

Noname manuscript No.
(will be inserted by the editor)

Recognition of A Highly Conserved DSRCPTQ Epitope in Envelope Protein of Zika Virus Through *in silico* Approaches

Anisur Rahman** · Md. Iqbal Hossain · Saheda Tamanna · Md Neamat Ullah

Received: date / Accepted: date

Abstract The recurrent and recent outbreak of Zika Virus (ZIKV) has turned into global concern as yet there is no appropriate preventive measure been found. Situation getting worse as this virus also associates with several birth defects in neonatal such as primary microcephaly as well as many other neurological disorders. ZIKV adopts a wide host range which has hastened its expansion more recklessly. Hence, now there is an acute demand for developing a preventive vaccine against ZIKV. Immunoinformatic techniques have been employed in this study to pick out a highly conserved versatile antigenic B-cell linear epitope for all strains of ZIKV. Capsid protein (C), Membrane Glycoprotein Precursor (PrM), envelope protein (E) and RNA-dependent RNA Polymerase (NS5) have investigated by the implementation of sequence analysis and different epitope prediction methods. Some potential linear peptides have been recognized and tested for hydrophilicity and conservancy. Peptide with best antigenic properties was selected as ultimate final epitope and further structural exploration revealed its compatible position in protein 3D structure. Being fully conserved in all strains of ZIKV and its position in Envelope protein suggested epitope DSRCPTQ can be a quantum leap in the advancement of ZIKV prevention. However, extensive *in*

vitro plus *in vivo* experimentations are needed to be clarified about the real potency of the selected epitope.

Keywords Zika Virus · epitope prediction · *in silico* · B-cell epitope · vaccine development

1 Introduction

Zika virus is a member of Flaviviridae virus family which contains a single-stranded positive-sense RNA (Malone et al., 2016). Its genome size is about 10.7 kb that encodes a single polyprotein that cleaved into different structural and non-structural proteins. Non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) are associated with virus replication and multiplication in the host cell. And, structural proteins (capsid protein (C), Membrane Glycoprotein Precursor (PrM)-that cleaved into mature membrane protein (M) and envelope protein (E)) are involved in viral assembly (Lindenbach and Rice, 2003). The name of Zika Virus is based on Zika Forest of Uganda, where the virus was isolated from a rhesus macaque monkey for the first time in 1947 (Sikka et al., 2016).

First ZIKV infected human was noticed from a serological survey in Uganda (Dick, 1952). Later on, several studies in African and Asian countries revealed that the virus had been widespread within the human population of these regions (Wikan and Smith, 2016). In 2015 there as an outbreak of ZIKV in Brazil since then it has spread through tropical Americas and in 2015 and 2016 it became epidemic. The spread of ZIKV to other countries around the world is also reported in several studies (Duffy et al., 2009; Ety, 2014).

Though preliminarily it was believed that ZIKV causes mild feverish infection with rash and headache (Dick et al., 1952), now it has also been reported to associate with Guillian-Barr Syndrome (GBS) (Cao-Lormeau et al., 2016) and several other neurological complications as well (Tappe

**Anisur Rahman

Department of Biotechnology and Genetic Engineering, Faculty of Life Science, Noakhali Science and Technology University, Sonapur, Noakhali-3814, Bangladesh.

ORCID: 0000-0001-8279-9810

Tel.: +880-1623092488

E-mail: arriyaz.nstu@gmail.com

Md. Iqbal Hossain

Department of Biochemistry and Molecular Biology, Faculty of Life Science, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh.

Saheda Tamanna · Md Neamat Ullah

Department of Biotechnology and Genetic Engineering, Faculty of life Science, Noakhali Science and Technology University, Sonapur, Noakhali-3814, Bangladesh.

et al., 2015). Primary microcephaly and miscarriage in the time of pregnancy also have been noted in many cases due to ZIKV infection (Brasil et al., 2016; Mlakar et al., 2016). Besides, the presence of ZIKV in infected patients semen indicates that it might transmit by sexual means (Mansuy et al., 2016). Till date, no therapeutics or vaccine has been approved against ZIKV which makes its expansion and prevention more difficult (Abbink et al., 2018; Fernandez and Diamond, 2017; NIH, 2018).

In the recent years research on ZIKV have remarkably increased (Aid et al., 2017; Faye et al., 2014). Structure of the virion has been resolved and complete genome sequence was identified (Cunha et al., 2016; Kostyuchenko et al., 2016). ZIKV mainly targets neuroprogenitor cells, which eventually causes Congenital Zika Syndrome (CZS) and apoptosis (Lin et al., 2017). ZIKV is closely related with many other member of Flavivirus family such as dengue virus (DENV), West Nile virus (WNV), tick-borne encephalitis virus (TBEV) Japanese encephalitis virus (JEV) which advanced its vaccine development exclusively (Lanciotti et al., 2008).

Currently, a number of vaccines (for example; DNA vaccine, mRNA vaccine, purified inactivated virus vaccine, and viral vector vaccines) for ZIKV are in preclinical and clinical development. Plasmids-coding the gene of interest to induce either of humoral and cellular immunity-are called DNA vaccine (Ferraro et al., 2011). Several ZIKV membrane glycoprotein precursor (PrM) and envelope protein coding DNA vaccines are still under clinical trial (Gaudinski et al., 2018; Tebas et al., 2017). About three ZIKV inactivated virus vaccine (ZPIV) were reported to be in phase I clinical trials (Modjarrad et al., 2018). mRNA vaccines are comparatively newer type of vaccine, which also been developed for ZIKV (Pardi et al., 2018). Additionally, viral vector-based approaches have been implemented to develop ZIKV vaccine. Various viruses (modified vaccinia virus Ankara (MVA), measles virus (MV) and adenovirus (Ad) vectors) are being engineered to function as ZIKV vaccine (Brault et al., 2017; Xu et al., 2018).

Researchers have been highly interested in Envelope protein (E), NS3 and NS5 as these proteins have a significant role in viral access and replication into the host cell (Malet et al., 2008; Perera et al., 2008). Studies on flavivirus vaccine revealed that antibodies designed to bind with Envelope protein (E) might be an efficient protection against ZIKV (Nowakowski et al., 2016). But, the risk of autoimmune response against potential epitope must be considered during vaccine development for ZIKV (Homan et al., 2016).

Various computer aided immunological approaches have been promising in vaccine design because of the availability of a large number of genomic and immunoinformatic data (De Gregorio and Rappuoli, 2012; Patronov and Doytchinova, 2013). Nowadays vaccine design has become more

facile than ever in the matter of both time and expenses due to the application of computational methods and algorithms in this field.

Generally, Epitopes are two types; continuous and discontinuous. Continuous epitopes are simple linear peptide sequences while discontinuous epitopes are structurally more complex, nonlinear and assembled together for folding in the natural form of protein. According to their corresponding receptors, epitopes can be also classified into B- and T cell epitopes. B cell epitopes contain both continuous (10%) and discontinuous (90%) epitopes, while most of T cell epitopes are continuous. Antibodies or B-cell receptors recognize B-cell epitopes. Four approaches have been employed for epitope prediction: sequence-based methods, structure-based methods, hybrid methods, and consensus methods. Though the hybrid method provides more accurate results than others, consensus methods give the best prediction results (Yang and Yu, 2009).

Though some investigations appeared to be auspicious in preclinical and clinical trials, no vaccine have been approved for use. The objectives of this computational approach were to (1) investigate the recent advancement in vaccine development against ZIKV throughout the world, (2) identify highly conserved region in ZIKV genome as this virus confer high versatility amongst different strains and region.

The main aim of the study was to suggest at least one precise epitope for ZIKV vaccine development, which would be universal for all strains of this virus with immunogenic properties.

In current investigation, a B-cell epitope *DSRCPTQ* has been discovered from the sequence of ZIKV which is 100% conserved in all strains from all region and hosts of this virus. Several sequence analyses, such as Multiple sequence alignment (MSA), retrieval of the most consensus sequence for ZIKV, determination of the conserved region as well as structural validation such as epitope mapping on protein 3D structure were accomplished to predict highly immunogenic, surface accessible and conserved epitope. Various computer programs and online server were implemented for our study.

2 Materials and Methods.

2.1 Retrieval of Protein Sequences

Primary amino acid sequences of the complete genome for full polyprotein of Zika Virus were collected using Virus Variation tool (<https://www.ncbi.nlm.nih.gov/genome/viruses/variation>) of NCBI (Hatcher et al., 2017). Only full-length sequences have been selected and fully identical sequences were excluded. All other filter options (Host, Region/Country, Genome Region, Isolation Source) were set to default. Thus, a sum of 396 protein sequences has been collected which includes all available strains of Zika Virus in Virus Variation database from all over the world. Sequences with ambiguous or exceptional amino acid code (e.g., B, J, O, U, Z, X, -) causes error during epitope conservancy analysis. Thereupon, sequences containing such codes were excluded from the list. Finally, total 305 full-length protein sequences were chosen for further analysis.

2.2 Identification of Conserved Region

UGENE v1.31.1 (<http://ugene.net>) software package was employed for Multiple Sequence Alignment (MSA) (Okonechnikov et al., 2012). In this software, MSA was done by using ClustalO program (Sievers et al., 2011). From this Multiple sequence alignment, the most consensus sequence of full-length polyprotein of ZIKV was retrieved and protein sequence region of Capsid Protein (C), Membrane Glycoprotein Precursor (PrM), Membrane Glycoprotein (M), Envelope Protein (E), RNA-dependent RNA Polymerase (NS5) were checked for conserved segments (Table-1). Sequences with the highest number of identical and similar amino acid and no gap were selected as conserved regions. Total 12 conserved regions were picked to use for epitope prediction (Table-2).

2.3 Epitope Prediction:

B-cell epitope prediction tool of the Immune Epitope Database (IEDB) was implemented to predict linear epitopes from selected conserved protein sequences (<http://tools.iedb.org/main/bcell1>). Initially, the following three methods from this tool have been applied for linear epitope prediction.

1. Bepipred Linear Epitope Prediction 2.0: This method predicts B-cell epitopes from a protein sequence, using a Random Forest algorithm trained on epitopes and non-epitope amino acids determined from crystal structures (Jespersen et al., 2017). This method was used with the default threshold value of 0.5.

2. Kolaskar & Tongaonkar Antigenicity: For predicting epitopes based on antigenicity this method was employed with a threshold value of 1.0 (Kolaskar and Tongaonkar, 1990). Kolaskar & Tongaonkar Antigenicity scheme uses physicochemical properties of amino acid residues and their frequencies of occurrence in protein sequences to predict B-cell epitopes. The method can predict antigenic determinants with about 75% accuracy which is ever best-known method.

3. Emini Surface Accessibility Prediction: Then, Emini Surface Accessibility Prediction method from IEDB tool was utilized for predicting epitope based on surface accessibility scale with the threshold value of 0.8 (Emini et al., 1985).

A number of peptides from each conserved region were generated for each method. Only peptides with overlapping between three prediction methods were picked for further analysis.

2.3.1 Prediction of Epitope Conservancy:

To calculate the Conservancy of selected epitopes, the IEDB epitope conservancy analysis tool (<http://tools.iedb.org/conservancy>) was applied (Bui et al., 2007). This tool measures the degree of conservancy of an epitope within given protein sequences based on sequence identity.

2.3.2 Prediction of Hydrophilicity:

Parker Hydrophilicity Prediction method was employed to determine Hydrophilicity of conserved regions containing candidate epitopes (Parker et al., 1986). This method uses the average score as default threshold value for calculation. In this study default threshold value for every conserved sequence were kept unchanged.

2.4 Evaluation in Protein 3D Model:

Phyre2 server (Protein Homology/analogy Recognition Engine V 2.0) was used to identify the 3D structure of protein E (Kelley et al., 2015). Based on homology modeling by Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html>) chain A of protein 5IRE was found in PDB Database (<https://www.rcsb.org>) with 100% confidence and coverage (Sirohi et al., 2016). Ramachandran plot for modeled protein was done by employing Uppsala Ramachandran Server (<http://eds.bmc.uu.se/ramachan.html>) to analyze its the validity and quality (Kleywegt and Jones, 1996). Later on, the protein structure has been processed, and the conserved epitope was mapped on protein 3D structure by the use of Discovery Studio 4.5 Visualizer software (<http://www.accelrys.com>) (Dassault Systèmes BIOVIA, 2017).

Table 1: Positions of Proteins in Consensus Sequence from Multiple Sequence Alignment (MSA)

From the Region of Protein	Position in Consensus Sequence from MSA
Capsid Protein (C)	3-106
Membrane Glycoprotein Precursor (PrM)	125-292
Membrane Glycoprotein (M)	218-292
Envelope Protein (E)	293-792
RNA-dependent RNA Polymerase (NS5)	2523-3425

Table 2: Conserved Sequence from consensus protein identified by Multiple Sequence Alignment (MSA) in UGENE software

No.	Conserved Sequence	From the Region of Protein	Position in Consensus Sequence from MSA
S1	GGLKRLPAGLLLGHGPIRMVLAFLRFTA PSLGLINRWGVSVKKEAMEIHKFKKDLAAML RIINARKE	Capsid Protein (C)	30-102
S2	QIMDLGHMCDATMSYECPLMDEGVPPDDVDC WCNTTSTWVVYGTCHHKKGAEARRSRAVTL SHSTRKLQTRSQTWLESREYTKHLI	Membrane Glycoprotein Precursor (PrM)	161-247
S3	IRCIQVSNRDFVEGMSGGTWVDVLEHGCV TVMAQDKPTVDIELVTTTNSMAEVRSYCYE ASISDMASDSRCPTQGEAYLDKQSDTQYVCKR TLVDRGWGNGCGLFGKGLVTCAKFACSKKM TGKSIQPENLEYRIMLSVHGSQHSGMI	Envelope Protein (E)	293-444
S4	TPNSPRAEATLGGFGSLGLDCEPRTGLDFSDLY YLTMNNKHVLVHKEWFHDIPLPWHAGADTG TPHWNNKEALVEFKDAHAKRQTVVVLGSQEG AVHTALAGALEAEMDGAKG	Envelope Protein (E)	462-574
S5	QMAVDMQTLTPVGRLITANPVITESTENSKMM LELDPPFGDSYIVIGVG	Envelope Protein (E)	636-684
S6	GGGTGETLGEKWKARLNQMSALEFYSYKKS ITEVCREEARALKDGVATGGHAVSRGSAKLR WLVERGYLQPYGK	RNA-dependent RNA Polymerase (NS5)	2523-2598
S7	LVQSYGWNIVRLKSGVDVFMMAEPCDILL DIGESSSPEVEEARTLRVLSMVGDWLEKRP AFCIKVLCPYTSTMMET	RNA-dependent RNA Polymerase (NS5)	2637-2716
S8	NGVVRLLSKPVDVVTGVTGIAMTDTTPYGQ RVFKEKVDTRVPDPQEGTRQVM	RNA-dependent RNA Polymerase (NS5)	2845-2897
S9	SCVYNMMGKREKKQGEFGKAKGSRAIYMW LGARFLEFEALGFLNEDHWMGRENSGGGVEG LGLQRLGY	RNA-dependent RNA Polymerase (NS5)	2972-3040
S10	GKTVMDIHQDQRGSGQVVTYALNTFTNLVV QLIRNMEAEVLEMQDLWLLR	RNA-dependent RNA Polymerase (NS5)	3110-3162
S11	WLQSNWDRDKRMAVSGDDCVVKPIDRFA HALRFLNDMGKVRKDTQEWKPGSTGW	RNA-dependent RNA Polymerase (NS5)	3170-3224
S12	LKDGRSIVPCRHQDELIGRARVSPGAGWSIRE TACLAKSYAQMWQLLYFHRRDLRLMANAICS	RNA-dependent RNA Polymerase (NS5)	3242-3305

3 Results

3.1 Selection of Conserved Regions

Multiple sequence alignment of full-length polyprotein of Zika virus have been checked for conserved peptide segments in the region of Capsid Protein (C), Membrane Glycoprotein Precursor/Membrane Glycoprotein (PrM/M), Envelope Protein (E), RNA-dependent RNA Polymerase (NS5) (See, Supplementary Materials: Sp1 to Sp4). And thus, 1, 1, 3 and 7 conserved peptide segments were identified in the region of C, PrM, E, NS5 protein respectively (Table-1). Total 12 conserved regions have been considered for epitope analysis. In UGENE a most consensus protein sequence of Zika virus polyprotein was found after performing MSA and conserved sequences were taken from this consensus sequence.

3.2 Prediction and Selection of B-cell Epitopes

To be an effective epitope, a peptide must confer immunogenic, antigenic and surface accessible properties. A competent B-cell epitope must have the surface accessibility property to bind with an antibody effectively (Caoili, 2010). Three methods (Bepipred Linear Epitope Prediction 2.0, Kolaskar & Tongaonkar Antigenicity and Emini Surface Accessibility Prediction) from the IEDB database have been utilized to recommend linear B-cell epitope from all conserved regions.

For every conserved amino acid sequence, a list of predicted peptides was generated in each method. Peptides which were common in all three methods were chosen as candidate epitope for further analysis. No epitope was found from conserved sequences in the region of Capsid Protein (Conserved Sequence-S1) and Membrane Glycoprotein Precursor (Conserved Sequence-S2). In envelope protein region, one epitope was identified from Conserved Sequence 3 (S3). Another, two epitopes were found from Conserved Sequence 6 and 12 (S6 and S12) in the NS5 protein region. All predicted results are listed in supplementary material Sp5.

1. Envelope Protein (E): Epitopes predicted by Bepipred 2.0, Antigenicity, and Surface Accessibility methods from Conserved Sequence 3 (S3) of Envelope Protein (E) region are listed in Table-3. An antigenic epitope *DSRCPTQ* was overlapped with Bepipred predicted peptide and surface accessible peptide *MASDSRCPTQGEAYLDKQSDT*. As this peptide fulfilled all three preliminary criteria to be a potential epitope, it was considered as a candidate epitope.

2. RNA-dependent RNA Polymerase (NS5): In this protein region 7 conserved amino acid segments (S6 to S12) were extracted. Every segment was analyzed for B-cell epitope prediction and two potential epitopes have been found out

in S6 and S12. All predicted peptides from conserved region 6 from each method (Bepipred 2.0, Antigenicity, and Surface Accessibility) are summarized in Table-4. In this case, a peptide *FYSYKK* was found to be overlapped among other prediction results. So, this peptide was considered as a candidate B-cell epitope.

Lastly, another candidate epitope *RHQDE* has been identified from S12 region on NS5 protein. This peptide also found to be overlapped with all three predictions. All predicted epitopes from Conserved Sequence 12 are listed in Table-5.

3.2.1 Conservancy Analysis:

Epitope Conservancy Analysis tool of IEDB has been implemented to find out the percentage of conservancy of all candidate epitope (*DSRCPTQ*, *FYSYKK* and *RHQDE*). All epitopes demonstrated 100% conservancy amid all 305 sequences of Zika virus (Figure-1).

3.2.2 Hydrophilicity of Epitopes:

Potential epitopes containing conserve sequence S3, S6, S12 were tested for hydrophilicity property using Parker Hydrophilicity Prediction method. Hydrophilicity of these three conserved sequences is demonstrated in Figure-2. Average hydrophilicity score of each sequence was used as default threshold value. For S3 sequence default threshold value was 2.137 and residue score of predicted epitope *DSRCPTQ* was 5.057, that is much higher than default threshold value. Average hydrophilicity of S6 sequence was 2.005 but predicted epitope *FYSYKK* from this region had much lower residue score (0.817). Though predicted epitope *RHQDE* from sequence S12 showed higher residue score 6.02, threshold value of this conserved region was very low 0.738 and length of this epitope is also quite short to be a suitable epitope. Summary of Parker Hydrophilicity results for all three epitopes and corresponding conserved sequences are tabulated in Table-6.

Considering all properties and criteria, *DSRCPTQ* in Envelope protein sequence was chosen as final desired B-cell epitope.

3.3 Epitope Mapping on Protein 3D Structure:

According to the result from Ramachandran plot analysis 84.3% amino acid residues of ZIKV Envelope protein are in the acceptable region (Figure-3). The epitope *DSRCPTQ* was found to be located on the surface of the 3D structure of Envelope Protein after mapping (Figure-4).

Table 3: Bepipred Linear Epitope, Antigenic and Surface Accessible epitopes predicted from Conserved Region S3 of Envelope Protein (E)

Predicted Peptides	Length
Bepipred Linear Epitope Prediction 2.0	
GVSNRDFVEGMSGGTWVDVVLEQDTTTSNMAEVRSYCYEASISDMAS DSRCPTQ GEAYLDKQSDTQYVCKRT LVDRGWGNGCGLFGKGS�VTCAKFACSKMTGKSIQPENLEVRIMLSVHGSQHS	145
Kolaskar & Tongaonkar Antigenicity	
WVDVVLEHGGCVTVM	15
KPTVDIELVTTT	12
VRSYCYEAS	9
DSRCPTQ	7
TQYVCKRTLVDL	12
KGS�VTCAKFACSK	14
RIMLSVHGSQ	10
Emini Surface Accessibility Prediction	
MAQDKPTV	8
MAS DSRCPTQ GEAYLDKQSDT	21
KSIQPENLEVR	11

Table 4: Bepipred Linear Epitope, Antigenic and Surface Accessible epitopes predicted from Conserved Region S6 in RNA-dependent RNA Polymerase (NS5)

Predicted Peptides	Length
Bepipred Linear Epitope Prediction 2.0	
ETLGEKWKARLNQMSALE FYSYK KSGITEVCREE RRALKDGVATGGHAVSRGSAK YLQ	34 21 3
Kolaskar & Tongaonkar Antigenicity	
ALEFYSYK K	9
ITEVCRE	7
TGGHAVSRGS	10
Emini Surface Accessibility Prediction	
LGEKWKARLN	10
FYSYK KSG	8
EEARRAL	7

Table 5: Bepipred Linear Epitope, Antigenic and Surface Accessible epitopes predicted from Conserved Region S12 in RNA-dependent RNA Polymerase (NS5)

Predicted Peptides	Length
Bepipred Linear Epitope Prediction 2.0	
RSIVVPCRHQDELIGRARVSPGAGWSIRETACLAKSYAQMWDLRLMAN	41
	7
Kolaskar & Tongaonkar Antigenicity	
RSIVVPCRHQDE	12
RARVSPG	7
ETACLAKSY	9
MWQLLYFHR	9
Emini Surface Accessibility Prediction	
RHQDEL	6
FHRRDL	6

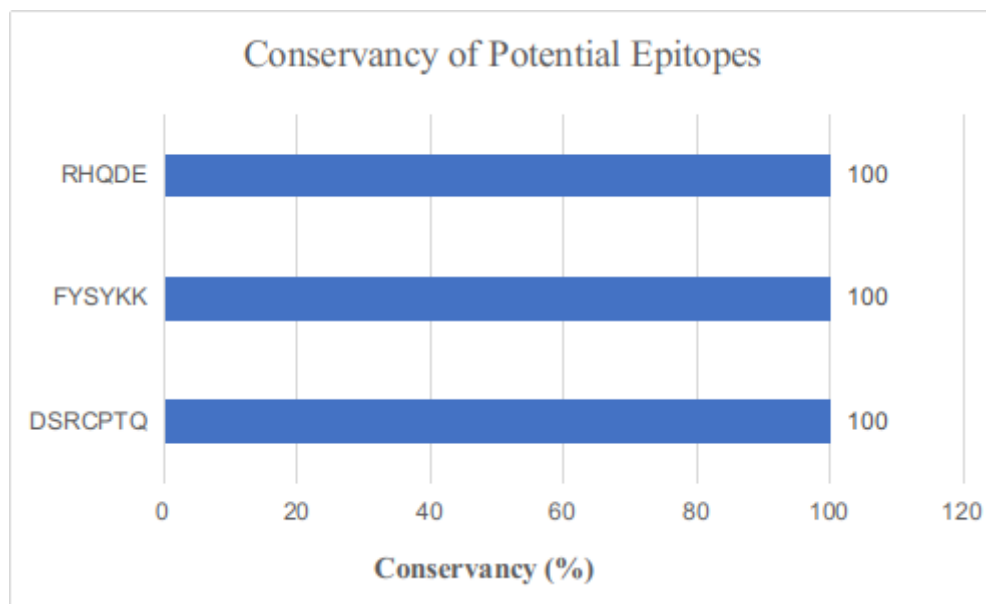


Fig. 1: Results of Conservancy Analysis for three predicted epitopes

Table 6: Parker Hydrophilicity results of candidate epitopes

Epitope	Conserved Sequence	Protein	Threshold Value of Conserved Sequence (Average Score)	Residue Score
DSRCTQ	S3	Envelope Protein (E)	2.137	5.057
FYSYKK	S6	RNA-dependent RNA Polymerase (NS5)	2.005	0.817
RHQDE	S12	RNA-dependent RNA Polymerase (NS5)	0.738	6.02

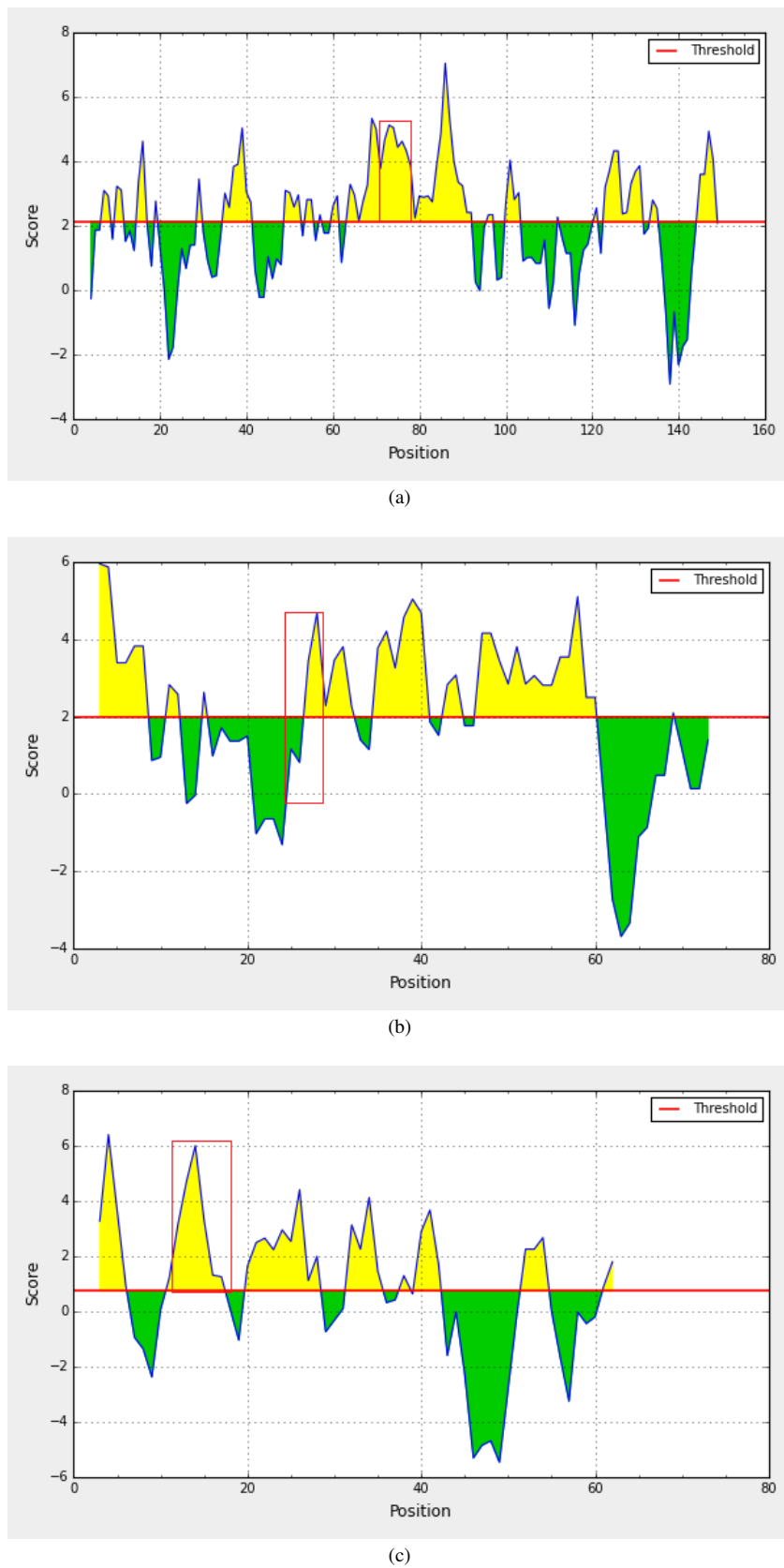


Fig. 2: Hydrophilicity plot of (a) conserved sequence S3 from Envelope protein, (b) S6 and (c) S12 from NS5 protein. In every graph yellow color represents hydrophilic regions. Red boxes represent the hydrophilicity of candidate epitope from corresponding conserved sequence

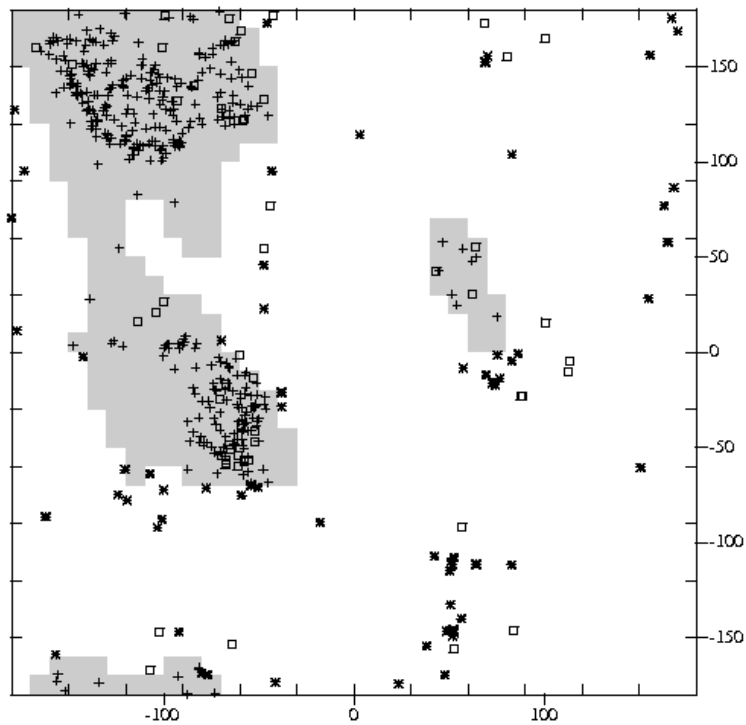


Fig. 3: Ramachandran plot for Envelope Protein by Uppsala Ramachandran Server

X = PHI mapped to [-180,180>
Y = PSI mapped to [-180,180>
PDB file : /pub/db/pub/pdb/data/structures/all/pdb/pdb5ire.ent
Glycines (open squares): 55 ; Start/end residues : 2
D-amino acids : 0 ; Residues with missing atoms : 0
Residues in Ramachandran plot checked : 444 out of 501
In core regions (plus signs): 376 ; Outliers (asterisks): 68
Percentage outliers: 15.3
An average <= 2.0 A model has ~0.5% outliers
See: Kleywegt, G.J. and Jones, T.A. (1996). Structure 4, 1395-1400.

Fig. 4: Epitope Mapping on E Protein 3D structure



4 Discussion

Zika virus epidemic has become a global threat in recent year. Over the past few decades, this virus has transmitted actively (local transmission) around different countries and territories (Countries, 2016). Zika virus uses a wide number of organisms as host including human, mosquito (mainly *Aedes* genus), monkeys and as well as other animals especially in several mammals (Olson et al., 1983; Wang et al., 2016). That's why full polyprotein sequences of all strains of Zika virus around the world from all listed hosts in virus variation were included in this study.

According to the World Health Organization vaccine for Zika virus should be inactivated and non-live as pregnant women are highly susceptible to this virus (Rasmussen et al., 2016). So, the peptide-based vaccine would be best effective preventive measure against this virus. Protein E and PrM coding plasmid DNA vaccine has been developed which will encode the outer protein coat of Zika virus virion (Dowd et al., 2016). Flavivirus envelope glycoproteins play a crucial role in the initiation of endocytosis into the host cell by binding with endosomal membrane (Dai et al., 2016). So, an antibody against envelope protein of Zika virus will be effective to prevent the entrance of virus into the host cell. Until July 2018, there is no appropriate vaccine for Zika virus has been approved for clinical use, though some vaccines are currently in the clinical trial (NIH, 2018). So, it is a crying need to construct a suitable preventive vaccine against this global threat.

in silico methods have been turning into popular and essential tools in vaccine model as it helps to curtail the time and the number of *in vitro* laboratory assays (Backert and Kohlbacher, 2015; Florea et al., 2003; Groot and Rappuoli, 2004). Now scientists are following various immunoinformatic tools as cost-efficient and rapid approaches to develop the vaccine against Zika virus, for example, antigenic epitope prediction from Zika proteins (Alam et al., 2016; Ashfaq and Ahmed, 2016; Badawi et al., 2016; Dar et al., 2016; Khan et al., 2015; Weltman, 2016). In this research work, a very potential epitope in the sequence of envelope protein was found which is universal for all the strains of Zika virus all over the world.

From MSA of full-length polyprotein of Zika Virus, several conserved regions (S1-S12) were selected in the coding sequence of Capsid Protein (C), Membrane Glycoprotein Precursor (PrM), Envelope Protein (E), RNA-dependent RNA Polymerase (NS5). The most consensus sequence of full-length polyprotein among all Zika Virus strains also been found from MSA result in UGENE software and conserved segments were taken from this consensus sequence. After analysis, each conserved segment in terms of Bepipred linear epitope prediction, antigenicity and surface accessibility three epitopes from conserved segment S3, S6 and S12

were identified, one from each segment (Table-6). Though all epitopes were 100% conserved in all sequence of Zika virus, epitope *DSRCPTQ* from conserved region S3 possesses better hydrophilicity property based on Parker Hydrophilicity Prediction method (Table-6). Additionally, the presence of S3 conserved segment in envelope protein sequence makes the epitope more acceptable.

Furthermore, after structural analysis by epitope mapping on Envelope protein 3D structure the epitope was located on the external surface of the protein, thereby it added further affirmation that *DSRCPTQ* epitope is endowed with all the compulsory properties to be a suitable B-cell epitope. Principle of virus detection through molecular immunoassay techniques is based on conserved epitopes of particular viral protein. Based on this principle Zika virus Envelope protein is of great interest for epitope detection among several researchers (Weltman, 2016).

As only computational approaches were implemented to identify this epitope, additional *in vivo* and *in vitro* evaluation is also required to determine the real efficacy, antigenicity and other immunogenic properties of the epitope. An additional adjuvant can be coupled with this peptide to enhance its immunogenicity and stability (Gershoni et al., 2007).

Defining the antigenic sites in viral protein is the fundamental basis for the generation of therapeutics and vaccine against viral diseases (Toyoda et al., 2000). By adopting sequence analysis and various computational prediction method *DSRCPTQ* peptide was identified from Zika virus envelope protein (E) as the best epitope to design new specific common antibody for all strains of this virus upon *in vitro* and *in vivo* test.

5 Conclusion

As ZIKV spreading throughout the world very rapidly and deadly, it is an immediate call for discovery of a standard universal vaccine, which will be effective against all strains of this virus. A number of researchers have been working tirelessly to develop preventive therapeutics and vaccine against this fatal virus and several remarkable outcomes also been noticed. This study was an effort to take this endeavor one step forward. A potential epitope has been identified in the envelope protein region of ZIKV which is conserved in all strains. To avert the further outbreak of ZIKV, vaccine, and therapeutics based on this suggested epitope can be an efficient arsenal. As it was an *in silico* approach consequential *in vitro* and *in vivo* validations are also necessary.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- (2014) Etymologia: Zika Virus. *Emerging infectious diseases* 20(6):1090
- (2018) Zika Virus Vaccines — NIH: National institute of allergy and infectious diseases. <https://www.niaid.nih.gov/diseases-conditions/zika-vaccines>, accessed: 02 Nov 2018
- Abbink P, Stephenson KE, Barouch DH (2018) Zika virus vaccines. *Nature Reviews Microbiology* 16(10):594–600
- Aid M, Abbink P, Larocca RA, Boyd M, Nityanandam R, Nanayakkara O, Martinot AJ, Moseley ET, Blass E, Borducchi EN, Chandrashekar A, Brinkman AL, Molloy K, Jetton D, Tartaglia LJ, Liu J, Best K, Perelson AS, De La Barrera RA, Lewis MG, Barouch DH (2017) Zika Virus Persistence in the Central Nervous System and Lymph Nodes of Rhesus Monkeys. *Cell* 169(4):610–620.e14
- Alam A, Ali S, Ahamad S, Malik MZ, Ishrat R (2016) From ZikV genome to vaccine: in silico approach for the epitope-based peptide vaccine against Zika virus envelope glycoprotein. *Immunology* 149(4):386–399
- Ashfaq UA, Ahmed B (2016) De Novo Structural Modeling and Conserved Epitopes Prediction of Zika Virus Envelop Protein for Vaccine Development. *Viral Immunology* 29(7):436–443
- Backert L, Kohlbacher O (2015) Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome Medicine* 7(1):119
- Badawi MM, Osman MM, Alla AAF, Ahmedani AM, hamed Abdalla M, Gasemelseed MM, Elsayed AA, Salih MA (2016) Highly Conserved Epitopes of ZIKA Envelope Glycoprotein May Act as a Novel Peptide Vaccine with High Coverage: Immunoinformatics Approach. *American Journal of Biomedical Research*, Vol 4, 2016, Pages 46-60 4(3):46–60
- Brasil P, Pereira JP, Moreira ME, Ribeiro Nogueira RM, Damasceno L, Wakimoto M, Rabello RS, Valderramos SG, Halai UA, Salles TS, Zin AA, Horovitz D, Daltro P, Boechat M, Raja Gabaglia C, Carvalho de Sequeira P, Pilotto JH, Medialdea-Carrera R, Cotrim da Cunha D, Abreu de Carvalho LM, Pone M, Machado Siqueira A, Calvet GA, Rodrigues Baião AE, Neves ES, Nassar de Carvalho PR, Hasue RH, Marschik PB, Einspieler C, Janzen C, Cherry JD, Bispo de Filippis AM, Nielsen-Saines K (2016) Zika Virus Infection in Pregnant Women in Rio de Janeiro. *New England Journal of Medicine* 375(24):2321–2334
- Braut AC, Domi A, McDonald EM, Talmi-Frank D, McCurley N, Basu R, Robinson HL, Hellerstein M, Duggal NK, Bowen RA, Guirakhoo F (2017) A Zika Vaccine Targeting NS1 Protein Protects Immunocompetent Adult Mice in a Lethal Challenge Model. *Scientific Reports* 7(1):14769
- Bui HH, Sidney J, Li W, Fusseder N, Sette A (2007) Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics* 8(1):361
- Cao-Lormeau VM, Blake A, Mons S, Lastère S, Roche C, Vanhomwegen J, Dub T, Baudouin L, Teissier A, Larre P, Vial AL, Decam C, Choumet V, Halstead SK, Willison HJ, Musset L, Manuguerra JC, Despres P, Fournier E, Mallet HP, Musso D, Fontanet A, Neil J, Ghawché F (2016) Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *The Lancet* 387(10027):1531–1539
- Caoili SEC (2010) B-cell epitope prediction for peptide-based vaccine design. In: *Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology - BCB '10*, ACM Press, New York, New York, USA, p 602
- Countries A (2016) Territories with Active Zika Virus Transmission [<https://www.cdc.gov/zika/geo/active-countries.html>]
- Cunha MS, Esposito DLA, Rocco IM, Maeda AY, Vasami FGS, Nogueira JS, de Souza RP, Suzuki A, Addas-Carvalho M, Barjas-Castro MdL, Resende MR, Stucchi RSB, Boin IdFSF, Katz G, Angerami RN, da Fonseca BAL (2016) First Complete Genome Sequence of Zika Virus (Flaviviridae, Flavivirus) from an Autochthonous Transmission in Brazil. *Genome announcements* 4(2):e00032–16
- Dai L, Song J, Lu X, Deng YQ, Musyoki AM, Cheng H, Zhang Y, Yuan Y, Song H, Haywood J, Xiao H, Yan J, Shi Y, Qin CF, Qi J, Gao GF (2016) Structures of the Zika Virus Envelope Protein and Its Complex with a Flavivirus Broadly Protective Antibody. *Cell Host & Microbe* 19(5):696–704
- Dar H, Zaheer T, Rehman MT, Ali A, Javed A, Khan GA, Babar MM, Waheed Y (2016) Prediction of promiscuous T-cell epitopes in the Zika virus polyprotein: An in silico approach. *Asian Pacific Journal of Tropical Medicine* 9(9):844–850
- Dassault Systèmes BIOVIA (2017) *Discovery Studio Visualizer 4.5*, San Diego: Dassault Systèmes 2016
- De Gregorio E, Rappuoli R (2012) Vaccines for the future: learning from human immunology. *Microbial Biotechnology* 5(2):149–155
- Dick G (1952) Zika virus (II). Pathogenicity and physical properties. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 46(5):521–534
- Dick G, Kitchen S, Haddock A (1952) Zika Virus (I). Isolations and serological specificity. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 46(5):509–520
- Dowd KA, Ko SY, Morabito KM, Yang ES, Pelc RS, DeMaso CR, Castilho LR, Abbink P, Boyd M, Nityanan-

- dam R, Gordon DN, Gallagher JR, Chen X, Todd JP, Tsybovsky Y, Harris A, Huang YJS, Higgs S, Vanlandingham DL, Andersen H, Lewis MG, De La Barrera R, Eckels KH, Jarman RG, Nason MC, Barouch DH, Roederer M, Kong WP, Mascola JR, Pierson TC, Graham BS (2016) Rapid development of a DNA vaccine for Zika virus. *Science (New York, NY)* 354(6309):237–240
- Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, Lanciotti RS, Pretrick M, Marfel M, Holzbauer S, Dubray C, Guillaumot L, Griggs A, Bel M, Lambert AJ, Laven J, Kosoy O, Panella A, Biggerstaff BJ, Fischer M, Hayes EB (2009) Zika Virus Outbreak on Yap Island, Federated States of Micronesia. *New England Journal of Medicine* 360(24):2536–2543
- Emini EA, Hughes JV, Perlow DS, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of virology* 55(3):836–9
- Faye O, Freire CCM, Iamarino A, Faye O, de Oliveira JVC, Diallo M, Zannotto PMA, Sall AA (2014) Molecular Evolution of Zika Virus during Its Emergence in the 20th Century. *PLoS Neglected Tropical Diseases* 8(1):e2636
- Fernandez E, Diamond MS (2017) Vaccination strategies against Zika virus. *Current Opinion in Virology* 23:59–67
- Ferraro B, Morrow MP, Hutnick NA, Shin TH, Lucke CE, Weiner DB (2011) Clinical Applications of DNA Vaccines: Current Progress. *Clinical Infectious Diseases* 53(3):296–302
- Florea L, Halldorsson B, Kohlbacher O, Schwartz R, Hoffman S, Istrail S (2003) Epitope prediction algorithms for peptide-based vaccine design. In: *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003, IEEE Comput. Soc, pp 17–26*
- Gaudinski MR, Houser KV, Morabito KM, Hu Z, Yamshchikov G, Rothwell RS, Berkowitz N, Mendoza F, Saunders JG, Novik L, Hendel CS, Holman LA, Gordon IJ, Cox JH, Edupuganti S, McArthur MA, Roupheal NG, Lyke KE, Cummings GE, Sitar S, Bailer RT, Foreman BM, Burgomaster K, Pelc RS, Gordon DN, DeMaso CR, Dowd KA, Laurencot C, Schwartz RM, Mascola JR, Graham BS, Pierson TC, Ledgerwood JE, Chen GL, Plummer S, Costner P, Zephir K, Casazza J, Ola A, Victorino M, Levinson C, Whalen W, Wang X, Cunningham J, Vasilenko O, Burgos Florez M, Hickman S, Pittman I, Le L, Larkin B, Andrews C, Apte P, Hicks R, Trelles Cartagena C, Williams P, Boyd CR, Conan-Cibotti M, Stein J, Kaltovich F, DeCederfelt H, McAdams S, Renehan P, Chen W, Greenberg N, Wymer N, Wadsworth L, Billington M, Robinson T, Boyce C, Pa’ahana Brown F, Chrisley L, Kwon A, Patel P, Kominou P, Dorsey B, Eddington S, Telscher S, Lee M, Mosely R, Ross A, Ford G, Domjahn B, Xu J, Beck A, Fineman R, Heeke S, Winter J, Nagar S, Kelley C, Mulligan M, Plummer S, Costner P, Zephir K, Casazza J, Ola A, Victorino M, Levinson C, Whalen W, Wang X, Cunningham J, Vasilenko O, Burgos Florez M, Hickman S, Pittman I, Le L, Larkin B, Andrews C, Apte P, Hicks R, Trelles Cartagena C, Williams P, Boyd CR, Conan-Cibotti M, Stein J, Kaltovich F, DeCederfelt H, McAdams S, Renehan P (2018) Safety, tolerability, and immunogenicity of two Zika virus DNA vaccine candidates in healthy adults: randomised, open-label, phase 1 clinical trials. *The Lancet* 391(10120):552–562
- Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Tarnovitski Freund N, Weiss Y (2007) Epitope Mapping. *BioDrugs* 21(3):145–156
- Groot ASD, Rappuoli R (2004) Genome-derived vaccines. *Expert Review of Vaccines* 3(1):59–76
- Hatcher EL, Zhdanov SA, Bao Y, Blinkova O, Nawrocki EP, Ostapchuck Y, Schäffer AA, Brister JR (2017) Virus Variation Resource improved response to emergent viral outbreaks. *Nucleic Acids Research* 45(D1):D482–D490
- Homan J, Malone RW, Darnell SJ, Bremel RD (2016) Antibody mediated epitope mimicry in the pathogenesis of Zika virus related disease. *bioRxiv* p 044834
- Jespersen MC, Peters B, Nielsen M, Marcatili P (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research* 45(W1):W24–W29
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10(6):845–858
- Khan MA, Hossain MU, Rakib-Uz-Zaman SM, Morshed MN (2015) Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study. *Scandinavian Journal of Immunology* 82(1):25–34
- Kleywegt GJ, Jones TA (1996) Phi/psi-chology: Ramachandran revisited. *Structure (London, England : 1993)* 4(12):1395–400
- Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Letters* 276(1-2):172–174
- Kostyuchenko VA, Lim EXY, Zhang S, Fibriansah G, Ng TS, Ooi JSG, Shi J, Lok SM (2016) Structure of the thermally stable Zika virus. *Nature* 533(7603):425–428
- Lanciotti RS, Kosoy OL, Laven JJ, Velez JO, Lambert AJ, Johnson AJ, Stanfield SM, Duffy MR (2008) Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerging infectious diseases* 14(8):1232–9
- Lin MY, Wang YL, Wu WL, Wolseley V, Tsai MT, Radic V, Thornton ME, Grubbs BH, Chow RH, Huang IC (2017) Zika Virus Infects Intermediate Progenitor Cells and Postmitotic Committed Neurons in Human Fetal Brain Tis-

- sues. *Scientific Reports* 7(1):14883
- Lindenbach BD, Rice CM (2003) Molecular biology of flaviviruses. *Advances in Virus Research* 59:23–61
- Malet H, Massé N, Selisko B, Romette JL, Alvarez K, Guillemot JC, Tolou H, Yap TL, Vasudevan SG, Lescar J, Canard B (2008) The flavivirus polymerase as a target for drug discovery. *Antiviral Research* 80(1):23–35
- Malone RW, Homan J, Callahan MV, Glasspool-Malone J, Damodaran L, Schneider ADB, Zimler R, Talton J, Cobb RR, Ruzic I, Smith-Gagen J, Janies D, Wilson J (2016) Zika Virus: Medical Countermeasure Development Challenges. *PLOS Neglected Tropical Diseases* 10(3):e0004530
- Mansuy JM, Dutertre M, Mengelle C, Fourcade C, Marchou B, Delobel P, Izopet J, Martin-Blondel G (2016) Zika virus: high infectious viral load in semen, a new sexually transmitted pathogen? *The Lancet Infectious Diseases* 16(4):405
- Mlakar J, Korva M, Tul N, Popović M, Poljšak-Prijatelj M, Mraz J, Kolenc M, Resman Rus K, Vesnaver Vipotnik T, Fabjan Vodušek V, Vizjak A, Pižem J, Petrovec M, Avšič Županc T (2016) Zika Virus Associated with Microcephaly. *New England Journal of Medicine* 374(10):951–958
- Modjarrad K, Lin L, George SL, Stephenson KE, Eckels KH, De La Barrera RA, Jarman RG, Sondergaard E, Tennant J, Ansel JL, Mills K, Koren M, Robb ML, Barrett J, Thompson J, Kosel AE, Dawson P, Hale A, Tan CS, Walsh SR, Meyer KE, Brien J, Crowell TA, Blazevic A, Mosby K, Larocca RA, Abbink P, Boyd M, Bricault CA, Seaman MS, Basil A, Walsh M, Tonwe V, Hoft DF, Thomas SJ, Barouch DH, Michael NL (2018) Preliminary aggregate safety and immunogenicity results from three trials of a purified inactivated Zika virus vaccine candidate: phase 1, randomised, double-blind, placebo-controlled clinical trials. *The Lancet* 391(10120):563–571
- Nowakowski TJ, Pollen AA, Di Lullo E, Sandoval-Espinosa C, Bershteyn M, Kriegstein AR (2016) Expression Analysis Highlights AXL as a Candidate Zika Virus Entry Receptor in Neural Stem Cells. *Cell Stem Cell* 18(5):591–596
- Okonechnikov K, Golosova O, Fursov M (2012) Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28(8):1166–1167
- Olson JG, Ksiazek TG, Gubler D, Lubis SI, Simanjuntak G, Lee VH, Nalim S, Juslis K, See R (1983) A survey for arboviral antibodies in sera of humans and animals in Lombok, Republic of Indonesia. *Annals of Tropical Medicine & Parasitology* 77(2):131–137
- Pardi N, Hogan MJ, Porter FW, Weissman D (2018) mRNA vaccines a new era in vaccinology. *Nature Reviews Drug Discovery* 17(4):261–279
- Parker JMR, Guo D, Hodges RS (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites. *Biochemistry* 25(19):5425–5432
- Patronov A, Doytchinova I (2013) T-cell epitope vaccine design by immunoinformatics. *Open Biology* 3(1):120139–120139
- Perera R, Khaliq M, Kuhn RJ (2008) Closing the door on flaviviruses: Entry as a target for antiviral drug design. *Antiviral Research* 80(1):11–22
- Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR (2016) Zika Virus and Birth Defects Reviewing the Evidence for Causality. *New England Journal of Medicine* 374(20):1981–1987
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* 7(1):539
- Sikka V, Chattu V, Popli R, Galwankar S, Kelkar D, Sawicki S, Stawicki S, Papadimos T (2016) The emergence of zika virus as a global health security threat: A review and a consensus statement of the INDUSEM Joint working Group (JWG). *Journal of Global Infectious Diseases* 8(1):3–15
- Sirohi D, Chen Z, Sun L, Klose T, Pierson T, Rossmann M, Kuhn R (2016) The 3.8 angstrom resolution cryo-EM structure of Zika virus. *Science* 352:467–470
- Tappe D, Nachtigall S, Kapaun A, Schnitzler P, Günther S, Schmidt-Chanasit J (2015) Acute Zika virus infection after travel to Malaysian Borneo, September 2014. *Emerging infectious diseases* 21(5):911–3
- Tebas P, Roberts CC, Muthumani K, Reuschel EL, Kudchodkar SB, Zaidi FI, White S, Khan AS, Racine T, Choi H, Boyer J, Park YK, Trottier S, Remigio C, Krieger D, Spruill SE, Bagarazzi M, Kobinger GP, Weiner DB, Maslow JN (2017) Safety and Immunogenicity of an AntiZika Virus DNA Vaccine Preliminary Report. *New England Journal of Medicine* p NEJMoa1708120
- Toyoda T, Masunaga K, Ohtus Y, Hara K, Hamada N, Kashiwagi T, Iwahashi J (2000) Antibody-Scanning and Epitope-Tagging Methods Molecular Mapping of Proteins Using Antibodies. *Current Protein & Peptide Science* 1(3):303–308
- Wang L, Valderramos SG, Wu A, Ouyang S, Li C, Brasil P, Bonaldo M, Coates T, Nielsen-Saines K, Jiang T, Aliyari R, Cheng G (2016) From Mosquitos to Humans: Genetic Evolution of Zika Virus. *Cell Host & Microbe* 19(5):561–565
- Weltman JK (2016) Computer-Assisted Vaccine Design by Analysis of Zika Virus E Proteins Obtained either from Humans or from Aedes Mosquitos. *Journal of Medical*

Microbiology & Diagnosis 5(3):1–6

Wikan N, Smith DR (2016) Zika virus: history of a newly emerging arbovirus. *The Lancet Infectious Diseases* 16(7):e119–e126

Xu K, Song Y, Dai L, Zhang Y, Lu X, Xie Y, Zhang H, Cheng T, Wang Q, Huang Q, Bi Y, Liu WJ, Liu W, Li X, Qin C, Shi Y, Yan J, Zhou D, Gao GF (2018) Recombinant Chimpanzee Adenovirus Vaccine AdC7-M/E Protects against Zika Virus Infection and Testis Damage. *Journal of virology* 92(6):JVI.01722–17

Yang X, Yu X (2009) An introduction to epitope prediction methods and software. *Reviews in Medical Virology* 19(2):77–96