Comparative Immunology, Microbiology and Infectious Diseases 65 (2019) 238–245

Contents lists available at ScienceDirect

Comparative Immunology, Microbiology and Infectious Diseases

journal homepage: www.elsevier.com/locate/cimid

A common conserved peptide harboring predicted T and B cell epitopes in domain III of envelope protein of Japanese Encephalitis Virus and West Nile Virus for potential use in epitope based vaccines

Parvez Singh Slathia*, Preeti Sharma

School of Biotechnology, Shri Mata Vaishno Devi University, Kakrial, Katra, J&K, India

ARTICLE INFO

Keywords:
Flavivirus
Envelope protein
Epitope prediction
Class I MHC
Class II MHC

ABSTRACT

Japanese encephalitis virus (JEV) and West Nile virus (WNV) are two major mosquito borne flaviviruses belonging to same serocomplex. JEV is transmitted by *Culex* mosquitoes and the reservoir host for the virus is pigs and/or water birds. WNV is also transmitted by *Culex* mosquitoes and reservoir host in this case is birds. It can also be transmitted through contact with other infected animals, their blood, or other tissues. The envelope protein of these viruses is the major source of epitopes and provides protective immunity. Bioinformatics tools were used to identify conserved epitopes in the envelope protein of these viruses. A conserved peptide "TPVGRLVTVNPFV" present in both the viruses containing predicted T and B cell epitopes was found. The model of one of the predicted epitope was generated and upon docking it bound in the groove of HLA-A0201 Class I MHC molecule. Further, it was amenable to proteasomal cleavage enhancing its chances of processing by cytosolic pathway. The peptide was found to be non toxic, non allergenic and stable in mammalian cells based on database search. The population coverage was pan world and nearly 70% identity of the peptide was found in the Zika virus envelope protein. The peptide was located in the domain III of envelope protein which is the exposed domain therefore B cell receptors may recognize this peptide easily. The conserved peptide containing T and B cell epitopes can have future application for designing epitope based vaccines for both JEV and WNV.

1. Introduction

Flaviviruses are arthropod borne viruses having small single stranded RNA and are transmitted by mosquitoes and ticks. Langat virus, Louping ill virus, Tick borne encephalitis virus are tick transmitted viruses comprising of single serocomplex whereas mosquito transmitted flaviviruses like dengue, yellow fever, West Nile virus (WNV), Japanese encephalitis virus (JEV) are serologically diverse [1]. JEV and WNV are members of the Japanese encephalitis serocomplex, which also includes St Louis encephalitis virus [2]. JEV is the main cause of viral encephalitis in many Asian countries and the estimated number of clinical cases in humans is 68,000 per year. The clinical disease progression occurs rarely, however the fatality rate in case of clinical encephalitis can be as high as 30%. The disease can result in permanent neurologic or psychiatric sequelae in 30%–50% cases of encephalitis. It is transmitted by *Culex tritaeniorhynchus* and the infection is an endemic problem in South-East Asia and Western Pacific regions [3]. More than 3 billion people in 24 countries mainly in Asia are prone to risks of infection [4]. With no specific antiviral therapy available for JEV, vaccination is the most important remedy for control of disease. The JEV vaccines have been in use for quite some time now. A list of vaccines that were used earlier or are currently in market is given in Table 1.The first JEV vaccine (JE-VAX) effectively controlled the disease in Japan and other Asian countries. The use of this vaccine has been discontinued because of hypersensitive responses, high production cost and some rare cases of severe neurological disorders associated with its use [5]. The new generation of JEV vaccines currently licensed for use comprise of inactivated vaccine (Ixiaro® or IC51) and chimeric vaccine (ChimeriVax-JE) [5,6].

WNV causes neuro invasive disease and death in people. WNV is commonly found in Africa, Europe, the Middle East, North America and West Asia. However, approximately 80% of people who are infected will not show any symptoms [7]. Since its emergence in 1999 in the U.S. it has spread to Asia, Europe and Middle East, and has been associated with 43,397 cases resulting in 1911 deaths in the U.S. till 2015 [8]. Culex are generally considered the principal vectors of WNV, in particular *Cx. pipiens* [9]. There is no available therapeutic intervention or vaccine available for the disease. Several WNV vaccines are licensed

* Corresponding author.
E-mail address: parvezsingh@gmail.com (P.S. Slathia).

https://doi.org/10.1016/j.cimid.2019.06.008
Received 15 February 2019; Received in revised form 19 June 2019; Accepted 20 June 2019

Table 1
List of vaccines (in market/under development) for Japanese Encephalitis Virus and West Nile Virus.

Japanese Encephalitis Virus			
Vaccine	Production Company	Design of Vaccine	Status
JE-VAX	BIKEN/ Sanofi-Pasteur	Inactivated vaccine derived from mouse brain	Production stopped
Ixiaro® or IC51	Intercell	Inactivated vaccine based on attenuated SA14-14-2 JEV strain grown in Vero cells	Licensed for use in US, Europe and Australia
ChimeriVax-JE	Sanofi-Pasteur	Chimeric vaccine based on the yellow fever virus (YFV) 17D virus back-bone containing the prM and E protein genes from JEV SA14-14-2	Licensed for use in Malaysia, Myanmar, Philippines, Korea, Cambodia, Thailand and Australia
West Nile Virus			
Vaccine	Developed by	Design of Vaccine	Status
ChimeriVax-WN02	Sanofi-Pasteur	Chimeric vaccine based on yellow fever virus (YFV) 17D vaccine strain containing WNV prM and E protein genes	Under clinical trials
HydroVax-001	Oregon Health & Science University	Inactivated WNV vaccine developed by using a novel, hydrogen peroxide-based process	Under clinical trials
VRC-302	Vaccine Research Center (VRC) & Vical	DNA plasmid vaccine expressing the WNV proteins prM and E	Under clinical trials
HBV-002	Hawaii Biotech	Purified recombinant WNV subunit vaccine containing proteins prM and truncated E protein	Under clinical trials

for veterinary use but no licensed human vaccine is commercially available [10]. There are several vaccines under clinical trials (enlisted in Table 1) which include live attenuated vaccine (ChimeriVax-WN02), inactivated WNV vaccine (HydroVax-001), DNA plasmid vaccine (VRC-302) and recombinant subunit vaccine (HBV-002) [11].

Most of the vaccines available for use in JEV or under clinical trials for WNV are based on traditional approaches of vaccine development. The conventional approaches are based on the killed or attenuated whole intact pathogen. Though these approaches have been successful in eradicating disease like small pox yet there are many drawbacks such as safety considerations during handling, difficulty in culturing of organisms in lab and most importantly the reversion into pathogenic forms. The genetic variation in pathogens can lead to decrease in efficacy of vaccines [12]. Novel approaches for vaccine development like DNA vaccines and epitope based vaccines can overcome these problems. Such vaccines based on appropriate epitopes have the potential to stimulate specific effective long lasting immunity while the side effects will be minimal [13].

In case of flaviviruses including WNV and JEV envelope (E) protein is postulated to be the major antigen harboring protein responsible for generating immunity. The E protein has an important role in cell receptor binding, generation of neutralizing antibodies, and induction of protective immunity. E protein has three domains designated as DI, DII and DIII. The structure of E protein in WNV and JEV follows the three-domain organization characteristic of flavivirus envelope proteins. The domain I is comprised of nine stranded β-barrel core, domain II is mostly made up of β strands whereas domain III forms an immunoglobulin like fold structure [14,15]. E protein is the major target for protecting antibodies and has been shown to be immunogenic and elicits protective immunity in both JEV and WNV [16–19]. In fact, the domain DIII is the major immunogenic region of this protein. Numerous studies have elucidated that DIII domain can induce strong immune response leading to production of neutralizing antibodies and providing long term immunity [18–21].

For designing epitope based vaccines, bioinformatics tools play an important role. The advances in bioinformatics tools coupled with increase in experimental data has led to the development of new field known as immunoinformatics [22]. Immunoinformatics studies for T cell and B cell epitope prediction to design epitope based vaccines has gained much importance. Studies on epitope prediction for vaccine design have carried out for viruses like Ebola virus [12], Zika Virus [23,24], bacteria like *Listeria monocytogenes* [25] and parasites like *Tryapnosoma brucei* [26], *Leishmania major* [27] and *Echinococcus granulosus* [28]. In the present study, E protein variants of WNV and JEV virus have been used to find the common conserved epitopes present in both the viruses. E protein sequences were retrieved from NCBI Identical protein groups database and were subjected to multiple sequence alignment. The alignment of sequences was used for finding out the conserved regions among them. The conserved regions were used to predict T and B cell epitopes. NetMHC 4.0 and NetMHCII 2.2 servers were used for predicting Class I MHC and Class II MHC epitopes, respectively. NetCTL 1.2 server was used for finding cytotoxic T cell (CTL) epitopes among Class I MHC epitopes. B cell epitopes were predicted by BCPREDS and ABCPred servers. Population coverage analysis, study of toxic and allergic properties of the predicted epitopes revealed worldwide immunity providing potential and none of the epitopes was toxic or allergenic. The epitope structure generated by molecular modeling was docked with HLA-A0201 which showed binding to the peptide binding cleft of the HLA molecule.

2. Materials and methods

2.1. Sequence retrieval and multiple sequence alignment

The NCBI Identical Protein Groups (IPG) database was used for retrieval of amino acid sequences of envelope protein of JEV and WNV.

The Identical Protein Groups resource contains a single entry for each protein translation found in several sources at NCBI, including annotated coding regions in GenBank and RefSeq, as well as records from SwissProt and PDB. This resource helps to obtain more targeted search results and quickly identifies a protein of interest (https://www.ncbi.nlm.nih.gov/ipg). The protein sequences retrieved were used for multiple sequence alignment by Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo) separately for each virus. The alignment files were analyzed for the presence of conserved regions common in both the viral protein variants. The common conserved regions were used for running predictions.

2.2. Prediction of peptides binding to MHC molecules

For prediction of peptides binding to Class I MHC molecules NetMHC4.0 server was used (http://www.cbs.dtu.dk/services/NetMHC/). NetMHC uses artificial neural network approach for peptide binding to Class I MHC [29,30]. Peptides binding to Class II MHC molecules were predicted by NetMHCII2.2 server (http://www.cbs.dtu.dk/services/NetMHCII/). NetMHCII 2.2 server predicts binding of peptides to HLA-DR, HLA-DQ, HLA-DP and mouse MHC class II alleles using artificial neural networks [31,32]. The peptides that were predicted as strong binders by NetMHC and NetMHCII were selected for further studies.

2.3. Confirmation of CTL epitopes

It has been observed that not all the Class I MHC binders can activate the Tc cells for their conversion into cytotoxic lymphocytes. Therefore, the peptides predicted as binders for Class I MHC molecules need to be confirmed as Cytotoxic T Lymphocyte (CTL) epitopes. NetCTL 1.2 server (http://www.cbs.dtu.dk/services/NetCTL/) was used to find out the CTL epitopes among the predicted binding peptides for Class I MHC [33]. NetCTL 1.2 server predicts CTL epitopes in protein sequences. Only those peptides which were predicted as epitopes were selected. However, there is no such prediction server available for predicting T_H cell epitopes peptides therefore all the predicted binding peptides for Class II MHC were taken as epitopes.

2.4. Prediction of B cell epitopes

Linear B cell epitope prediction was carried out using BCPREDS (http://ailab.ist.psu.edu/bcpred/predict.html) [34,35] and ABCPred servers (http://crdd.osdd.net/raghava/abcpred/) [36]. BCPREDS allows the user to select among three prediction methods: AAP method; BCPred and FBCPred. Users provide an antigen sequence and optionally can specify desired epitope length and specificity threshold. ABCpred server predicts B cell epitope(s) in an antigen sequence, using artificial neural network. This was the first server developed based on recurrent neural network (machine based technique) using fixed length patterns. The linear epitopes jointly predicted by both the servers were used.

2.5. IEDB analysis, self-epitope detection, toxicity prediction and allergenic properties

The Immune Epitope Database (IEDB) contains experimentally validated T and B cell epitopes. It offers easy searching of experimental data characterizing antibody and T cell epitopes studied in humans, non-human primates, and other animal species. Epitopes involved in infectious disease, allergy, autoimmunity, and transplant are included (http://www.iedb.org/). The predicted epitopes were studied for their presence in IEDB. These epitopes were then used for BLAST search against the NCBI protein database to check if they have similarity to any of the human proteins. To predict the toxic nature of the peptide ToxinPred server (http://crdd.osdd.net/raghava/toxinpred/) was used. ToxinPred is an in silico method, which is developed to predict and design toxic/non-toxic peptides [37]. The main dataset used in this method consists of 1805 toxic peptides (< = 35 residues). For finding if the peptide contains any allergenic sequence AlgPred server (http://crdd.osdd.net/raghava/algpred/) was used [38]. Prediction at this server was done by using hybrid option available in the server. The hybrid option of server allows predicting allergen using combined approach *viz.* Support Vector Machine (SVM) + IgE epitope + Allergen Representative Peptides BLAST (ARPs BLAST) + Multiple-antigen simultaneous test (MAST) motif search.

2.6. Population coverage

Since HLA alleles are highly polymorphic and their distribution varies in different ethnic groups. Further, the interaction between a T cell and its epitope is MHC restricted. The peptides that can be used for vaccine design should be able to bind a broad range of HLA molecules so as to make them effective in different population groups. Population coverage tool at IEDB analysis resource is designed on the basis of HLA allele frequencies of population groups of the world obtained from Allele Frequency database (http://www.allelefrequencies.net/). This tool (http://tools.immuneepitope.org/population/) was used to find out different population groups whose HLA alleles can respond to these predicted epitopes.

2.7. Peptide analysis in other flaviviruses

For checking the conservancy of the peptide 'TPVGRLVTVNPFV' in other flavivirus envelope protein, the envelope protein sequences of yellow fever virus, dengue virus, zika virus were retrieved from NCBI Identical Protein database. One sequence representative each of envelope protein was used. Epitope conservancy analysis tool available at IEDB analysis resource was run for finding out the conservancy of the peptide in different flavivirus envelope proteins (http://tools.immuneepitope.org/conservancy/). This tool computes the degree of conservancy of an epitope within a given protein sequence set at a given identity level. Since there was nearly 70% identity of the peptide in zika virus envelope protein, all full length protein sequences of its envelope protein were retrieved from NCBI Identical Protein database and run for conservancy analysis.

2.8. Peptide modeling, peptide properties and proteasomal cleavage prediction

The peptide 'TPVGRLVTVNPFV' was modeled using PEP-FOLD server (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/). PEP-FOLD is a de novo approach aimed at predicting peptide structures from amino acid sequences [39,40]. For in silico determination of properties of the peptide Protparam tool available at Expasy server was used. NetChop server (http://tools.immuneepitope.org/netchop/) at IEDB analysis resource was used for prediction of proteasomal cleavage sites in the peptide. The NetChop server produces neural network predictions for cleavage sites of the human proteasome [41]. NetChop has been trained on human data only, and will therefore presumably have better performance for prediction of the cleavage sites of the human proteasome.

2.9. Docking of peptide to HLA-A0201

The structure of epitope "RLVTVNPFV" was modeled by PEP-FOLD and used for docking to HLA-A0201 molecule. The PDB structure of HLA-A0201 was obtained from PDB database bearing ID 3MRK. This is a co crystallized structure of HLA-A0201 molecule with a nonapeptide. The nonapeptide bound in the cleft was removed and the PDB file so obtained was used for docking. Z-Dock server (http://zdock.umassmed.edu/) was used for carrying out docking experiments [42]. It is a server that carries out interactive protein-protein docking prediction.

2.10. Designing of vaccine

The peptide "TPVGRLVTVNPFV" containing T and B cell epitopes can be used for design of DNA vaccine. In order to design DNA vaccine the peptide needs to be reverse translated. For this purpose Reverse Translate tool at Expasy tools (http://www.bioinformatics.org/sms2/rev_trans.html) can be used. The designed DNA vaccine should express in humans and as such the codon table of humans could be used while reverse translating the peptide. Cloning into DNA vaccine vectors like pcDNA3.1 vector can be done. pcDNA3.1 is a mammalian expression vector and has been widely used for DNA vaccine design. The codon optimized DNA sequence of the peptide can be inserted into the multiple cloning site of the vector. Alternatively, synthetic peptide vaccine can also be constructed. The peptide can be synthesized but needs to administered along with a suitable adjuvant.

3. Results

3.1. Sequence retrieval, multiple sequence alignment and conservancy analysis

358 sequences of envelope protein of JEV and 27 sequences of envelope protein of WNV were retrieved from NCBI Identical protein database. These were the total number of entries in the database for these proteins at the time when study was conducted. The benefit of using identical protein database is that it reduces the redundancy of data. There can be different entries for the protein having the same sequence in NCBI protein database and as such the protein with same sequence can be over represented. The full length sequences having length of more than 495 amino acid residues were selected and the sequences of less amino acid residue length were omitted for carrying out multiple sequence alignment. Alignments for JEV E protein and WNV E protein were carried out separately and the alignment files obtained were analysed for the presence of common conserved sequences ≥ 9 amino acid residues in length. The three common conserved peptides found are shown in Table 2 along with their position in the respective envelope proteins.

3.2. Prediction of peptides binding to MHC molecules

For Class I MHC, three epitopes were predicted for peptide no.2 (mentioned in Table 2) whereas neither peptide 1 nor peptide 3 generated any Class I MHC binding peptides. In case of Class II MHC only one epitope was predicted from peptide 2 while no epitope was predicted for the other two conserved peptides. The peptides that were predicted as strong binders by the servers were selected. The epitopes binding to different MHC alleles are shown in Table 3. From Table 3 it becomes clear that two Class I MHC epitopes are promiscuous in nature. Epitope 1 "RLVTVNPFV" showing strong binding to 10 alleles and epitope 2 "GRLVTVNPF" showing strong binding affinity for 4 alleles whereas epitope 3 is binding to only one MHC allele. In case of Class II MHC the only predicted epitope shows promiscuous binding for five different alleles. The promiscuous nature of the epitopes is a desirable property as one epitope can generate immunity among different ethnic populations.

Table 2
Conserved peptide sequences common for JEV and WNV Envelope protein.

S.No.	Peptide Sequence	Position in JEV E protein	Position in WNV E protein
1.	DRGWGNGCGL	98–107	122–131*, 98–107
2.	TPVGRLVTVNPFV	349–361	374–386*, 350–362
3.	EPPFGDSYIVVGRG	375–388	400–413*, 376–389

* WNV E protein sequences retrieved from the IPG database formed two groups. One sequence group had 24 amino acid residues more than the other in the start of the sequence. Thus, in one group of sequences the position of the conserved peptides exceeds by 24 amino residues than the other. So, the position of the peptides is represented twice for WNV one for each group.

3.3. Confirmation of CTL epitopes

Cytotoxic T lymphocyte epitopes are the ones that can be recognized by Tc cells and can trigger their conversion to CTLs. When Class I MHC epitopes were analyzed using NetCTL1.2 server all the three epitopes were found out to be CTL epitopes. For the want of any such prediction servers that can do the same for Class II MHC the predictions for T_H cell epitopes could not be run.

3.4. B cell epitope prediction

Since peptide no. 1 was only 10 amino acid residues in length and BCPREDS allows prediction for minimum 12 amino acid residue epitope therefore it was not used for linear B cell epitope prediction runs. For the other two peptides one linear B cell epitope each was reported by both the servers (Table 4). These peptides could serve as linear B cell epitopes as they lie in the domain III of envelope protein of the viruses. This domain is the major antibody binding site and is highly immunogenic. Epitopes located on domain III of WNV are likely to be highly significant in the development of protective immunity.

3.5. Epitope analysis in IEDB, allergenic and toxic properties

Both T and B cell epitopes when checked for their presence in IEDB revealed that only one epitope *viz.* "GRLVTVNPF" was found to be present in the database and has been shown experimentally binding to HLA-B27:05. None of the other epitopes were present in the IEDB. Since the conserved peptide no. 2 contains both T and B cell epitopes, this peptide was used for further analysis and conserved peptide no. 3 was not processed further. The epitopes did not show any similarity to human protein sequences thus negating the chances of being treated as self-epitopes. The selected conserved peptide was predicted to be non-toxin using Support Vector Machine (Swiss-Prot) method in ToxinPred server. AlgPred results showed that there was no allergenic sequence present in the peptide. Hence, it can be characterized as non-allergenic, non-toxin peptide and can be used for vaccine designing purpose.

3.6. Peptide modeling and properties

The best model generated by PEP-FOLD for the peptide "TPVGRLVTVNPFV" is shown in Fig. 1. The properties of peptide generated by Protparam tool are depicted in Table 5. The peptide was predicted to be stable and was estimated to have half-life of 7.2 h in mammalian reticulocytes *in vitro*. Proteasomal cleavage prediction revealed by Netchop revealed the presence of three potential cleavage sites in the peptide thus making it amenable to be processed by the cytosolic pathway (Fig. 2).

3.7. Docking

Z dock server generated 10 bound complexes for the epitope and HLA-A0201 wherein the epitope "RLVTVNPFV" binds to the peptide binding cleft of HLA- A0201 molecule and the binding score is 1111.346 (Fig. 3). This further confirms that the epitope can bind to HLA molecule and can be presented to Tc cell receptor for recognition.

3.8. Population coverage and peptide presence in other flaviviruses

The population coverage of both Class I and Class II MHC combined is shown in Fig. 4. The results showed that the coverage range of the epitopes is worldwide and the highest coverage is for European population (70%). The epitopes have the potential to cover 60% of the world

Table 3
Epitopes predicted for Class I and Class II MHC from the conserved peptides.

S.No.	Epitope	HLA Class	Allele	NetCTL
1.	RLVTVNPFV	Class I MHC	HLA-A0201, HLA-A0202, HLA-A0205, HLA-A0206, HLA-A0207, HLA-A0211, HLA-A0212, HLA-A0216, HLA-A0217, HLA-A0219	Epitope
2.	GRLVTVNPF	Class I MHC	HLA-B2705, HLA-B2720, HLA-B4801, HLA-C0702	Epitope
3.	TPVGRLVTV	Class I MHC	HLA-B0702	Epitope
4.	TPVGRLVTVNPFV	Class II MHC	HLA-DRB-10110, HLA-DRB10120, HLA-DRB10123, HLA-DRB10102, HLA-DRB10404	Not applicable

Table 4
B cell epitopes predicted by BCPREDS and ABCPred servers.

Conserved Peptide no.	Sequence of Peptide	BCPREDS linear B cell epitope	ABCPred linear B cell epitope
2.	TPVGRLVTVNPFV	PVGRLVTVNPFV	PVGRLVTVNPFV
3.	EPPFGDSYIVVGRG	EPPFGDSYIVVG	EPPFGDSYIVVG

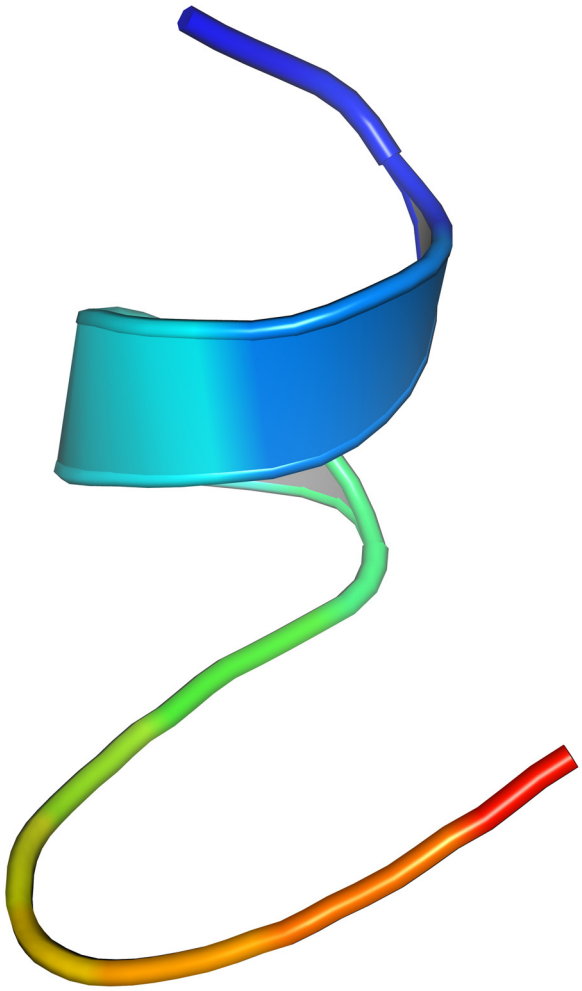

Fig. 1. Structure of the conserved peptide "TPVGRLVTVNPFV".

population. The presence of the conserved peptide when checked in other flaviviruses revealed that it was present in the other flavivirus envelope protein sequences used in the study with varying identity

Table 5
Properties of the peptide predicted by Protparam tool.

S.No.	Property	Value for the peptide
1.	Number of amino acids	13
2.	Molecular weight	1398.67
3.	Theoretical pI	9.41
4.	Formula	$C_{65}H_{107}N_{17}O_{17}$
5.	Total number of atoms	206
6.	Instability index	23.58
7.	Stable/Unstable	Stable protein
8.	Aliphatic index	119.230
9.	GRAVY	0.800

(Table 6). The maximum identity was found in Zika virus envelope protein whereas minimum in case of Dengue virus 4 envelope protein. The peptide showed the same identity (~70%) when compared with all Zika virus envelope proteins obtained from NCBI IPG database.

4. Discussion

Vaccine development against pathogenic organisms is of prime importance for reducing the disease burden to ease human life on earth. Viruses pose a great threat because they use the host cell machinery for growth and have small genome size. There is presence of fewer targets for drug development because of this. In addition, the faster rate of mutation in viruses further complicates the problem. Thus, vaccines become the major prophylactic measure for viruses. The advance in sequencing technology has made available a wealth of genome sequence data. The use of genome sequences can be made to develop vaccines in silico. This approach known as reverse vaccinology can significantly reduce the timeline for finding out new candidate vaccines. There is no requirement of growing the pathogen in lab. The development of vaccine for *Neisseria meningitidis* serogroup B has been successfully achieved by reverse vaccinology methods clearly illustrating the potential of this approach. The protein sequence data can be obtained from genome by in silico translation of annotated genome sequences. This has led to the development of several immunoinformatics tools that have the potential to accelerate the process of vaccine development. In our study JEV and WNV viral envelope protein was used for vaccine design using immunoinformatics tools. E protein of JEV and WNV is the site of binding for many neutralizing antibodies and provides protective immunity against infection. The protein sequences used in the study were obtained from NCBI IPG database. The E protein sequences retrieved from NCBI IPG database have the advantage over those obtained from NCBI Protein database. There could be many entries for the protein with same sequence in NCBI protein database however; in case of IPG database all the protein entries having the same sequence form a group. One sequence can be used as representative for the whole group. This leads to avoidance of repetition

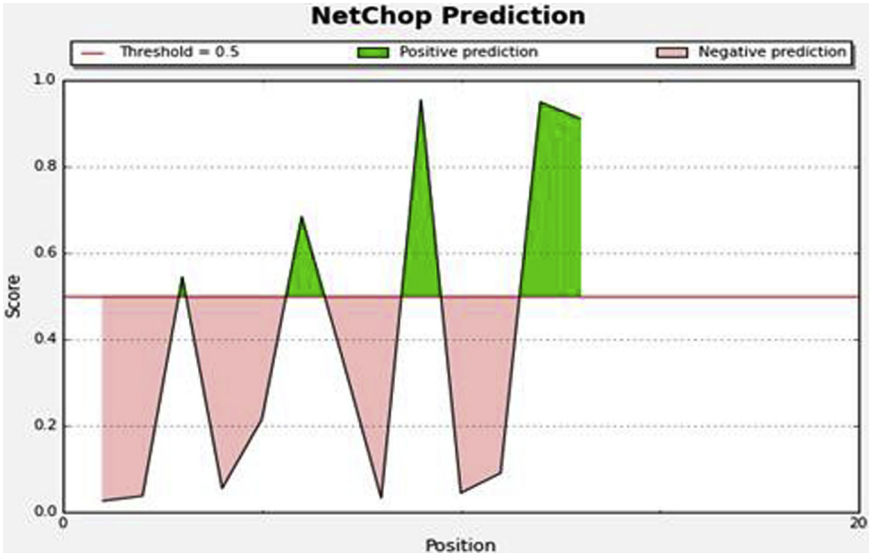

Fig. 2. Predicted proteasomal cleavage sites in peptide by NetChop.

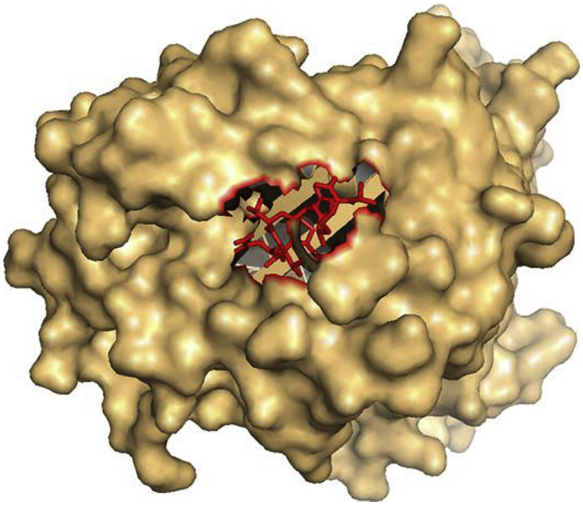

Fig. 3. Class I MHC Epitope "RLVTVNPFV" bound to cleft of HLA-A0201 molecule (peptide in red color and the HLA molecule in wheat color).

of data and overrepresentation of certain sequences. The peptides that can act as T cell epitopes must bind to MHC molecules. The binding of peptides to Class I and Class II MHC molecules can be predicted by many methods like Position-Specific Scoring Matrix (PSSM), Support Vector Machine (SVM) and Artificial Neural Network (ANN) methods. SYFPEITHI [43] RANKPEP [44] are the PSSM predictors, SVMHC [45] is SVM predictor and NetMHC [30] is ANN prediction server. ANN and SVM methods outperform the PSSM methods. PSSM methods fail to model the nonlinearity of the binding process and the interrelationship between different binding positions. SVM and ANN methods have the ability to model these effects and thus show superior performance [46]. In the benchmarking studies conducted for peptide MHC binding prediction servers it was found that among all the methods NetMHC is the best performing method [47]. Thus, NetMHC was chosen for running MHC peptide binding predictions. The prediction efficiency of MHC Class II epitope predictors is not as good as MHC Class I epitope predictors and among the available methods NetMHCII has shown better performance [46]. The processing of antigens is necessary for recognition by T cells. The prediction methods for processing of Class I MHC antigens are available but unfortunately the same is not true for Class II MHC antigens. NetCTL predicts the binding to MHC class I, proteasomal cleavage and TAP transport efficiency [33]. Class I MHC binding and proteasomal cleavage is predicted using ANN whereas TAP transport efficiency is predicted using weight matrix. Based on the results of NetCTL all the binding peptides predicted by NetMHC turned out to be epitopes as they were susceptible to processing and presentation pathways. The results from NetChop revealed three potential

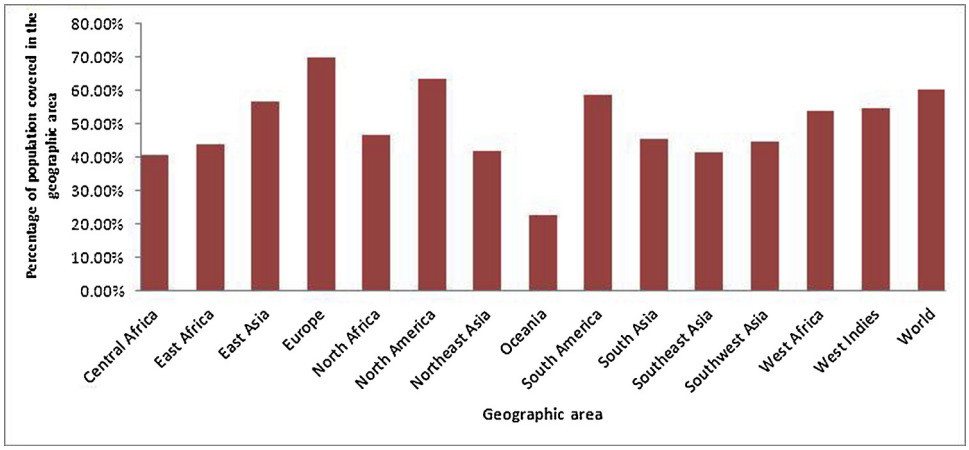

Fig. 4. Geographical area population coverage percentage of the epitopes.

Table 6
Identity of the conserved peptide "TPVGRLVTVNPFV" in other flavivirus envelope protein.

Protein Accession Number & Name	Positions in sequence	Identity
ATP66566.1 envelope protein Dengue virus 1	346–358	61.54%
ATP66571.1 envelope protein Dengue virus 2	346–358	61.54%
ATP66578.1 envelope protein Dengue virus 3	344–356	53.85%
ATP66580.1 envelope protein Dengue virus 4	346–358	38.46%
YP_009227198.1 envelope protein Zika virus	349–361	69.23%
AAY68346.1 envelope protein Yellow fever virus	344–356	53.85%

proteasomal cleavage sites in the peptide further confirming the vulnerability to proteasomal cleavage. B cell epitopes are both linear/continuous and conformatiobnal/discontinuous. Most of the B cell epitopes reported so far are linear in nature and as such the methods for B cell epitope prediction are similar to that of T cell epitopes [46,48]. Prediction of discontinuous or structural epitopes is difficult and the available prediction methods have very low accuracy. We have predicted only the linear/continuous B cell epitopes in this work.

There have been many reports of cross reactivity in flaviviruses. The evidence of cross protection between WNV and JEV is well established [49,50]. In a recent study in mice it was seen that Zika virus infection provided protection against WNV infection [51]. Since most of the Japanese population is seropositive for anti- JEV antibodies because of previous JEV vaccination or natural infection it was demonstrated recently that intravenous immunoglobulin (IVIG) originating from the plasma of Japanese individuals, where WNV endemic infection has not been found, has neutralizing activity against both JEV and WNV [52]. In our study we found out a region harboring both T and B cell epitopes located in domain III of the E protein. Domain III of the E protein is the major immunogenic part of the protein as mentioned in various earlier studies. The epitopes revealed in the study were searched for in IEDB wherein except one epitope all other were not present in the database suggesting that these were novel epitopes. In one of the earlier works WNV polyprotein has been used to for identification of Tc cell epitopes in which one of these epitopes 'GRLVTVNPF' was reported and none of the other two epitopes find a mention [51]. The same epitope has been experimentally validated and is present in IEDB. However, Class II MHC epitope and B cell epitope prediction was not carried out in the above study. The conserved peptide 'TPVGRLVTVNPFV' is non allergen and non-toxin in nature as per AlgPred and ToxinPred servers and thus no undesirable responses are predicted to be generated in the host against these peptides. The peptide was found to be stable in nature. The stability of protein provides the time period required for interaction with immune effector molecules. The binding of epitopic peptide in the cleft of HLA molecule is required for its presentation to T cells. The modeled epitope 'RLVTVNPFV' upon docking to HLA-0201 showed binding in the peptide harboring cleft of the molecule further enhance the chances of its acting as epitope and being presented to Tc cells. The coverage of a wide range of population groups is essential for developing a T cell based which would be successful throughout the world populations. The population coverage of these epitopes is pan world and the vaccine designed by using these epitopes has the potential to provide immunity universally. The maximum identity of the conserved epitopic peptide was found out with envelope protein of Zika virus (70% approximately). Recently, Zika virus infection has shown to be protective against WNV in mice [52]. There are chances that this epitope based vaccine may work against Zika virus as well. For different dengue virus strains the identity ranged from 38% to 61% which could limit the efficiency against this virus. The cross neutralizing activity is very less against Dengue virus for (IVIG) originating from the plasma of Japanese individuals [53].

Finally, from this study we were able to find a stretch of amino acids located in the domain III of envelope protein which is conserved in WNV and JEV and also having 70% identity in the Zika virus E protein. The peptide contains both T and B cell epitopes. It harbors sites for processing pathway required for presentation of antigens to T cells. The wide coverage of world population groups further enhances the usefulness of the peptide. Besides, it is non-toxic and non-allergenic in nature. This conserved peptide harboring both T and B cell epitopes has the potential to be used in epitope based vaccines for both JEV and WNV. The DNA and peptide vaccines can be designed based on the peptide and suitable *in vivo* studies need to be carried out to check the immunity generated.

References

[1] M.R. Holbrook, Historical perspectives on flavivirus research, Viruses 9 (2017) 97.
[2] B.K. Schweitzer, N.M. Chapman, P.C. Iwen, Overview of the flaviviridae with an emphasis on the Japanese encephalitis group viruses, Lab. Med. 40 (2009) 493–499.
[3] WHO, Japanese encephalitis. Factsheet No 386, December (2015) http://www.who.int/mediacentre/factsheets/fs386/en/.
[4] T.E. Erlanger, S. Weiss, J. Keiser, J. Utzinger, K. Wiedenmayer, Past, present, and future of Japanese encephalitis, Emerg. Infect. Dis. 15 (2009) 1–7.
[5] S.I. Yun, Y.M. Lee, Japanese encephalitis: the virus and vaccines, Hum. Vaccin. Immunother. 10 (2014) 263–279.
[6] F.X. Heinz, K. Stiasny, Flaviviruses and flavivirus vaccines, Vaccine 30 (2012) 4301–4306.
[7] WHO, West Nile Virus. Fact Sheet No. 354, July (2011) http://www.who.int/mediacentre/factsheets/fs354/en/.
[8] CDC: West Nile Virus, (2016) https://www.cdc.gov/westnile/statsmaps/index.html.
[9] L.R. Petersen, A.C. Brault, R.S. Nasci, West Nile virus: review of the literature, JAMA 310 (2013) 308–315.
[10] M.D. Filette, S. Ulbert, M.S. Diamond, N.N. Sanders, Recent progress in West Nile virus diagnosis and vaccination, Vet. Res. 43 (2012) 16.
[11] I. Scherwitzl, J. Mongkolsapaja, G. Screaton, Recent advances in human flavivirus vaccines, Curr. Opin. Virol. 23 (2017) 95–101.
[12] M.A. Khan, M.U. Hossain, S.M. Rakib-Uz-Zaman, M.N. Morshed, Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study, Scand. J. Immunol. 82 (2015) 25–34.
[13] R. Arnon, A novel approach to vaccine design - Epitope-based vaccines, Biotech. Int. 18 (2006) 10–13.
[14] R. Kanai, K. Kar, K. Anthony, L.H. Gould, M. Ledizet, E. Fikrig, W.A. Marasco, R.A. Koski, Y. Modis, Crystal structure of the West Nile Virus Envelope glycoprotein reveals viral surface epitopes, J. Virol. 80 (2006) 11000–11008.
[15] V.C. Luca, J.A. Mansour, C.A. Nelson, D.H. Fremont, Crystal structure of the japanese encephalitis virus envelope protein, J. Virol. 86 (2012) 2337–2346.
[16] K.W. Taeg, W.S. Lee, P.J. Park, T.K. Park, H. Kang, Protective immunity of *Pichia pastoris*-expressed recombinant envelope protein of japanese encephalitis virus, J. Microbiol. Biotechnol. 22 (2012) 1580–1587.
[17] M. Ledizet, K. Kar, H.G. Foellmer, T. Wang, S.L. Bushmich, J.F. Anderson, E. Fikrig, R.A. Koski, A recombinant envelope protein vaccine against West Nile virus, Vaccine 23 (2005) 3915–3924.
[18] A. Schneeweiss, S. Chabierski, M. Salomo, N. Delaroque, S. Al-Robaiy, T. Grunwald, K. Burki, U.G. Liebert, S. Ulbet, A DNA vaccine encoding the E protein of West Nile virus is protective and can be boosted by recombinant domain DIII, Vaccine 29 (2011) 6352–6357.
[19] J.H.J. Chu, C.C.S. Chiang, M.L. Ng, Immunization of flavivirus West Nile recombinant envelope domain III protein induced specific immune response and protection against West Nile virus infection, J. Immunol. 178 (2007) 2699–2705.
[20] B.M. Friedrich, D.W.C. Beasley, J.S. Rudra, Supramolecular peptide hydrogel adjuvanted subunit vaccine elicits protective antibody responses against West Nile virus, Vaccine 34 (2016) 5479–5482.
[21] S.A. Liu, M. Haque, B. Stanfield, F.M. Andrews, A.A. Roy, K.G. Kousoulas, A recombinant fusion protein consisting of West Nile virus envelope domain III fused in-frame with equine CD40 ligand induces antiviral immune responses in horses, Vet. Microbiol. 198 (2017) 51–58.
[22] A. Patronov, I. Dotchinova, T- cell epitope vaccine design by immunoinformatics, Open Biol. 3 (2013) 120139.
[23] A. Alam, S. Ali, S. Ahamad, M.Z. Malik, R. Ishrat, From ZikV genome to vaccine: in silico approach for the epitope-based peptide vaccine against Zika virus envelope glycoprotein, Immunology 149 (2016) 386–399.
[24] P.S. Slathia, P. Sharma, Epitiope based DNA vaccine design using epitopes predicted from Zika virus polyprotein, Int. J. Infect. Dis. 53 (2016) 149.
[25] A. Jahangiri, I. Rasooli, S.L. Gargari, P. Owlia, M.R. Rahbar, J. Amani, S. Khalili, An in silico DNA vaccine against *Listeria monocytogenes*, Vaccine 29 (2011) 6948–6958.
[26] B.A. Akhoon, P.S. Slathia, P. Sharma, S.K. Gupta, V. Verma, In silico identification of novel protective VSG antigens expressed by *Trypanosoma brucei* and an effort for

designing a highly immunogenic DNA vaccine using IL-12 as adjuvant, Microb. Pathog. 51 (2011) 77–87.
[27] P. Slathia, P. Sharma, Predicting conserved epitopes in the variants of amastin protein of *Leishmania major* and designing an epitope based immunogenic DNA vaccine, Int. J. Infect. Dis. 16 (2012) e314.
[28] Y.J. Lu, D.S.H. Chen, W.T. Hao, H.W. Xu, Y.W. Zhang, F.F. Sun, W. Pan, In silico characterization of *Echinococcus granulosus* paramyosin nucleotide sequence for the development of epitope vaccine against cystic echinococcosis, Helminthologia 54 (2017) 275–283.
[29] M. Andreatta, M. Nielsen, Gapped sequence alignment using artificial neural networks: application to the MHC class I system, Bioinformatics 32 (2016) 511–517.
[30] M. Nielsen, C. Lundegaard, P. Worning, S.L. Laumoller, K. Lamberth, S. Buus, S. Brunak, O. Lund, Reliable prediction of T-cell epitopes using neural networks with novel sequence representations, Protein Sci. 12 (2003) 1007–1017.
[31] M. Nielsen, O. Lund, NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction, BMC Bioinf. 10 (2009) 296.
[32] M. Nielsen, C. Lundegaard, O. Lund, Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method, BMC Bioinf. 8 (2007) 238.
[33] M.V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, NetCTL-1.2: large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction, BMC Bioinf. 8 (2007) 424.
[34] J. Chen, H. Liu, J. Yang, K. Chou, Prediction of linear B-cell epitopes using amino acid pair antigenicity scale, Amino Acids 33 (2007) 423–428.
[35] Y. EL-Manzalawy, D. Dobbs, V. Honavar, Predicting linear B-cell epitopes using string kernels, J. Mol. Recognit. 21 (2008) 243–255.
[36] S. Saha, G.P.S. Raghava, Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, Proteins 65 (2006) 40–48.
[37] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, In silico approach for predicting toxicity of peptides and proteins, PLoS One 8 (2013) e73957.
[38] S. Saha, G.P.S. Raghava, AlgPred: prediction of allergenic proteins and mapping of IgE epitopes, Nucleic Acids Res. 34 (2006) W202–W209.
[39] Y. Shen, J. Maupetit, P. Derreumaux, P. Tufféry, Improved PEP-FOLD approach for peptide and miniprotein structure prediction, J. Chem. Theor. Comput. 10 (2014) 4745–4758.
[40] P. Thévenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux, P. Tuffery, PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides, Nucleic Acids Res. 40 (2012) W288–293.
[41] M. Nielsen, C. Lundegaard, O. Lund, C. Kesmir, The role of the proteasome in generating cytotoxic T cell epitopes: insights obtained from improved predictions of proteasomal cleavage, Immunogenetics 57 (2005) 33–41.
[42] B.G. Pierce, K. Wiehe, H. Hwang, B.H. Kim, T. Vreven, Z. Weng, ZDOCK Server: interactive docking prediction of protein-protein complexes and symmetric multimers, Bioinformatics 30 (2014) 1771–1773.
[43] H. Rammensee, J. Bachmann, N.P. Emmerich, O.A. Bachor, S. Stevanovic, SYFPEITHI: database for MHC ligands and peptide motifs, Immunogenetics 50 (1999) 213–219.
[44] P.A. Reche, J.P. Glutting, E.L. Reinherz, Prediction of MHC class I binding peptides using profile motifs, Hum. Immunol. 63 (2002) 701–709.
[45] P. Dönnes, A. Elofsson, Prediction of MHC class I binding peptides, using SVMHC, BMC Bioinf. 3 (2002) 25.
[46] L. Backert, O. Kohlbacher, Immunoinformatics and epitope prediction in the age of genomic medicine, Genome Med. 7 (2015) 119.
[47] T. Trolle, I.G. Metushi, J.A. Greenbaum, Y. Kim, J. Sidney, O. Lund, A. Sette, B. Peters, M. Nielsen, Automated benchmarking of peptide-MHC class I binding predictions, Bioinformatics 31 (2015) 2174–2181.
[48] J.V. Kringelum, M. Nielsen, S.B. Padkjær, O. Lund, Structural analysis of B-cell epitopes in antibody:protein complexes, Mol. Immunol. 53 (2013) 24–34.
[49] S.H. Li, X.F. Li, H. Zhao, T. Jiang, Y.Q. Deng, X.D. Yu, Q.Y. Zhu, E.D. Qin, C.F. Qin, Cross protection against lethal West Nile virus challenge in mice immunized with recombinant E protein domain III of Japanese encephalitis virus, Immunol. Lett. 138 (2011) 156–160.
[50] N.M. Nemeth, A.M. Bosco-Lauth, R.A. Bowen, Cross-protection between West Nile and Japanese encephalitis viruses in red-winged blackbirds (*Agelaius phoeniceus*), Avian Dis. 53 (2009) 421–425.
[51] M.V. Larsen, A. Lelic, R. Parsons, M. Nielsen, I. Hoof, K. Lamberth, M.B. Loeb, S. Buus, J. Bramson, O. Lund, Identification of CD8 + t cell epitopes in the West Nile Virus Polyprotein by reverse-immunology using NetCTL, PLoS One 5 (9) (2010) e12697.
[52] A.V. Calvo, A.B. Blázquez, E.E. Romero, M.T. Ramos, J.C. Saiz, M.A.M. Acebes, J.N. de Oya, Zika virus infection confers protection against West Nile virus challenge in mice, Emerg. Microbes Infect. 6 (2017) e81.
[53] N. Shinohara, T. Owada, C. Matsumoto, S. Uchida, T. Nagai, M. Satake, K. Tadokoro, Evaluation of the protective ability of plasma from Japanese individuals against mosquito-borne viral infections, Trans. R. Soc. Trop. Med. Hyg. 111 (2017) 393–401.