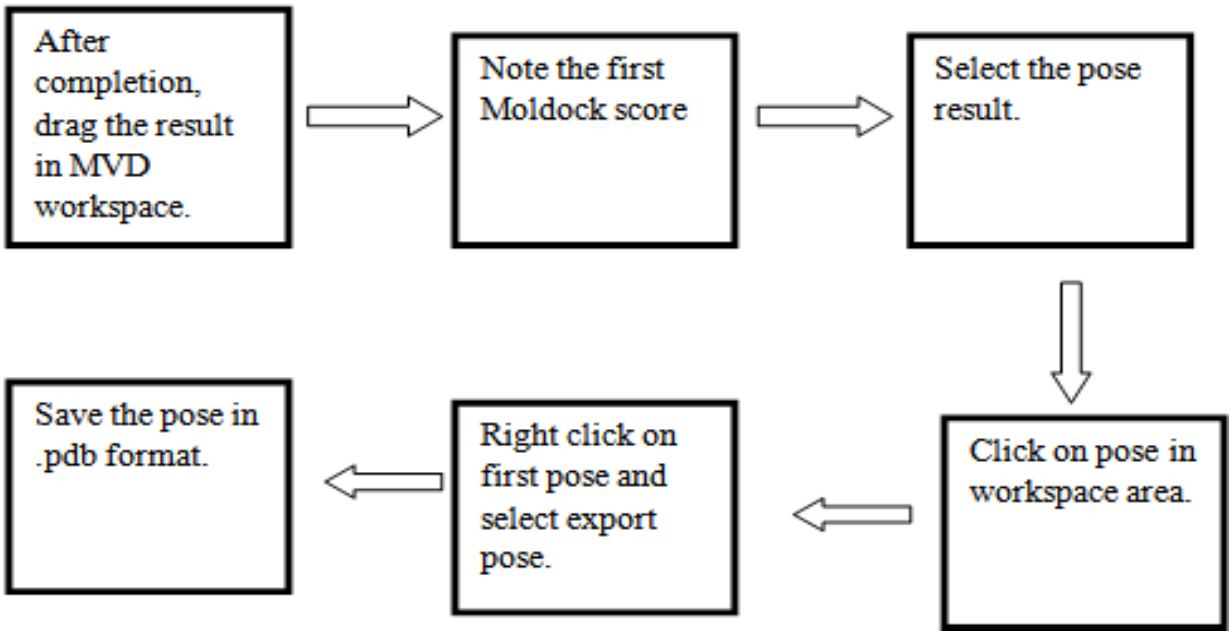


Figure 16 c): Procedure for result analysis and export pose

4.9 Conservancy analysis in other dengue serotypes and flaviviruses

The conservancy percentage of the identified peptide was evaluated individually among the total number of sequences for each virus mentioned above by searching peptides in AVANA. The conservancy of each peptide other than identified peptides was also recorded (Table 17).

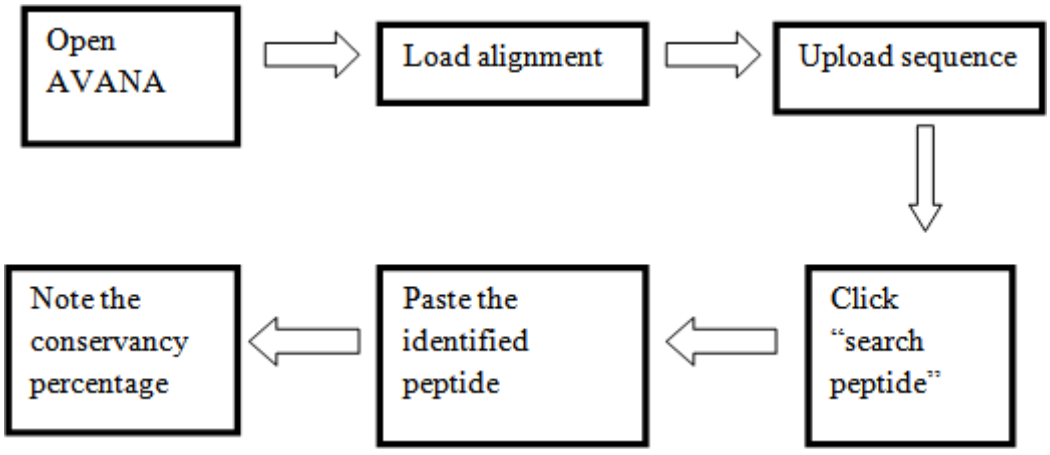


Figure 17: Procedure for evaluating conservancy of identified peptides using AVANA

5. Results

5.1 Identification of conserved regions in Dengue envelope protein

342 unique sequences of DENV serotype II E protein (1944 to Sep, 2017) were downloaded from the virus pathogen database and analysis resource (ViPRbrc) database in FASTA format, followed by multiple sequence alignment by using MUSCLE tool. The resultant .fasta file was used as input for AVANA tool to find regions displaying ≥70% conservancy. Nine such peptide fragments were identified (Table 8) and used as input for various epitope prediction tools.

Table 8: Conserved peptide fragments of Dengue envelope protein

	Conserved peptide fragments	Length
E1	MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTE AKQPATLRKYCIEAKLTNTTT	70
E2	SRCPTQGEPSLNEEQDKRF	19
E3	CKHSMVDRGWGNGCGLFGKGGIVTCAMFTCKKNMEGK	37
E4	VQPENLEYTIVITPHSGEE	19
E5	AVGNDTGKHGKEIK	14
E6	TPQSSITEAELTGYGTVTMECSPRTGLDFNEMVLLQME	38
E7	KAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQD VVVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCRLRMDKLQLK GMSYSMCTGKFK	104
E8	VKEIAETQHGTIVIRVQYEGDGSPCKIPFEIMDLEKRHVLGRLITVNPIVT EKDSPVNIEAEPPFGDSYIIIGVEPGQLKLNWFKKGSSIGQMFETTMRG AKRMAILGDTAWDFGSLGGVFTSIGKALHQVFGAIYGAAFSGVSWTM KILIGV	153
E9	ITWIGMNSRSTSLSVSLVLVGVVTLYLG	28

5.2 Epitope prediction

5.2.1 Prediction of $CD8^+$ T cell epitopes

A consensus approach was applied for epitope prediction so as to increase the prediction stringency and cater to different immunological parameters. Three tools SYFPEITHI, NetCTL 1.2 and IEDB were employed for $CD8^+$ T cell epitope prediction. Total number of epitopes predicted by SYFPEITHI, NetCTL1.2 and IEDB tool were 172, 129 and 234 respectively. Fifty two epitopes commonly predicted by the three tools were identified (Table 9). 14 peptide fragments (Table 10) containing multiple $CD8^+$ T cell epitopes were obtained by merging the overlapping epitopes.

Table 9: $CD8^+$ T cell epitopes in Dengue envelope protein predicted commonly by three epitope prediction tools

$CD8^+$ T-cell epitopes			
TMAKNKPTL	RLRMDKLQL	KAWLVHRQW	AEPPFGDSY
GSCVTTMAK	WLVHRQWFL	MDLEKRHVL	GRLITVNPI
KQPATLRKY	KLQLKGMSY	GTIVIRVQY	GQLKLNWFK
TEAKQPATL	NLLFTGHLK	SIGKALHQV	KRHVLGRLI
RKYCIEAKL	VTFKNPHAK	HVLGRLITV	AETQHGTIV
CPTQGEPSL	MSYSMCTGK	SLGGVFTSI	SLSVSLVLV
VTCAMFTCK	YSMCTGKFK	RLITVNPIV	LVLVGVVTL
LEYTIVITP	SYSMCTGKF	WTMKILIGV	SVSLVLVGV
ITEAELTGY	IQKETLVTF	TVNPIVTEK	GMNSRSTSL
GLDFNEMVL	RQWFLDLPL	QLKLNWFKK	STSLSVSLV
MECSPRTGL	IQMSSGNLL	ILGDTAWDF	VLVGVVTLY
AELTGYGTV	LRMDKLQLK	FSGVSWTMK	TSLSVSLVL
FLDLPLPWL	QEGAMHTAL	DTAWDFGSL	RSTSLSVSL

Table 10: Dengue envelope peptide fragments containing multiple $CD8^+$ T cell epitopes.

Peptide fragments containing multiple $CD8^+$ T cell epitopes	Position	Epitopes
GSCVTTMAKNKPTL	28-41	GSCVTTMAK, TMAKNKPTL
TEAKQPATLRKYCIEAKL	48-65	TEAKQPATL, RKYCIEAKL
ITEAELTGYGTV	170-181	ITEAELTGY, AELTGYGTV
MECSPRTGLDFNEMVL	183-198	MECSPRTGL, GLDFNEMVL
KAWLVHRQWFLDLPLPWL	204-221	KAWLVHRQW,WLVHRQWFL, RQWFLDLPL, FLDLPLPWL
IQKETLVTFKNPHAK	232-246	IQKETLVTF, VTFKNPHAK
IQMSSGNLLFTGHLK	270-284	IQMSSGNLL, NLLFTGHLK
RLRMDKLQLKGMSYSMCTGKFK	286-307	RLRMDKLQL,LRMDKLQLK, KLQLKGMSY,MSYSMCTGK, SYSMCTGKF, YSMCTGKFK
AETQHGTIVIRVQY	313-326	AETQHGTIV, GTIVIRVQY
MDLEKRHVLGRLITVNPIVTEK	340-361	MDLEKRHVL, KRHVLGRLI, HVLGRLITV,GRLITVNPI, RLITVNPIV, TVNPIVTEK
GQLKLNWFKK	385-394	GQLKLNWFK,QLKLNWFKK
ILGDTAWDFGSLGGVFTSIGKALHQV	414-439	ILGDTAWDF,DTAWDFGSL, SLGGVFTSI, SIGKALHQV
FSGVSWTMKILIGV	448-461	FSGVSWTMK,WTMKILIGV
GMNSRSTSLSVSLVLVGVVTLY	467-488	GMNSRSTSL,RSTSLSVSL, TSLSVSLVL,SLSVSLVLV, STSLSVSLV,SVSLVLVGV, LVLVGVVTL,VLVGVVTLY

5.2.2 Prediction of $CD4^+$ T cells epitopes

Three tools viz., MHC2Pred, ProPred and IEDB were used for $CD4^+$ T cell epitope prediction. The total number of $CD4^+$ T cell epitopes predicted from ProPred, MHC2Pred and IEDB were 559, 1829 and 1899 respectively (Data not shown). Fifty six epitopes commonly predicted by the three tools were identified (Table11) and twelve peptide fragments were obtained (Table 12) by merging the overlapping epitopes.

Table 11: $CD4^+$ T cell epitopes in Dengue envelope protein identified commonly by three epitope prediction tools

$CD4^+$ T cell epitopes			
IVLEHGSCV	LRMDKLQLK	MRGAKRMAI	VLGRLITVN
IGISNRDFV	WFLDLPLPW	FGSLGGVFT	VFTSIGKAL
FVEGVSGGS	WIQKETLVT	FGAIYGAAF	VSWTMKILI
FELIKTEAK	LKGMSYSMC	ILGDTAWDF	IMDLEKRHV
YCIEAKLTN	WLVHRQWFL	IIIGVEPGQ	FEIMDLEKR
LRKYCIEAK	LVHRQWFLD	FKKGSSIGQ	YGAAFSGVS
FGKGGIVTC	LKCRLRMDK	FTSIGKALH	FETTMRGAK
FTCKKNMEG	LQLKGMSYS	LGRLITVNP	FSGVSWTMK
YTIVITPHS	VTFKNPHAK	IVIRVQYEG	LNWFKKGSS
LTGYGTVTM	LLFTGHLKC	LEKRHVLGR	WIGMNSRST
FNEMVLLQM	FTGHLKCRL	LGGVFTSIG	LVLVGVVTL
LVTFKNPHA	YSMCTGKFK	VNPIVTEKD	LVGVVTLYL
VVVLGSQEG	LITVNPIVT	IIGVEPGQL	IGMNSRSTS
IQMSSGNLL	YIIIGVEPG	WFKKGSSIG	VLVGVVTLY

Table 12: Dengue envelope peptide fragments containing multiple $CD4^+$ T cell epitopes.

Peptide fragments containing multiple $CD8^+$ T cell epitopes	Position	Epitopes
IGISNRDFVEGVSGGS	4-19	IGISNRDFV,FVEGVSGGS
LRKYCIEAKLTN	56-67	LRKYCIEAK,YCIEAKLTN
WLVHRQWFLDLPLPW	206-220	WLVHRQWFL,LVHRQWFLD, WFLDLPLPW
WIQKETLVTFKNPHAK	231-246	WIQKETLVT, LVTFKNPHA, VTFKNPHAK,
IQMSSGNLLFTGHLKCRLRMDKLQL KGMSYSMCTGKFK	270-307	IQMSSGNLL, LLFTGHLKC, FTGHLKCRL,LKCRLRMDK, LRMDKLQLK,LQLKGMSYS, LKGMSYSMC , YSMCTGKFK
FEIMDLEKRHVLGRLITVNPIVTEKD	337-362	FEIMDLEKR, IMDLEKRHV, LEKRHVLGR, LGRLITVNP, LITVNPIVT ,VLGRLITVN, VNPIVTEKD
YIIIGVEPGQL	377-387	YIIIGVEPG, IIIGVEPGQ, IIGVEPGQL
LNWFKKGSSIGQ	389-400	LNWFKKGSS, WFKKGSSIG, FKKGSSIGQ
FETTMRGAKRMAILGDTAWDFGSL GGVFTSIGKALH	402-437	FETTMRGAK,MRGAKRMAI, ILGDTAWDF,FGSLGGVFT, LGGVFTSIG,VFTSIGKAL, FTSIGKALH
FGAIYGAAFSGVSWTMKILI	440-459	FGAIYGAAF,YGAAFSGVS, FSGVSWTMK,VSWTMKILI
WIGMNSRSTS	465-474	WIGMNSRST,IGMNSRSTS
LVLVGVVTLYL	479-489	LVLVGVVTL,VLVGVVTLY, LVGVVTLYL

5.3 Identification of peptides containing $CD8^+$ and $CD4^+$ T-cell epitopes

The peptide fragments containing $CD4^+$ and $CD8^+$ T cell epitopes individually were further analyzed for presence of common regions so as to obtain peptide fragments containing both, $CD8^+$ and $CD4^+$ T cells (Table 13). Nine such fragments were identified and used for further analysis (Table 14).

Table 13: Peptides containing both, $CD8^+$ and $CD4^+$ T-cell epitopes

Peptides containing $CD8^+$ T cell epitopes	Peptides containing $CD4^+$ T cell epitopes	Peptides containing both $CD4^+$ and $CD8^+$ T cell epitopes
TEAKQPAT**LRKYCIEAKL**	**LRKYCIEAKL**TN	**LRKYCIEAKL**
KA**WLVHRQWFLDLPLPW**L	**WLVHRQWFLDLPLPW**	**WLVHRQWFLDLPLPW**
IQKETLVTFKNPHAK	W**IQKETLVTFKNPHAK**	**IQKETLVTFKNPHAK**
IQMSSGNLLFTGHLK	**IQMSSGNLLFTGHLK**	**IQMSSGNLLFTGHLK**
RLRMDKLQLKGMSYSMCTGKFK	**RLRMDKLQLKGMSYSMCTGKFK**	**RLRMDKLQLKGMSYSMCTGKFK**
MDLEKRHVLGRLITVNPIVTEK	FEI**MDLEKRHVLGRLITVNPIVTEK**D	**MDLEKRHVLGRLITVNPIVTEK**
ILGDTAWDFGSLGGVFTSIGKALHQV	FETTMRGAKRMA**ILGDTAWDFGSLGGVFTSIGKALH**	**ILGDTAWDFGSLGGVFTSIGKALH**
FSGVSWTMKILIGV	FGAIYGAA**FSGVSWTMKILI**	**FSGVSWTMKILI**
LVLVGVVTLY	**LVLVGVVTLY**L	**LVLVGVVTLY**

Table 14: $CD8^+$ and $CD4^+$ T cell epitopes present in the identified peptides

Peptides containing both $CD4^+$ and $CD8^+$ T cell epitopes	Number of $CD8^+$ T cell epitopes	Number of $CD4^+$ T cell epitopes
LRKYCIEAKL	1	1
WLVHRQWFLDLPLPW	2	3
IQKETLVTFKNPHAK	2	1
IQMSSGNLLFTGHLK	2	2
RLRMDKLQLKGMSYSMCTGKFK	6	4
MDLEKRHVLGRLITVNPIVTEK	6	**3**
ILGDTAWDFGSLGGVFTSIGKALH	3	4
FSGVSWTMKILI	1	2
LVLVGVVTLY	2	2

5.4 Screening the identified peptides for autoimmune, allergic and toxic responses

The identified peptides (Table 13) were analyzed with the help of AlgPred and ToxinPred for allergic and toxic responses respectively. All peptides were found to be non-allergic as well as non-toxic. Four peptides: P3, P4, P5 and P6 exhibited similarity with human chloride channel protein, immunoglobulin heavy chain variable region, myosin protein and glucocorticoid receptor respectively during BLAST analysis. Therefore, these peptide fragments were eliminated from further studies. The rest of the peptides were chosen for further analysis (Table 15).

Table 15: Peptide fragments selected after screening for undesired responses.

Peptide	Location	Length
LRKYCIEAKL (P1)	56-65	10
WLVHRQWFLDLPLPW (P2)	204-220	15
ILGDTAWDFGSLGGVFTSIGKALH (P3)	414-437	24
FSGVSWTMKILI (P4)	448-459	12
LVLVGVVTLY (P5)	479-488	10

5.5 Mapping of identified peptides

The finally selected peptide fragments were located on the DENV E protein structure downloaded from protein data bank (PDB). The downloaded structure (PDBid: 3J2P) was edited in Discovery visual studio v2.0 for removing additional sequences and monomeric E protein structure was obtained. All the peptides were located on this structure (Figure 18).

Also, a schematic presentation of the peptide location was made so as to better understand the peptide functions (Figure 19). P1 and P2 peptide located in the domain II (DII) of E protein between 52-131 amino acids, participates in determination of serotype specificity, P3 and P4 peptide located in the stem segment between 392-449 amino acids, provides stability to three envelop protein domains viz., DI, DII and DIII and P5 peptide located in the transmembrane anchor region between 449-495 amino acids, facilitates interaction between three domains and stem segment.

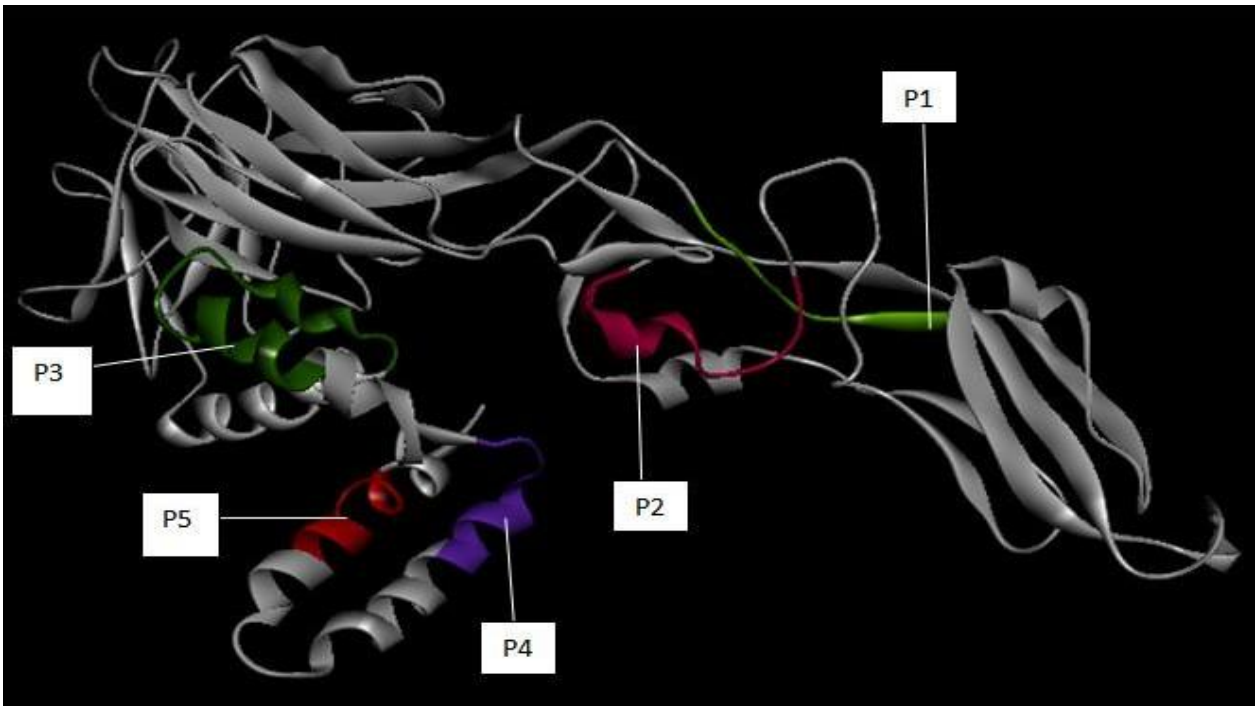


Figure 18: Mapping of peptides in three dimensional Dengue envelope protein

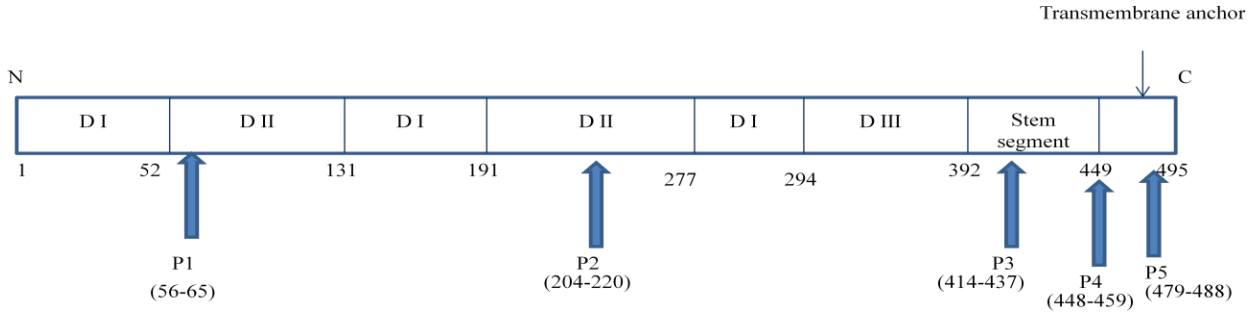


Figure 19: Location of peptides on dengue envelope protein

5.6 Population coverage analysis

For development of an effective peptide based vaccine, the percentage of population coverage across the various geographical regions should be high. Population coverage analysis deciphers the potential of peptides to induce an immune response in populations belonging to diverse geographical regions. IEDB population coverage tool was used for this analysis and evaluation was done for three continents (comprising of 13 geographical regions) viz., Asia, Africa and America. The population coverage in Europe, Oceania, West Indies and World was also analyzed. P1, P2. P3, P4 and P5 peptides exhibits 51.94, 98.14, 97.82, 91.71 and 96.37 percent average population coverage respectively in the six continents. P1 and P2 peptide have displayed least and highest population coverage respectively among six continents. Population coverage shown by P1, P2, P3, P4 and P5 peptides in Asia, Africa, America, Europe, Oceania, West Indies and World have been mentioned in table 16 and figure 20.

Table 16: Average population coverage in six continents and world

	Population coverage (%)				
Continents	**P1**	**P2**	**P3**	**P4**	**P5**
Asia	55.67	99.26	97.76	94.61	94.02
Africa	33.31	93.18	95.55	85.33	95.63
America	55.67	98.32	97.87	99.01	99.21
Europe	55.67	100	100	99.99	99.66
Oceania	55.67	99.82	98.42	98.14	91.67
West Indies	55.67	98.23	97.32	73.18	98.03
Average of all continents	51.94	98.14	97.82	91.71	96.37
World	55.67	99.95	80.89	99.32	99.13

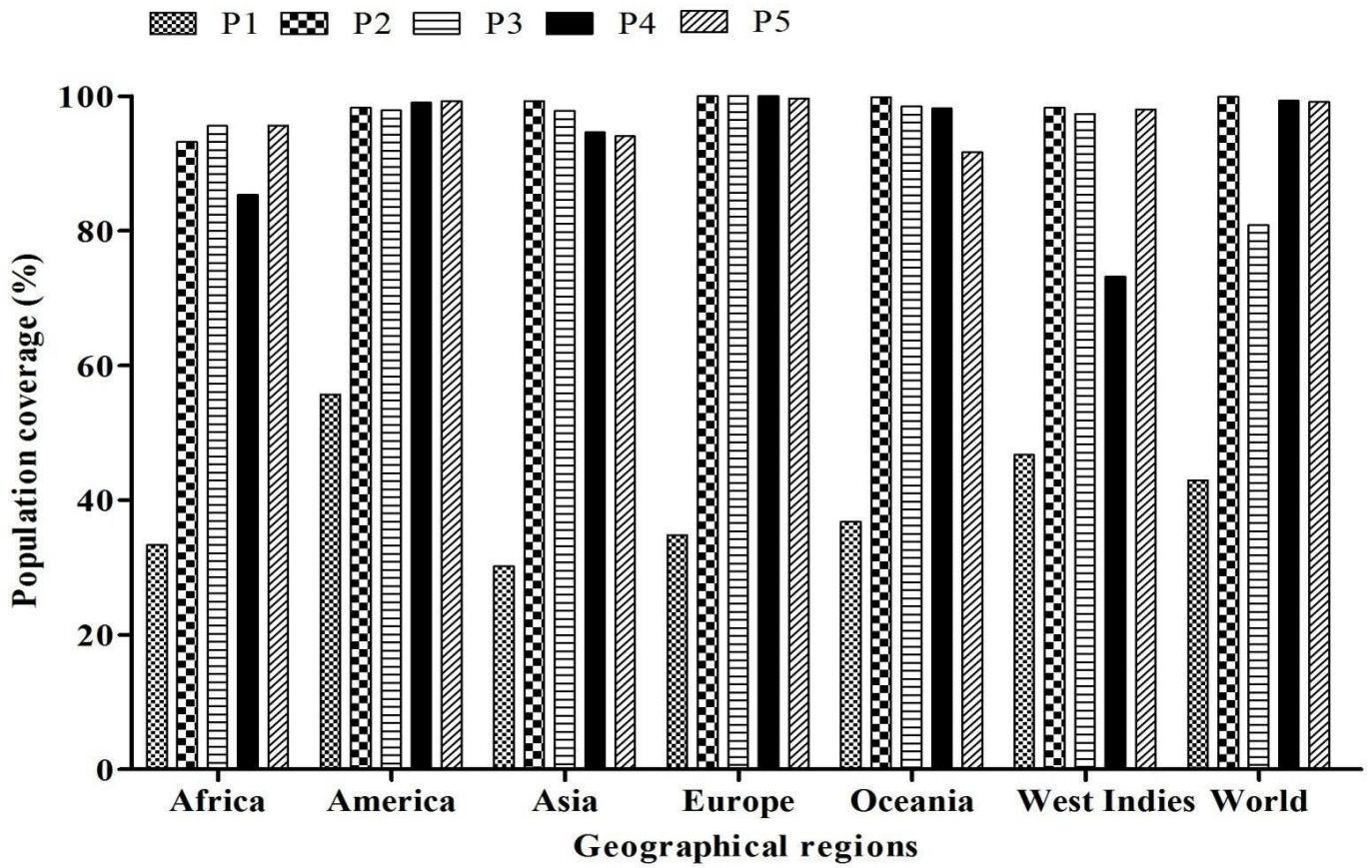


Figure 20: Population coverage analysis of identified peptides in three continents, Europe, Oceania, West Indies and world.

5.7 Molecular docking

Upon viral entry into host cells, its proteins are proteolyzed into small peptide fragments. Some of the peptides (epitopes) are presented by HLA molecules to T-cells which results in cell mediated immune response. Therefore, the binding efficiency of a peptide towards HLA molecule plays a significant role in induction of an immune response. This binding efficacy of a peptide for various HLA molecules can be evaluated by carrying out docking studies.

In this study, ten HLA class I and six HLA class II molecules representing different HLA alleles were selected for docking. Docking was carried out by using Molecular virtual docker and For CABS-dock and molecular virtual docker the native peptides (peptide bound to the downloaded HLA molecule) were separated from their respective HLA molecule with the help of Discovery Studio Visualizer 4.1 tool. CABS-dock and Molecular virtual docker results were analyzed by observing the binding energy and RMSD value respectively. The binder and non-binder peptides with HLA molecules were also analyzed by CABS-dock. Binding energy and RMSD value of native peptide was used as standard in case of Molecular virtual docker.

Results of Molecular virtual docker deciphers, high binding affinity of P2 peptide with all HLA class I molecules (table 17) as compared to native peptide. P1, P3, P4 and P5 have shown high binding affinity with eight, seven, five and four HLA class I molecules respectively with respect to native peptide (figure 18). In case of HLA class II molecules, four peptides have shown high binding affinity with all the six selected HLA class II molecules (table 18) whereas P5 peptide has shown high binding affinity for three HLA class II molecules with respect to native peptide (Figure 19).

Results of CABS-dock with HLA class I and II molecules deciphers that some of peptides were found to be binding outside the binding groove with few HLA molecules, which were considered as non binders and were not plotted (Figure 20 and 21).

Binder peptides displaying high RMSD and binding energy with respect to native RMSD value with HLA class I molecules mentioned in table 17, concludes that P2 and P3 peptide have shown high binding energy and RMSD values respectively with maximum number HLA class I molecules but P1 peptide have shown high RMSD value with minimum number of HLA class I alleles with respect to native RMSD values.

Binder peptides displaying high RMSD and binding energy with respect to native RMSD value with HLA class II molecules, mentioned in table 18, concludes that P2 peptide has shown high RMSD value with six HLA class II alleles whereas P1 peptide has shown high RMSD value

with minimum number of HLA class II alleles with respect to native RMSD values. P1, P2, P3 and P4 peptides have shown high binding energy with respect to native binding energies.

In majority of the cases the binding energy and RMSD values of peptides-HLA were found to be similar to corresponding native peptide-HLA representing the strong potential of these peptides to be presentated by HLA molecules to elicit T cel immune response.

Table 17: Peptide exhibiting higher binding energy and RMSD value with native peptide for HLA class I molecules

	Peptide	**RMSD**	B**inding energy**
		Alleles	
P1	LRKYCIEAK	A*1101, B*3508, A2	B27, A3, B58, B7, B62, A2, B8, A24
P2	WLVHRQWFLDLPLPW	A*1101, B*5703, A1, B*3508, B*1501, B8, A24	B27, A3, B58, B7, B62, A1,A2, B8, A24, B44
P3	ILGDTAWDFGSLGGVFTSIGKALH	A*1101, B*3508, B*1801, B*1501, B*2709, B*5703, A1, B8, A24	B27, A3, B58, B7, B62, B8, A24
P4	FSGVSWTMKILI	A1, B*3508, B8, A24, B*1801	A3, B7, B62, B8, A24
P5	LVLVGVVTLY	B*2709, A*1101, B*5708, B*3508, B8, A24, B*1501, B*1801	B7, B62, B8, A24

Table 18: Peptide exhibinting higher binding energy and RMSD value with native peptide for HLA class II molecules

	Peptide	**RMSD**	B**inding energy**
		Alleles	
P1	LRKYCIEAK	DP	DR4, DQ2, DR52c, DRB1*01:01, DQ6, DP
P2	WLVHRQWFLDLPLPW	DQ2, DR52c, DP, DR4, DRB1*04:01, DQ6	DR4, DQ2, DR52c, DRB1*01:01, DQ6, DP
P3	ILGDTAWDFGSLGGVFTSIGKALH	DR4, DR52c, DRB1*04:01, DP, DQ2	DR4, DQ2, DR52c, DRB1*01:01, DQ6, DP
P4	FSGVSWTMKILI	DR4, DP, DQ2, DR52c, DRB1*04:01	DR4, DQ2, DR52c, DRB1*01:01, DQ6, DP
P5	LVLVGVVTLY	DR4, DR52c,	DR52c, DQ6, DP

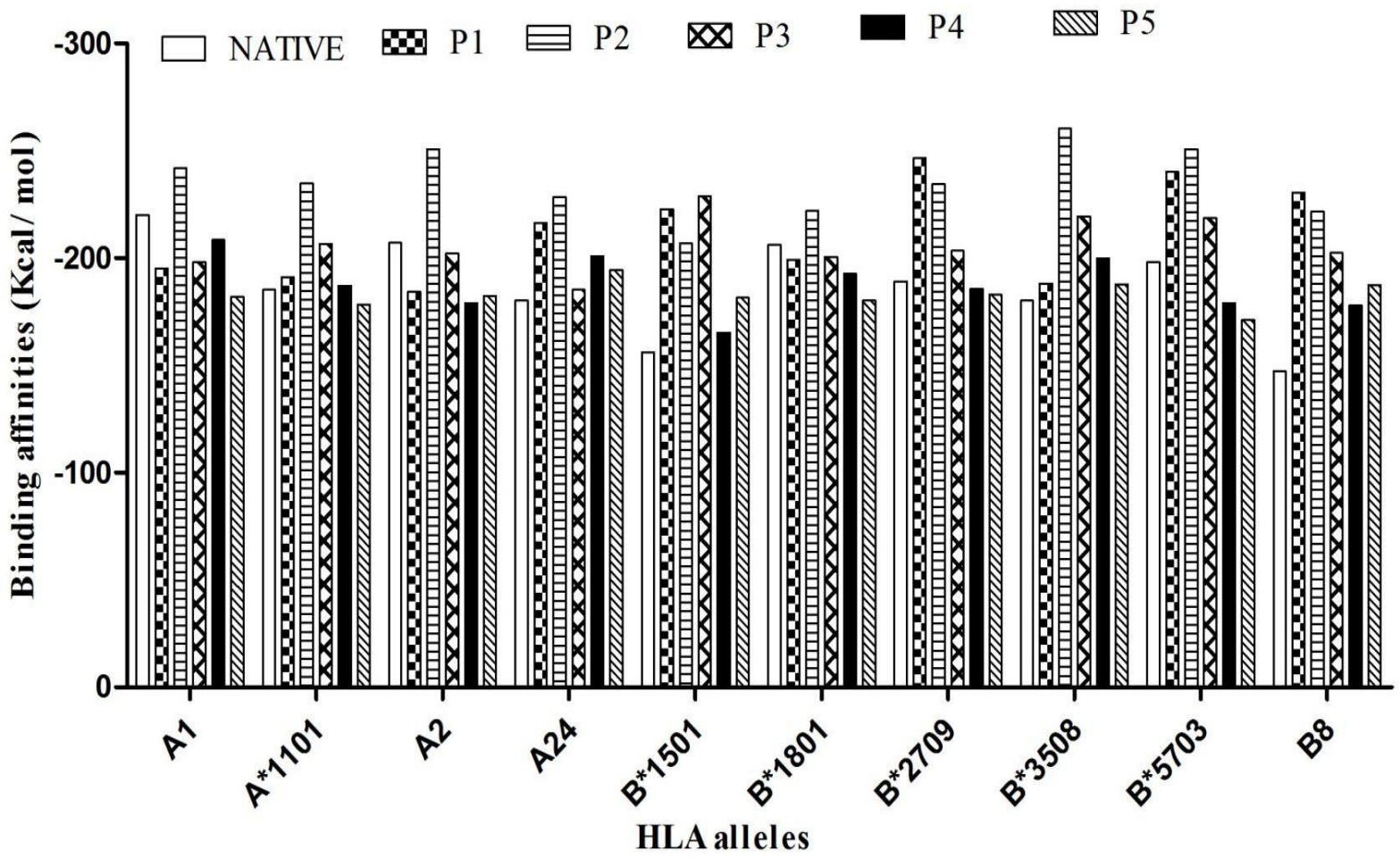


Figure 18: Binding energies of native and identified peptides with HLA class I molecules.

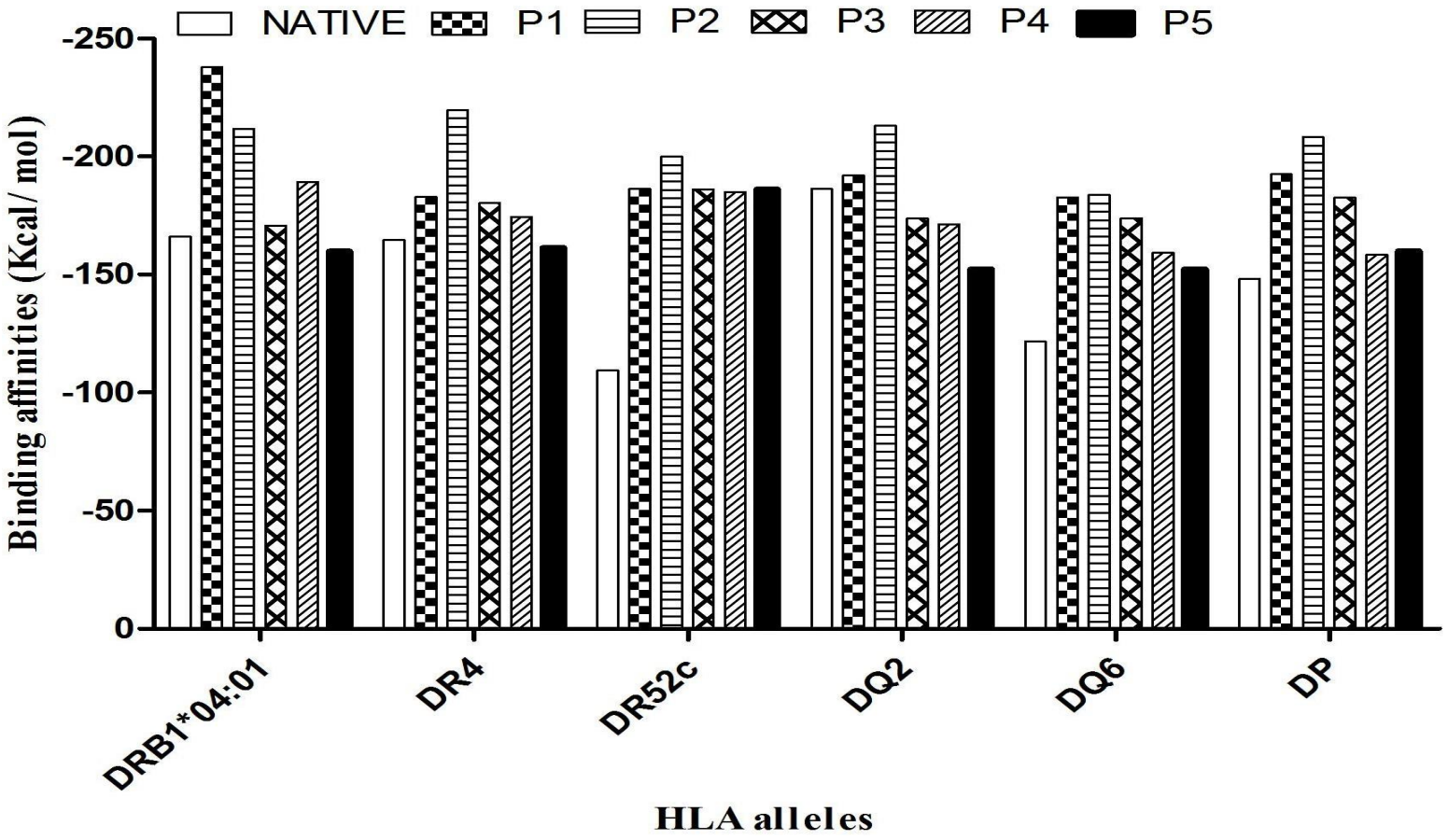


Figure 19: Binding energies of native and identified peptides with HLA class II molecules.

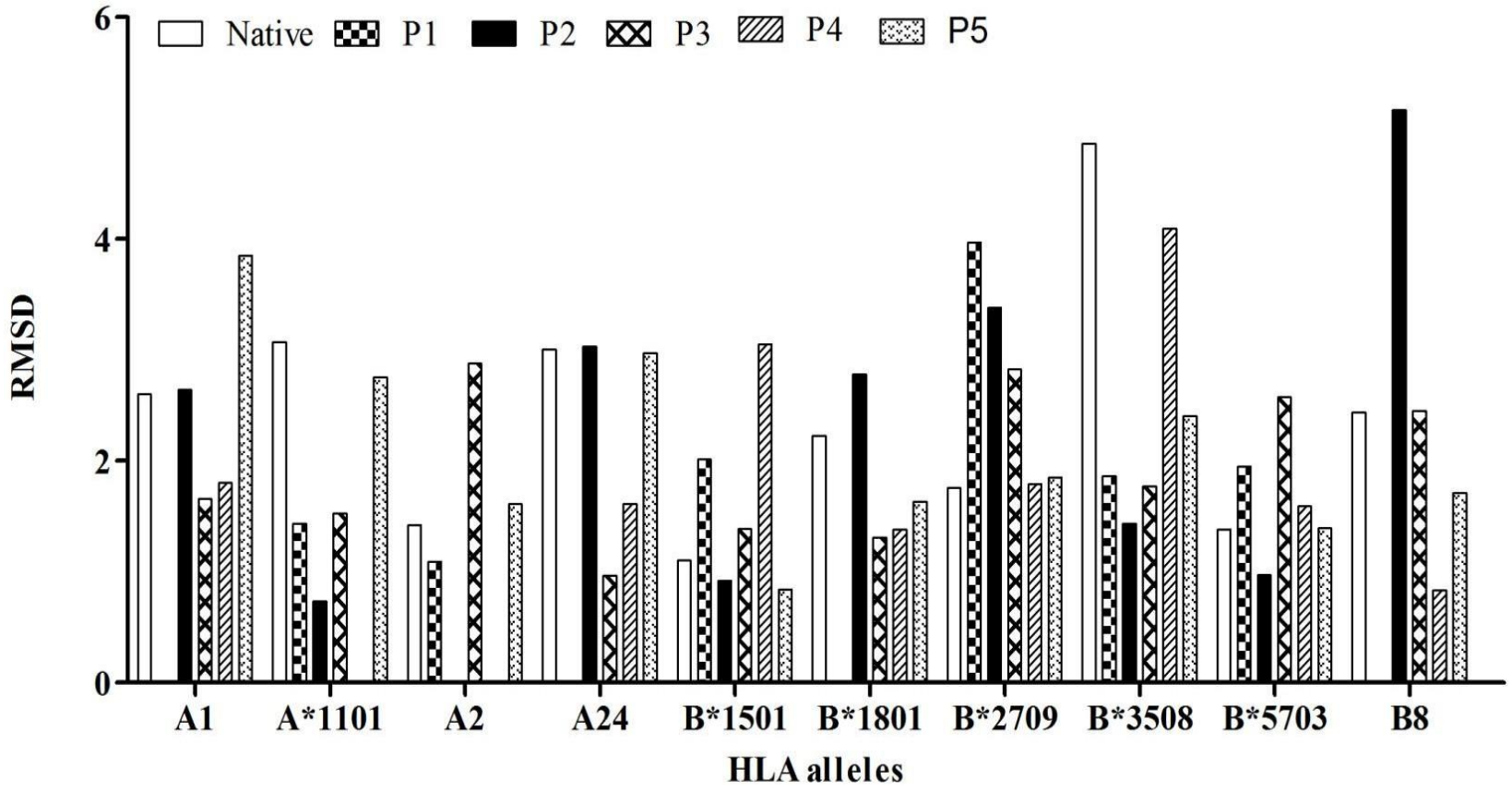


Figure 20: RMSD values of native and identified peptide with HLA class I molecules

.

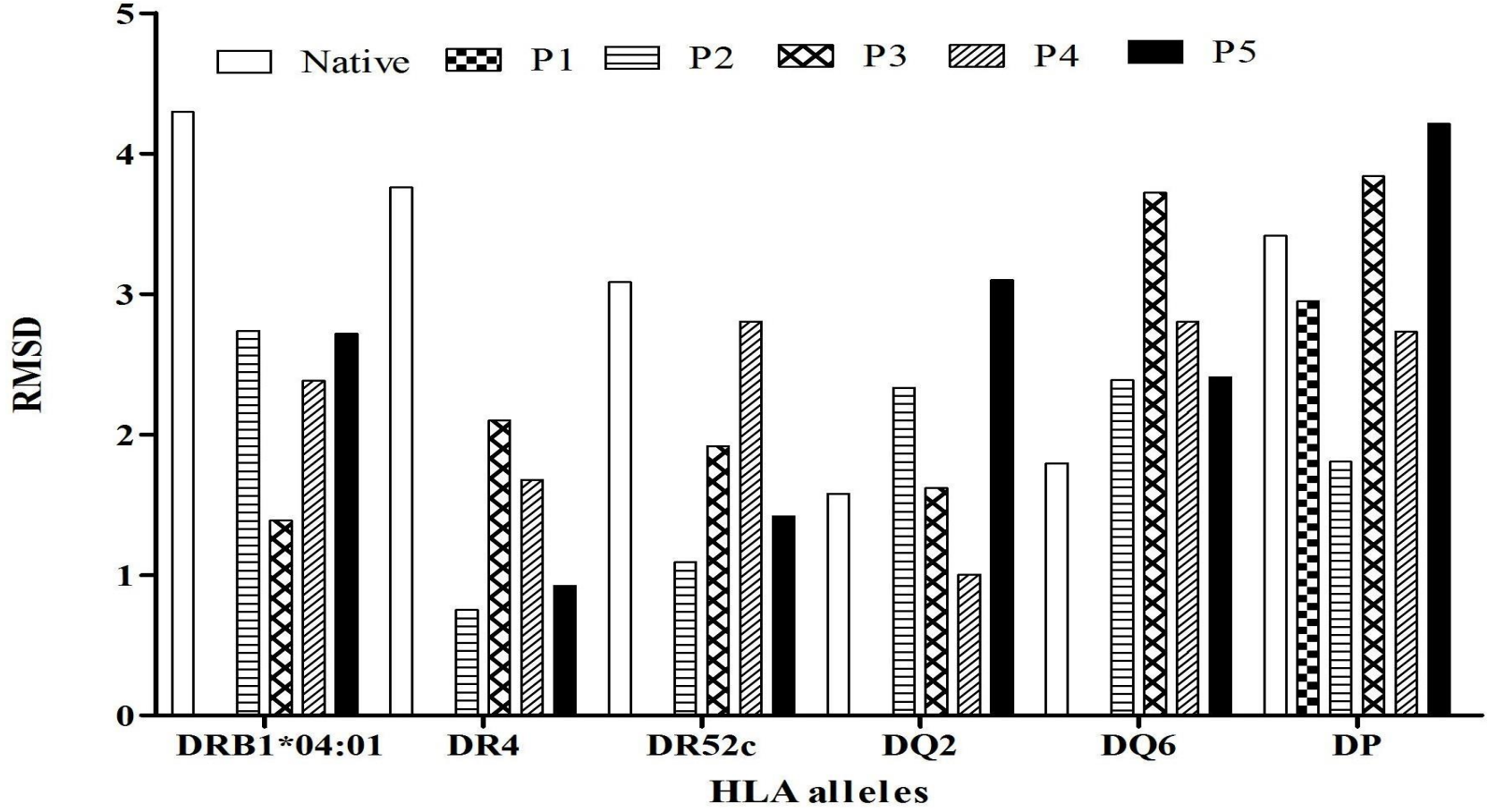


Figure 21: RMSD values of native and identified peptide with HLA class II molecules

5.8 Conservancy analysis

In order to circumvent the chances of antibody dependent enhancement upon infection with other dengue serotype or flavivirus, development of an effective anti-viral treatment entails employment of candidate which can elicit immune response against most of the members of *flavivirus*.

The conservancy of the identified DENV-2 E peptide fragments w.r.t. other dengue serotypes and flaviviruses viz., Zika virus (ZIKV), West Nile virus (WNV) and Japanese encephalitis virus (JEV) was studied for evaluating the percentage of amino acid homology.

The number of downloaded sequences of other DENV serotypes and flavivirus members and their time frame has been mentioned in table 19. Multiple sequence alignment of sequences of each species was carried out with the help of MUSCLE.

Table 19: Number of downloaded sequences and the time frame of flaviviruses

Virus	Total number of sequences	Number of unique sequences	Time frame
Dengue			
Serotype I	1732	463	1944 - APR 2017
Serotype III	815	228	1956 - JUNE 2016
Serotype IV	192	82	1956 - JAN 2016
Japanese Encephalitis virus	16	13	1952- MAR 2016
West Nile virus	140	57	1953-OCT 2017
Zika virus	516	114	1968-OCT 2018

Five peptides were found to be highly conserved in four different serotypes of Dengue, Zika virus, Japanese Encephalitis virus and West Nile virus (Table 20) except in some cases. Interestingly, P4 peptide was found to be 100% conserved in Japanese Encephalitis virus and West Nile virus and while some have peptides have shown ≥ 90 % conserved in different members of flavivirus (Table 20).

Table 20: Conservancy percentage of each peptide other dengue serotypes and flaviviruses.

	P1	**P2**	**P3**	**P4**	**P5**
DENV-2	92.40%	97.95%	92.28%	95.61%	71.05%
DENV-1	98.70%	99.57%	86.61%	97.40%	36.29%
DENV-3	99.12%	97.37%	96.05%	59.21%	59.65%
DENV-4	64.63%	97.56%	76.83%	65.85%	93.90%
ZIKV	97.25%	97.27%	96.04%	98.06%	95.19%
JEV	92.31%	61.54%	84.62%	100%	84.62%
WNV	36.84%	98.25%	38.60%	100%	96.49%

6. Discussion

Dengue virus (DENV) infection affects 3.9 million people per year worldwide. It occurs annually in the America, Asia and Australia. In India, an average incidence of DENV infection is > 50 per million populations. Among the four serotypes of DENV, DENV-2 is most infectious serotypes because it usually leads to secondary infection. Major challenge in the development of treatment for DENV infection is chances of antibody dependent enhancement upon infection with other DENV serotypes or other flaviviruses. Currently, only supportive treatment for DENV infection is available. A live attenuated vaccine, Dengvaxia® (CYD-TDV) has been licensed, but it is not effective in children younger than 9 years. Other potent vaccines are in various clinical trial stages.

Identification of potential binding sites using classical methods is a tedious and time consuming process but immunoinformatics approach reduces the experimental time. Limited studies have been done for the development of dengue vaccine using immunoinformatics approach. Class I epitopes and class II epitopes have been predicted by using NetCTL1.2 and IEDB database respectively against all the structural and non-structural proteins of DENV-2 (Ali et al., 2017). ProPred tool have been used for prediction of epitopes against the structural proteins of DENV (Somvanshi et al., 2009). In this study, the peptides displaying ≥70% conservancy were taken for prediction of T cell epitopes by using six prediction tools for enhancing the reliability of predicted epitopes.

Allergenicity analysis of predicted epitopes against all the structural and non-structural protein of dengue virus was carried out by AlgPred server (Ali et al., 2017). In this study, allergenicity, toxicity and blast analysis was done for screening of safe synthetic peptide vaccine candidates. None of predicted peptides were allergic and toxic but four of the predicted peptides exhibiting similarity with human proteome were eliminated for further analysis.

According to IEDB previous report analysis, among the identified DENV-2 E protein; P1,and P2 peptides have been reported by Simmons et al., 2005 and Rivino et al. 2013 respectively whereas both P4 and P5 have been reported by Ramanathan et al., 2016. ILGDTAWDF, ILGDTAWDFG, TAWDFGSLGGVFTSIG partial peptides of P3 peptide have been reported (Lund et al., 2011; Simmons et al., 2005; Innis et al., 1968).

IEDB population coverage tool was used for this analysis and evaluation was done for three continents (comprising of 13 geographical regions) viz., Asia, Africa and America. The population coverage in Europe, Oceania, West Indies and World was also analyzed. P1, P2. P3,

P4 and P5 peptides exhibits 51.94, 98.14, 97.82, 91.71 and 96.37 percent average population coverage respectively in the six continents. P1 and P2 peptide have displayed least and highest population coverage of 51.94 and 98.14 respectively.

P1 and P2 peptide located in the domain II (DII) of E protein between 52-131 amino acids, which participates in determination of serotype specificity., P3 and P4 peptide located in the stem segment between 392-449 amino acids, provides stability to three envelop protein domains viz., DI, DII and DIII and P5 peptide located in the transmembrane anchor region between 449-495 amino acids, facilitates interaction between three domains and stem segment.

Very few studies have reported HLA-Peptide docking using Molecular virtual docker tool. The binding affinities of peptides predicted against human coranavirus with HLA molecules have been evaluated by using Molecular Virtual Docker (MVD) tool (Oany et al., 2014). In this study, the binding affinities of HLA-peptide complexes was calculated by employing MVD tool which have shown high binding affinity of P2 peptide with all class I HLA molecules whereas P1, P3, P4 and P5 have shown high binding affinity with eight, seven, five and four class I HLA molecules respectively with respect to the binding affinities of native class I HLA molecules.

Out of five peptides, four peptides have shown high binding affinity with all the six HLA class II molecules whereas P5 peptides has shown high binding affinity with three out of six HLA class II molecules with respect to binding affinity of native class II HLA molecules

In this study, CABS-dock was used for calculating the RMSD values of HLA-peptide complexes for determining the quality of prediction. The RMSD value <3A°, between 3A° to 5.5A° and more than 5.5A° corresponds to high, medium and low quality prediction respectively and predicts the whether the predicted peptide is binder or non-binder with the HLA molecules.

Binder peptides displaying high RMSD and binding energy with respect to native RMSD value with HLA class I molecules, concludes that P2 and P3 peptide have shown high binding energy and RMSD values respectively with maximum number HLA class I molecules but P1 peptide have shown high RMSD value with minimum number of HLA class I alleles with respect to native RMSD values.

Binder peptides displaying high RMSD and binding energy with respect to native RMSD value with HLA class II molecules, concludes that P2 peptide has shown high RMSD value with six HLA class II alleles whereas P1 peptide has shown high RMSD value with minimum number of HLA class II alleles with respect to native RMSD values. P1, P2, P3 and P4 peptides have shown high binding energy with respect to native binding energies.

In majority of the cases the binding energy and RMSD values of peptides-HLA were found to be similar to corresponding native peptide-HLA representing the strong potential of these peptides to be presentated by HLA molecules to elicit T cel immune response.

Conservation analysis of these peptides showed that these peptides are also highly conserved in other three serotypes of Dengue, Zika virus, Japanese Encephalitis virus and West Nile virus representing that the potential of cross protective immunity in different member of flavivirus.

Conclusion

In conclusion, five peptides were identified which are enriched in $CD4^+$ and $CD8^+$ T cell epitopes and devoid of allergic, autoimmune and toxic response. Further, these peptides have shown strong binding interaction with HLA class I and II alleles and also represent good population coverage among different continents. Hence, these peptides need to be validated in different experimental model to confirm their immunogenic potential to consider for vaccine candidate against Dengue virus.

References

1. Adhikari, U. K., Tayebi, M., & Rahman, M. M. (2018). Immunoinformatics Approach for Epitope-Based Peptide Vaccine Design and Active Site Prediction against Polyprotein of Emerging Oropouche Virus. *Journal of immunology research*, *2018*.

2. Ali, M., Pandey, R. K., Khatoon, N., Narula, A., Mishra, A., & Prajapati, V. K. (2017). Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Scientific reports*, *7*(1), 9232.

3. Andreatta, M., & Nielsen, M. (2018). Bioinformatics tools for the prediction of T-cell epitopes. In *Epitope Mapping Protocols*(pp. 269-281). Humana Press, New York, NY.

4. Aruna, R,. Review on Dengue viral Replication, assembly and entry into the host cells. *Int.J.Curr.Microbiol.App.Sci* (2014) *3*(11): 1025-1039.

5. Alcalá, A. C., Palomares, L. A., & Ludert, J. E. (2018). Secretion of nonstructural protein 1 of dengue virus from infected mosquito cells: facts and speculations. *Journal of virology*, *92*(14), e00275-18.

6. Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... & Myers, M. F. (2013). The global distribution and burden of dengue. *Nature*, *496*(7446), 504.

7. Backert, L., & Kohlbacher, O. (2015). Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome medicine*, *7*(1), 119.

8. Bhasin, M., & Raghava, G. P. S. (2004). SVM based method for predicting HLA-DRB1* 0401 binding peptides in an antigen sequence. *Bioinformatics*, *20*(3), 421-423.

9. Bhuvanakantham, R., Li, J., Tan, T. T., and Ng, M. L. (2010). Human Sec3 protein is a novel transcriptional and translational repressor of flavivirus. Cell. Microbiol. 12:453–472.

10. Blacksell, S. D. (2012). Commercial dengue rapid diagnostic tests for point-of-care application: recent evaluations and future needs?. *BioMed Research International*, *2012*.

11. Blaszczyk, M., Kurcinski, M., Kouza, M., Wieteska, L., Debinski, A., Kolinski, A., & Kmiecik, S. (2016). Modeling of protein–peptide interactions using the CABS-dock web server for binding site search and flexible docking. *Methods*, *93*, 72-83.

12. Bollati, M., Alvarez, K., Assenberg, R., Baronti, C., Canard, B., Cook, S., ... & Grard, G. (2010). Structure and functionality in flavivirus NS-proteins: perspectives for drug design. *Antiviral research*, *87*(2), 125-148.

13. Chakraborty, S., Chakravorty, R., Ahmed, M., Rahman, A., Waise, T. M., Hassan, F., & Shamsuzzaman, S. (2010). A computational approach for identification of epitopes in dengue

virus envelope protein: a step towards designing a universal dengue vaccine targeting endemic regions. *In silico biology*, *10*(5, 6), 235-246.

14. Chen, S., Wu, Z., Wang, M., & Cheng, A. (2017). Innate immune evasion mediated by flaviviridae non-structural proteins. *Viruses*, *9*(10), 291.

15. Christofferson, R. C. (2015). Editor's chice: A Reevaluation of the Role of Aedes albopictus in Dengue Transmission. *The Journal of infectious diseases*, *212*(8), 1177.

16. Collins, M. H., & Metz, S. W. (2017). Progress and works in progress: update on flavivirus vaccine development. *Clinical therapeutics*, *39*(8), 1519-1536.

17. Cruz-Oliveira, C., Freire, J. M., Conceição, T. M., Higa, L. M., Castanho, M. A., & Da Poian, A. T. (2015). Receptors and routes of dengue virus entry into the host cells. *FEMS microbiology reviews*, *39*(2), 155-170.

18. Daep, C. A., Muñoz-Jordán, J. L., & Eugenin, E. A. (2014). Flaviviruses, an expanding threat in public health: focus on dengue, West Nile, and Japanese encephalitis virus. *Journal of neurovirology*, *20*(6), 539-560.

19. De Brito, R. C., Jamille, M. D. O., Reis, L. E., Vieira, J. F., Mathias, F. A., Roatt, B. M., ... & Reis, A. B. (2018). Peptide vaccines for leishmaniasis. *Frontiers in immunology*, *9*.

20. De La Guardia, C., & Lleonart, R. (2014). Progress in the identification of dengue virus entry/fusion inhibitors. *BioMed research international*, *2014*.

21. Dhanda, S. K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M. C., ... & Sette, A. (2019). IEDB-AR: immune epitope database—analysis resource in 2019. *Nucleic acids research*.

22. Diaz, C., Lin, L., Martinez, L. J., Eckels, K. H., Campos, M., Jarman, R. G., ... & Innis, B. L. (2018). Phase I randomized study of a tetravalent dengue purified inactivated vaccine in healthy adults from Puerto Rico. *The American journal of tropical medicine and hygiene*, *98*(5), 1435-1443.

23. Dick, O. B., San Martín, J. L., Montoya, R. H., del Diego, J., Zambrano, B., & Dayan, G. H. (2012). The history of dengue outbreaks in the Americas. *The American journal of tropical medicine and hygiene*, *87*(4), 584-593.

24. Diamond, M. S., & Pierson, T. C. (2015). Molecular insight into dengue virus pathogenesis and its implications for disease control. *Cell*, *162*(3), 488-492.

25. Dos Santos Francisco, R., Buhler, S., Nunes, J. M., Bitarello, B. D., França, G. S., Meyer, D., & Sanchez-Mazas, A. (2015). HLA supertype variation across populations: new insights into the role of natural selection in the evolution of HLA-A and HLA-B polymorphisms. *Immunogenetics*, *67*(11-12), 651-663.

26. Duan, X., Lu, X., Li, J., & Liu, Y. (2008). Novel binding between pre-membrane protein and vacuolar ATPase is required for efficient dengue virus secretion. *Biochemical and biophysical research communications*, *373*(2), 319-324.

27. Dutra, N. R., de Paula, M. B., de Oliveira, M. D., de Oliveira, L. L., & De Paula, S. O. (2009). The laboratorial diagnosis of dengue: applications and implications. *Journal of global infectious diseases*, *1*(1), 38.

28. Edeling, M. A., Diamond, M. S., & Fremont, D. H. (2014). Structural basis of Flavivirus NS1 assembly and antibody recognition. *Proceedings of the National Academy of Sciences*, *111*(11), 4285-4290.

29. Edinur, H. A., Manaf, S. M., & Mat, N. F. C. (2016). Genetic barriers in transplantation medicine. *World journal of transplantation*, *6*(3), 532.

30. El Sahili, A., and J. Lescar. "Dengue virus non-structural protein 5. Viruses 9: 91." (2017).

31. Fahimi, H., Mohammadipour, M., Kashani, H. H., Parvini, F., & Sadeghizadeh, M. (2018). Dengue viruses and promising envelope protein domain III-based vaccines. *Applied microbiology and biotechnology*, *102*(7), 2977-2996.

32. Ferreira, L., dos Santos, R., Oliva, G., & Andricopulo, A. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, *20*(7), 13384-13421.

33. Fibriansah, G., Ng, T. S., Kostyuchenko, V. A., Lee, J., Lee, S., Wang, J., & Lok, S. M. (2013). Structural changes in dengue virus when exposed to a temperature of 37 C. *Journal of virology*, *87*(13), 7585-7592.

34. Fleri, W., Paul, S., Dhanda, S. K., Mahajan, S., Xu, X., Peters, B., & Sette, A. (2017). The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. *Frontiers in immunology*, *8*, 278.

35. Fried, J. R., Gibbons, R. V., Kalayanarooj, S., Thomas, S. J., Srikiatkhachorn, A., Yoon, I. K., & Cummings, D. A. (2010). Serotype-specific differences in the risk of dengue hemorrhagic fever: an analysis of data collected in Bangkok, Thailand from 1994 to 2006. *PLoS neglected tropical diseases*, *4*(3), e617.

36. Fernandez-Garcia, M. D., Mazzon, M., Jacobs, M., & Amara, A. (2009). Pathogenesis of flavivirus infections: using and abusing the host cell. *Cell host & microbe*, *5*(4), 318-328.

37. Ghosh, A., & Dar, L. (2015). Dengue vaccines: challenges, development, current status and prospects. *Indian journal of medical microbiology*, *33*(1), 3.

38. Gould, E. A., & Solomon, T. (2008). Pathogenic flaviviruses. *The Lancet*, *371*(9611), 500-509.

39. Govindarajan, D., Meschino, S., Guan, L., Clements, D. E., ter Meulen, J. H., Casimiro, D. R., ... & Bett, A. J. (2015). Preclinical development of a dengue tetravalent recombinant subunit vaccine: immunogenicity and protective efficacy in nonhuman primates. *Vaccine*, *33*(33), 4105-4116.

40. Gubler, D. J. (2006, August). Dengue/dengue haemorrhagic fever: history and current status. In *New Treatment Strategies for Dengue and Other Flaviviral Diseases: Novartis Foundation Symposium 277* (pp. 3-22). Chichester, UK: John Wiley & Sons, Ltd.

41. Gupta, N., Srivastava, S., Jain, A., & Chaturvedi, U. (2012). Dengue in India. *Indian Journal of Medical Research*, *136*(3), 373.

42. Guo, C., Zhou, Z., Wen, Z., Liu, Y., Zeng, C., Xiao, D., ... & Ye, X. (2017). Global epidemiology of dengue outbreaks in 1990–2015: a systematic review and meta-analysis. *Frontiers in cellular and infection microbiology*, *7*, 317.

43. Guy, B., & Jackson, N. (2016). Dengue vaccine: hypotheses to understand CYD-TDV-induced protection. *Nature reviews Microbiology*, *14*(1), 45.

44. Gromowski, G. D., Henein, S., Kannadka, C. B., Barvir, D. A., Thomas, S. J., de Silva, A. M., & Jarman, R. G. (2018). Delineating the serotype-specific neutralizing antibody response to a live attenuated tetravalent dengue vaccine. *Vaccine*, *36*(18), 2403-2410.

45. Halenius, A., Gerke, C., & Hengel, H. (2015). Classical and non-classical MHC I molecule manipulation by human cytomegalovirus: so many targets—but how many arrows in the quiver?. *Cellular & molecular immunology*, *12*(2), 139.

46. Huang, X., Karabudak, A., Comber, J. D., Philip, M., Morcol, T., & Philip, R. (2017). A novel immunization approach for dengue infection based on conserved T cell epitopes formulated in calcium phosphate nanoparticles. *Human vaccines & immunotherapeutics*, *13*(11), 2612-2625.

47. Hsieh, S. C., Tsai, W. Y., and Wang, W. K. (2010). The length of and nonhydrophobic residues in the transmembrane domain of dengue virus envelope protein are critical for its retention and assembly in the endoplasmic reticulum. J. Virol. 84:4782–4797.

48. Neglected tropical disease http://www.searo.who.int/entity/vector_borne_tropical_ diseases/data/data_factsheet/en/)

49. Symptoms and Treatment https://www.cdc.gov/dengue/symptoms/index.html

50. Flaviviridae https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/flaviviridae

51. Dengue and severe dengue https://www.who.int/news-room/fact sheets/detail/dengue-and-severe-dengue

52. Communicable disease threats report, https://ecdc.europa.eu/sites/portal/files /documents/CDTR_week47-2018.pdf

53. Dengue situation update number 568, https://iris.wpro.who.int/bitstream/ handle/10665.1/14329/Dengue-20190509.pdf

54. Immunization, Vaccines and Biologicals https://www.who.int/immunization/ research/dev elopment/dengue_vaccines/en/

55. Symptoms and treatment https://www.cdc.gov/dengue/symptoms/index.html

56. IPD-IMGT/HLA https://www.ebi.ac.uk/ipd/imgt/hla/stats.html

57. MUSCLE Help and documentation https://www.ebi.ac.uk/seqdb/ confluence display/JDSAT/MUSCLE+Help+and+Documentation

58. http://www.syfpeithi.de/bin/MHCServer.dll/Info.html

59. NetCTL 1.2 Server http://www.cbs.dtu.dk/services/NetCTL/

60. http://tools.iedb.org/mhci/help/

61. MHC2pred help http://crdd.osdd.net/raghava/mhc2pred/info.html

62. Introduction to server 1.0 https://webs.iiitd.edu.in/raghava/propred/page3.html

63. AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes https://webs.iiitd.edu.in/raghava/algpred/help.html

64. http://crdd.osdd.net/raghava/toxinpred/

65. PEP-FOLD server ,http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/

66.Innis, B. L., Thirawuth, V., & Hemachudha, C. (1989). Identification of continuous epitopes of the envelope glycoprotein of dengue type 2 virus. *The American journal of tropical medicine and hygiene*, *40*(6), 676-687.

67. Jain, S., & Baranwal, M. (2019). Computational analysis in designing T cell epitopes enriched peptides of Ebola glycoprotein exhibiting strong binding interaction with HLA molecules. *Journal of theoretical biology*, *465*, 34-44.

68. Jespersen, M. C., Mahajan, S., Peters, B., Nielsen, M., & Marcatili, P. (2019). Antibody specific B-cell epitope predictions: leveraging information from antibody-antigen protein complexes. *Frontiers in immunology*, *10*.

69. Junglen, S., Korries, M., Grasse, W., Wieseler, J., Kopp, A., Hermanns, K., ... & Kümmerer, B. M. (2017). Host range restriction of insect-specific flaviviruses occurs at several levels of the viral life cycle. *Msphere*, *2*(1), e00375-16.

70. Kao, Y. T., Lai, M., & Yu, C. Y. (2018). How dengue virus circumvents innate immunity. *Frontiers in immunology*, *9*, 2860.

71. Klein, D. E., Choi, J. L., & Harrison, S. C. (2013). Structure of a dengue virus envelope protein late-stage fusion intermediate. *Journal of virology*, *87*(4), 2287-2293.

72. Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A., & Kmiecik, S. (2015). CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. *Nucleic acids research*, *43*(W1), W419-W424.

73. Kusumaningrum, S., Budianto, E., Kosela, S., Sumaryono, W., & Juniarti, F. (2014). The molecular docking of 1, 4-naphthoquinone derivatives as inhibitors of polo-like kinase 1 using Molegro Virtual Docker. *J App Sci*, *4*, 47-53.

74. Kroschewski, H., Sagripanti, J. L., & Davidson, A. D. (2009). Identification of amino acids in the dengue virus type 2 envelope glycoprotein critical to virus infectivity. *Journal of General Virology*, *90*(10), 2457-2461.

75. Laureti, M., Narayanan, D., Rodriguez-Andres, J., Fazakerley, J. K., & Kedzierski, L. (2018). Flavivirus receptors: diversity, identity, and cell entry. *Frontiers in immunology*, *9*.

76. Larsen, M. V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., & Nielsen, M. (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC bioinformatics*, *8*(1), 424.

77. Lata, S., Bhasin, M., & Raghava, G. P. (2007). Application of machine learning techniques in predicting MHC binders. In *Immunoinformatics* (pp. 201-215). Humana Press.

78. Lee, E., Leang, S. K., Davidson, A., and Lobigs, M. (2010). Both E protein glycans adversely affect dengue virus infectivity but are beneficial for virion release. J. Virol. 84:5171–5180.

79. Li, W., Joshi, M., Singhania, S., Ramsey, K., & Murthy, A. (2014). Peptide vaccine: progress and challenges. *Vaccines*, *2*(3), 515-536.

80. Limjindaporn, T., Wongwiwat, W., Noisakran, S., Srisawat, C., Netsawang, J., Puttikhunt, C., ... & Sittisombut, N. (2009). Interaction of dengue virus envelope protein with endoplasmic reticulum-resident chaperones facilitates dengue virus production. *Biochemical and biophysical research communications*, *379*(2), 196-200.

81. Liu, Y., Zhou, J., Yu, Z., Fang, D., Fu, C., Zhu, X., ... & Jiang, L. (2014). Tetravalent recombinant dengue virus-like particles as potential vaccine candidates: immunological properties. *BMC microbiology*, *14*(1), 233.

82. Liu, M. A. (2010). Immunologic basis of vaccine vectors. *Immunity*, *33*(4), 504-515.

83. Lohia, N., & Baranwal, M. (2015). Identification of conserved peptides comprising multiple T cell epitopes of Matrix 1 protein in H1N1 influenza virus. *Viral immunology*, *28*(10), 570-579.

84. Lundegaard, C., Lund, O., Buus, S., & Nielsen, M. (2010). Major histocompatibility complex class I binding predictions as a tool in epitope discovery. *Immunology*, *130*(3), 309-318.

85. Lund, O., Nascimento, E. J., Maciel Jr, M., Nielsen, M., Larsen, M. V., Lundegaard, C., ... & August, T. J. (2011). Human leukocyte antigen (HLA) class I restricted epitope discovery in yellow fewer and dengue viruses: importance of HLA binding strength. *PloS one*, *6*(10), e26494.

86. Madden, T. (2013). The BLAST sequence analysis tool. In *The NCBI Handbook [Internet]. 2nd edition*. National Center for Biotechnology Information (US).

87. Marintcheva, B. (2018). *Viruses as Tools for Vaccine Development. Harnessing the Power of Viruses, 217–242.* doi:10.1016/b978-0-12-810514-6.00008-8.

88. Medigeshi, G. R. (2011). Mosquito-borne flaviviruses: overview of viral life-cycle and host–virus interactions. *Future Virology*, *6*(9), 1075-1089.

89. Messer, W. B., De Alwis, R., Yount, B. L., Royal, S. R., Huynh, J. P., Smith, S. A., ... & White, L. J. (2015). Retraction: Dengue virus envelope protein domain I/II hinge determines long-lived serotype-specific dengue immunity (Proc Natl Acad Sci USA (2014) 111,(1939-1944. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(20), E2738.

90. Messina, J. P., Brady, O. J., Scott, T. W., Zou, C., Pigott, D. M., Duda, K. A., ... & Simmons, C. P. (2014). Global spread of dengue virus types: mapping the 70 year history. *Trends in microbiology*, *22*(3), 138-146.

91. Miotto, O., Heiny, A. T., Tan, T. W., August, J. T., & Brusic, V. (2008, February). Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large-scale mutual information analysis. In *BMC bioinformatics* (Vol. 9, No. 1, p. S18). BioMed Central.

92. Monos, D. S., & Winchester, R. J. (2019). The Major Histocompatibility Complex. In *Clinical Immunology* (pp. 79-92). Content Repository Only!.

93. Mukhopadhyay, S., Kuhn, R. J., & Rossmann, M. G. (2005). A structural perspective of the flavivirus life cycle. *Nature Reviews Microbiology*, *3*(1), 13.

94. Muller, D. A., Depelsenaire, A. C., & Young, P. R. (2017). Clinical and laboratory diagnosis of dengue virus infection. *The Journal of infectious diseases*, *215*(suppl_2), S89-S95.

95. Mutheneni, S. R., Morse, A. P., Caminade, C., & Upadhyayula, S. M. (2017). Dengue burden in India: recent trends and importance of climatic parameters. *Emerging microbes & infections*, *6*(1), 1-10.

96. Nandy, A., Dey, S., Nandy, P., Das, S., & Nandy, A. (2016). Comparative Study of Envelope Proteins of Dengue Virus of All Four Serotypes Isolated In India. *Journal of Bioinformatics, Proteomics and Imaging Analysis*, *2*(2), 0-0.

97. Nadugala, M. N., Premaratne, P. H., & Goonasekara, C. L. (2016). Systematic bioinformatic approach for prediction of linear B-cell epitopes on dengue E and prM protein. *Advances in bioinformatics*, *2016*.

98. Nemésio, H., Palomares-Jerez, F., & Villalaín, J. (2012). NS4A and NS4B proteins from dengue virus: membranotropic regions. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1818*(11), 2818-2830.

99.Nielsen, M., & Lund, O. (2009). NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC bioinformatics*, *10*(1), 296.

100. Nemésio, H., Palomares-Jerez, F., & Villalaín, J. (2011). The membrane-active regions of the dengue virus proteins C and E. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1808*(10), 2390-2402.

101. Oany, A. R., Emran, A. A., & Jyoti, T. P. (2014). Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. *Drug design, development and therapy*, *8*, 1139.

102. Omokoko, M. D., Pambudi, S., Phanthanawiboon, S., Masrinoul, P., Setthapramote, C., Sasaki, T., ... & Ikuta, K. (2014). A highly conserved region between amino acids 221 and 266 of dengue virus non-structural protein 1 is a major epitope region in infected patients. *The American journal of tropical medicine and hygiene*, *91*(1), 146-155.

103. Pang, X., Zhang, M., & Dayton, A. I. (2001). Development of dengue virus replicons expressing HIV-1 gp120 and other heterologous genes: a potential future tool for dual vaccination against dengue virus and HIV. *BMC microbiology*, *1*(1), 28.

104. Perera, R., & Kuhn, R. J. (2008). Structural proteomics of dengue virus. *Current opinion in microbiology*, *11*(4), 369-377.

105. Perera-Lecoin, M., Meertens, L., Carnec, X., & Amara, A. (2014). Flavivirus entry receptors: an update. *Viruses*, *6*(1), 69-88.

106. Petersdorf, E., & Socié, G. (2019). The HLA System in Hematopoietic Stem Cell Transplantation. In *Immune Biology of Allogeneic Hematopoietic Stem Cell Transplantation* (pp. 15-32). Academic Press.

107. Posch, P. E., & Hurley, C. K. (2011). Histocompatibility: HLA and other systems. In *Blood and Bone Marrow Pathology* (pp. 641-676). Churchill Livingstone.

108. Potocnakova, L., Bhide, M., & Pulzova, L. B. (2016). An introduction to B-cell epitope mapping and in silico epitope prediction. *Journal of immunology research*, *2016*.

109. Pratheek, B. M., Nayak, T. K., Sahoo, S. S., Mohanty, P. K., Chattopadhyay, S., Chakraborty, N. G., & Chattopadhyay, S. (2014). Mammalian non-classical major histocompatibility complex I and its receptors: important contexts of gene, evolution, and immunity. *Indian journal of human genetics*, *20*(2), 129.

110. Rajapakse, S., Rodrigo, C., & Rajapakse, A. (2012). Treatment of dengue fever. *Infection and drug resistance*, *5*, 103.

111. Rammensee, H. G., Bachmann, J., Emmerich, N. P. N., Bachor, O. A., & Stevanović, S. S. Y. F. P. E. I. T. H. I. (1999). SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, *50*(3-4), 213-219.

112. Ramanathan, B., Poh, C. L., Kirk, K., McBride, W. J. H., Aaskov, J., & Grollo, L. (2016). Synthetic B-cell epitopes eliciting cross-neutralizing antibodies: strategies for future dengue vaccine. *PloS one*, *11*(5), e0155900.

113. Ranjbar,M.M., Dashpour, N., Ebrahimi, M.M.(2019). Novel applications of Immuno-bioinformatics in vaccine and bio-product developments at research institutes. *Archives of Razi Institute*, *Vol*

114. Rastogi, M., Sharma, N., & Singh, S. K. (2016). Flavivirus NS1: a multifaceted enigmatic viral protein. *Virology journal*, *13*(1), 131.

115. Reginald, K., Chan, Y., Plebanski, M., & Poh, C. L. (2018). Development of peptide vaccines in dengue. *Current pharmaceutical design*, *24*(11), 1157-1173.

116. Rico-Hesse, R. (2003). Microevolution and virulence of dengue viruses. *Advances in virus research*, *59*, 315.

117. Rivino, L., Kumaran, E. A., Jovanovic, V., Nadua, K., Teo, E. W., Pang, S. W., ... & Hanson, B. J. (2013). Differential targeting of viral components by CD4+ versus CD8+ T lymphocytes in dengue virus infection. *Journal of virology*, *87*(5), 2693-2706.

118. Rey, F. A., Stiasny, K., Vaney, M. C., Dellarole, M., & Heinz, F. X. (2018). The bright and the dark side of human antibody responses to flaviviruses: lessons for vaccine design. *EMBO reports*, *19*(2), 206-224.

119. Rocha, R. P., Livonesi, M. C., Fumagalli, M. J., Rodrigues, N. F., da Costa, L. C. F., dos Santos, M. C. S. G., ... & Coelho, L. F. L. (2014). Evaluation of tetravalent and conserved synthetic peptides vaccines derived from Dengue virus Envelope domain I and II. *Virus research*, *188*, 122-127.

120. Rodriguez-Roche, R., Blanc, H., Bordería, A. V., Díaz, G., Henningsson, R., Gonzalez, D., ... & Vignuzzi, M. (2016). Increasing clinical severity during a dengue virus type 3 Cuban epidemic: deep sequencing of evolving viral populations. *Journal of virology*, *90*(9), 4320-4333

121. Salles, T. S., da Encarnação Sá-Guimarães, T., de Alvarenga, E. S. L., Guimarães-Ribeiro, V., de Meneses, M. D. F., de Castro-Salles, P. F., ... & Moreira, M. F. (2018). History, epidemiology and diagnostics of dengue in the American and Brazilian contexts: a review. *Parasites & vectors*, *11*(1), 264.

122. Saha, S., & Raghava, G. P. S. (2006). AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research*, *34*(suppl_2), W202-W209.

123. Sasazuki, T., Inoko, H., Morishima, S., & Morishima, Y. (2016). Gene map of the HLA region, Graves' disease and Hashimoto thyroiditis, and hematopoietic stem cell transplantation. In *Advances in immunology* (Vol. 129, pp. 175-249). Academic Press.

124. Sessions, O. M., Wilm, A., Kamaraj, U. S., Choy, M. M., Chow, A., Chong, Y., & Ooi, E. E. (2015). Analysis of dengue virus genetic diversity during human and mosquito infection reveals genetic constraints. *PLoS neglected tropical diseases*, *9*(9), e0004044.

125. Shim, E. (2019). Optimal dengue vaccination strategies of seropositive individuals. *Mathematical Biosciences and Engineering*, *16*(3), 1171-1189

126. Shrivastava, S., Tiraki, D., Diwan, A., Lalwani, S. K., Modak, M., Mishra, A. C., & Arankalle, V. A. (2018). Co-circulation of all the four dengue virus serotypes and detection of a novel clade of DENV-4 (genotype I) virus in Pune, India during 2016 season. *PloS one*, *13*(2), e0192672.

127. Simmonds, P., Becher, P., Bukh, J., Gould, E. A., Meyers, G., Monath, T., ... & Stapleton, J. T. (2017). ICTV virus taxonomy profile: Flaviviridae. *The Journal of general virology*, *98*(1), 2.

128. Simmons, C. P., Dong, T., Chau, N. V., Dung, N. T. P., Chau, T. N. B., Dung, N. T., ... & Farrar, J. (2005). Early T-cell responses to dengue virus epitopes in Vietnamese adults with secondary dengue virus infections. *Journal of virology*, *79*(9), 5665-5675.

129. Singh, H., & Raghava, G. P. S. (2001). ProPred: prediction of HLA-DR binding sites. *Bioinformatics*, *17*(12), 1236-1237.

130. Skwarczynski, M., & Toth, I. (2016). Peptide-based synthetic vaccines. *Chemical Science*, *7*(2), 842-854.

131. Snyder, J. E., Jose, J., & Kuhn, R. J. (2014). The Togaviridae and the Flaviviridae.

132. Stern, L. J., & Calvo-Calle, J. M. (2009). HLA-DR: molecular insights and vaccine design. *Current pharmaceutical design*, *15*(28), 3249-3261.

133. Swaminathan, S., & Khanna, N. (2019). Dengue vaccine development: Global and Indian scenarios. *International Journal of Infectious Diseases*.

134. Thomsen, R., & Christensen, M. H. (2006). MolDock: a new technique for high-accuracy molecular docking. *Journal of medicinal chemistry*, *49*(11), 3315-3321.

135. Tomar, N., & De, R. K. (2010). Immunoinformatics: an integrated scenario. *Immunology131*(2), 153-168.

136. Torresi, J., Ebert, G., & Pellegrini, M. (2017). Vaccines licensed and in clinical trials for the prevention of dengue. *Human vaccines & immunotherapeutics*, *13*(5), 1059-1072.

137. Tuiskunen Bäck, A., & Lundkvist, Å. (2013). Dengue viruses–an overview. *Infection ecology & epidemiology*, *3*(1), 19839.

138. Uno, N., & Ross, T. M. (2018). Dengue virus and the host innate immune response. *Emerging microbes & infections*, *7*(1), 1-11.

139. Urcuqui-Inchima, S., Patino, C., Torres, S., Haenni, A. L., & Díaz, F. J. (2010). Recent developments in understanding dengue virus replication. In *Advances in virus research* (Vol. 77, pp. 1-39). Academic Press.

140. Wahala, W. M., & de Silva, A. M. (2011). The human antibody response to dengue virus infection. *Viruses*, *3*(12), 2374-2395.

141. Watterson, D., Kobe, B., & Young, P. R. (2012). Residues in domain III of the dengue virus envelope glycoprotein involved in cell-surface glycosaminoglycan binding. *Journal of General Virology*, *93*(1), 72-82.

142. Wilder-Smith, A., Vannice, K., Durbin, A., Hombach, J., Thomas, S. J., Thevarjan, I., & Simmons, C. P. (2018). Zika vaccines and therapeutics: landscape analysis and challenges ahead. *BMC medicine*, *16*(1), 84.

143. Xie, X., Gayen, S., Kang, C., Yuan, Z., & Shi, P. Y. (2013). Membrane topology and function of dengue virus NS2A protein. *Journal of virology*, *87*(8), 4609-4622.

144. Yan, Q. (2010). Immunoinformatics and systems biology methods for personalized medicine. In *Systems Biology in Drug Discovery and Development* (pp. 203-220). Humana Press, Totowa, NJ.

145. Yang, H., & Kim, D. S. (2015). Peptide immunotherapy in vaccine development: from epitope to adjuvant. In *Advances in protein chemistry and structural biology* (Vol. 99, pp. 1-14). Academic Press.

146. Zeidler, J., Fernandes-Siqueira, L., Barbosa, G., & Da Poian, A. (2017). Non-canonical roles of Dengue virus non-structural proteins. *Viruses*, *9*(3), 42.

Thesis-1

ORIGINALITY REPORT

16%	8%	8%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to National University of Singapore Student Paper	2%
2	www.frontiersin.org Internet Source	1%
3	Carlo Amorin Daep, Jorge L. Muñoz-Jordán, Eliseo Alberto Eugenin. "Flaviviruses, an expanding threat in public health: focus on dengue, West Nile, and Japanese encephalitis virus", Journal of NeuroVirology, 2014 Publication	1%
4	Submitted to Associatie K.U.Leuven Student Paper	1%
5	Svetlana F. Khaiboullina, Fabiola M. Ribeiro, Timsy Uppal, Ekaterina V. Martynova, Albert A. Rizvanov, Subhash C. Verma. "Zika Virus Transmission Through Blood Tissue Barriers", Frontiers in Microbiology, 2019 Publication	1%
6	etd.auburn.edu Internet Source	1%

7	Sahil Jain, Manoj Baranwal. "Computational analysis in designing T cell epitopes enriched peptides of Ebola glycoprotein exhibiting strong binding interaction with HLA molecules", Journal of Theoretical Biology, 2019 Publication	1%
8	E. Lee. "Both E Protein Glycans Adversely Affect Dengue Virus Infectivity but Are Beneficial for Virion Release", Journal of Virology, 05/15/2010 Publication	1%
9	Submitted to University of Liverpool Student Paper	1%
10	Submitted to University of Glasgow Student Paper	<1%
11	talk.ictvonline.org Internet Source	<1%
12	Lohia, Neha, and Manoj Baranwal. "Conserved Peptides Containing Overlapping CD4+ and CD8+ T-Cell Epitopes in the H1N1 Influenza Virus: An Immunoinformatics Approach", Viral Immunology, 2014. Publication	<1%
13	healthdocbox.com Internet Source	<1%

14	Submitted to Universiti Tunku Abdul Rahman Student Paper	<1%
15	link.springer.com Internet Source	<1%
16	Submitted to University of Hong Kong Student Paper	<1%
17	Submitted to Medizinischen Universität Wien Student Paper	<1%
18	Nicholas J. Barrows, Rafael K. Campos, Kuo-Chieh Liao, K. Reddisiva Prasanth et al. "Biochemistry and Molecular Biology of Flaviviruses", Chemical Reviews, 2018 Publication	<1%
19	Submitted to University of Birmingham Student Paper	<1%
20	oro.open.ac.uk Internet Source	<1%
21	www.nature.com Internet Source	<1%
22	Submitted to University of Bedfordshire Student Paper	<1%
23	microbiologysociety.org Internet Source	<1%
	Submitted to Savitribai Phule Pune University	

24	Student Paper	<1%
25	Submitted to Laureate Higher Education Group Student Paper	<1%
26	Submitted to Mahidol University Student Paper	<1%
27	journals.plos.org Internet Source	<1%
28	Raja, Kalpana, Suresh Subramani, and Jeyakumar Natarajan. "Template Filling, Text Mining", Encyclopedia of Systems Biology, 2013. Publication	<1%
29	repositorio.ufrn.br Internet Source	<1%
30	etheses.whiterose.ac.uk Internet Source	<1%
31	Submitted to Boston University Student Paper	<1%
32	Submitted to University of Oxford Student Paper	<1%
33	pubs.sciepub.com Internet Source	<1%
34	"Poster Abstracts", Immunology, 2013.	

	Publication	<1%
35	scholarworks.gsu.edu Internet Source	<1%
36	Sahil Jain, Manoj Baranwal. "Conserved peptide vaccine candidates containing multiple Ebola nucleoprotein epitopes display interactions with diverse HLA molecules", Medical Microbiology and Immunology, 2019 Publication	<1%
37	Submitted to University of Wolverhampton Student Paper	<1%
38	emedicine.medscape.com Internet Source	<1%
39	Submitted to CSU, Long Beach Student Paper	<1%

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		