HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage: http://ees.elsevier.com/apjtm

Original research http://dx.doi.org/10.1016/j.apjtm.2016.07.004

Prediction of promiscuous T-cell epitopes in the Zika virus polyprotein: An *in silico* approach

Hamza Dar[1], Tahreem Zaheer[1], Muhammad Talha Rehman[1], Amjad Ali[1], Aneela Javed[1*], Gohar Ayub Khan[1], Mustafeez Mujtaba Babar[2], Yasir Waheed[3*]

[1]*Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology, Islamabad 44000, Pakistan*

[2]*Shifa College of Pharmaceutical Sciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan*

[3]*Foundation University Medical College, Foundation University Islamabad, DHA-I, Islamabad 44000, Pakistan*

ARTICLE INFO

Article history:
Received 17 May 2016
Received in revised form 16 Jun 2016
Accepted 1 Jul 2016
Available online 26 Jul 2016

Keywords:
Zika virus
B-cell epitopes
T-cell epitopes
Vaccine
Antigenicity

ABSTRACT

Objective: To predict immunogenic promiscuous T cell epitopes from the polyprotein of the Zika virus using a range of bioinformatics tools. To date, no epitope data are available for the Zika virus in the IEDB database.

Methods: We retrieved nearly 54 full length polyprotein sequences of the Zika virus from the NCBI database belonging to different outbreaks. A consensus sequence was then used to predict the promiscuous T cell epitopes that bind MHC 1 and MHC II alleles using PorPred1 and ProPred immunoinformatic algorithms respectively. The antigenicity predicted score was also calculated for each predicted epitope using the VaxiJen 2.0 tool.

Results: By using ProPred1, 23 antigenic epitopes for HLA class I and 48 antigenic epitopes for HLA class II were predicted from the consensus polyprotein sequence of Zika virus. The greatest number of MHC class I binding epitopes were projected within the NS5 (21%), followed by Envelope (17%). For MHC class II, greatest number of predicted epitopes were in NS5 (19%) followed by the Envelope, NS1 and NS2 (17% each). A variety of epitopes with good binding affinity, promiscuity and antigenicity were predicted for both the HLA classes.

Conclusion: The predicted conserved promiscuous T-cell epitopes examined in this study were reported for the first time and will contribute to the imminent design of Zika virus vaccine candidates, which will be able to induce a broad range of immune responses in a heterogeneous HLA population. However, our results can be verified and employed in future efficacious vaccine formulations only after successful experimental studies.

1. Introduction

Zika virus is a single stranded RNA virus belonged to Flaviviridae family [1]. The genome of the virus is 10794 nucleotides long, which is translated into 3410 amino acids [2]. The large polypeptide chain that is encoded by long and single ORF is cleaved into: Envelope, a membrane precursor, a capsid and non-structural proteins including NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B, and NS5. The envelope protein of the virus is involved in the process of fusion of the virus with the receptor of host cells and is also involved in the replication cycle of the virus. The NS5 protein has two terminals: N terminus and C terminus, the N terminus has a role in protection of RNA while the C terminus encodes RNA dependant RPA activity [3].

During 1947–2006, more than twenty cases of Zika virus infection were reported, but the research on them was not given prime importance because of its geographical spread limited to the countries in Africa and South Asia, and mild clinical signs and symptoms of the Zika virus infection [4]. After 2006, a sudden outbreak of Zika virus was reported in 2007 in the Yap Island, where 73% of the population was infected with Zika virus [5]. In 2013 a major outbreak of Zika virus was reported in the French Polynesia [6]. The infectious Zika virus then started spreading into the other islands of Pacific Ocean and in 2014 it arrived in Chile and Eastern island of Western Hemisphere [7] and in Latin America probably due to infected travellers. The virus is a mosquito borne virus, and mosquito

*Corresponding authors. Yasir Waheed, PhD, Foundation University Medical College, Foundation University Islamabad, DHA-I, Islamabad 44000, Pakistan; Aneela Javed, PhD, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology, Islamabad 44000, Pakistan.
E-mails: javedaneela@gmail.com, yasir_waheed_199@hotmail.com
Peer review under the responsibility of Hainan Medical College.

plays a key role in the transmission of the Zika virus infection in humans, which is the primary host of Zika virus. The transmission of Zika virus is carried out by *Aedes* species that includes *Aedes albopictus (A. albopictus), Aedes aegypt (A. aegypt)* [8], *Aedes luteocephalus (A. luteocephalus), Aedes furcifer (A. furcifer), Aedes taylori (A. taylori), Aedes africanus (A. africanus),* and monkeys (Rhesus Macaques) [9]. The studies on the transmission of Zika virus show that the virus can be transmitted through sexual contact [10] due to its extended persistence in the semen [11], and also through blood transfusion [12]. Viral load is greater than other arboviruses and commences about ten days before the clinical manifestation of the disease [13]. The acute symptoms of Zika virus infections are arthralgia, maculopapular rash, myalgia, conjunctivitis, emesis, retro-orbital pain and headache; however, 80% of the patients are asymptomatic during the initial stages of infection. Recent reports about the outbreak of Zika virus in Brazil are linked to microcephaly and Guillain–Barre syndrome [14]. This association poses serious teratogenic and neuropathic risks to the health of fetus. A fatal Zika virus infection has also been reported that shows increased risk of disease and mortality in individual having compromised immune system [15].

The infection of Zika virus is fatal and can cause serious health threatening issues, so an antiviral vaccine or antiviral therapy needs to be designed in order to control the disease state. Antiviral therapies need to be designed by targeting enzymes that are involved in post translational packaging of viral protein [16] or by targeting enzymes that are essential for the replication of virus [17]. Development of vaccine for the treatment of Zika virus is extremely important in current situation as the virus has caused a great number of deaths in Brazil and is spreading in the other parts of world. Currently there is no prophylactic or therapeutic vaccine available in the market to curtail this infection.

Though the development of live attenuated YFV vaccine was a milestone but with the new advancements, epitope-based vaccine are gaining more importance, as the live attenuated vaccine may prove fatal in immunocompromised patients [18].

Advances in immunoinformatics research found that many conservative and highly immunogenic T/B cell epitopes (antigenic determinants that are recognized by host immune cells and can elicit both a humoral and cellular immune response) on the virus antigen could be used as potential vaccine targets. These epitopes can induce a protective immune response against a wide range of pathogenic microorganisms. After the artificial T-cell epitope is presented via the appropriate MHC molecule on the surface of the target cell to its corresponding T-cell, the epitope is recognized by T cells through TCR recognition, thereby activating the T-cell to proliferate and generate an appropriate immune response. Based on this scenario, the use of different pathogenic microorganisms and their corresponding T cell epitopes can be used to develop a $CD4^+$ T cell epitope vaccine (mostly for exogenous antigens that are degraded in the APCs after phagocytosis, thereafter binding to MHC II molecules, and finally presentation to $CD4^+$ T cells) or a $CD8^+$ T cell epitope vaccine (mostly for endogenous antigen that are digested following uptake by the APCs, and subsequent presentation to $CD8^+$ T cells via MHC-I molecules) [19].

Epitope vaccine or an epitope based subunit-vaccine has lesser side effects when compared to conventional vaccines, is easier to produce, is cheaper to manufacture, is easier to get rid of in the *in vitro* restriction cultures when compared to engineered subunit vaccines, does not contain any complete component of the pathogens, allows for the *in vitro* incorporation of sugar analogs which is difficult to achieve through engineered subunit vaccines and also takes less time to produce along with improved stability, specificity and sustainability [20]. However, due to the highly polymorphic nature of the HLA genes in the human population, the epitope specific HLA restricted vaccine is not normally expected to cause an immune response in all individuals within a given population. Thus, there is a need for the development of promiscuous epitopes that can bind to multiple HLA alleles within a heterogeneous population thereby catering to the need of a wide range of individuals [21].

The present study targets the near full length polyprotein of the Zika virus containing key structural and non-structural proteins, for prediction of promiscuous and antigenic epitopes using a range of online tools for the development of a safe and effective epitope based subunit vaccine.

2. Materials and methods

2.1. Sequence retrieval

54 Zika virus polyprotein sequences derived from 54 different genomes were retrieved from the NIAID Virus Pathogen Database and Analysis Resource (ViPR) through the web site at [22] as shown in S1 Table. The sequences were aligned and consensus sequence was generated using the multiple sequence alignment tool, Jalview [23].

2.2. Prediction of T cell epitopes

To determine the T cell epitopes, both HLA I and HLA II binding peptide sequences were required. ProPred I (www.imtech.res.in/raghava/ProPred1/) [24]was used to predict the HLA class I binding promiscuous epitopes in the consensus sequence. 4% default threshold value was selected and proteasome and immunoproteasome filters were enabled at 5% threshold value to maximize the efficiency of finding T cell epitopes. ProPred I determines epitopes that can bind to 47 HLA class I alleles. To predict epitopes for HLA class II alleles, ProPred [25] was used at a cut off value of 3% threshold. ProPred allows the prediction of antigenic epitopes for 51 HLA class II alleles.

2.3. Antigenic prediction

All the promiscuous T cell epitopes obtained from ProPred and ProPred1 tools were analysed for their antigenic properties using VaxiJen version 2.0 at [26]. Threshold value of 0.5 antigenic score was kept to filer probable non-antigenic sequences. Moreover, 87% accurate results are obtained for viruses at this default threshold. Vaxijen server performs alignment-independent prediction of protective antigens on the basis of their physicochemical properties.

2.4. Class I immunogenicity prediction

All the HLA 1 binding antigenic epitopes were scanned for MHC 1 immunogenicity using IEDB Analysis tool [27]. Default

parameters were selected to perform the immunogenicity prediction. The tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a peptide MHC complex.

2.5. Validation of predicted epitopes

The epitopes predicted in our study were submitted to IEDB database to check if they had been tested earlier. The IEDB database contains experimentally confirmed data characterizing antibody and T cell epitopes studied in homo sapiens, NHPs, and other animal species.

3. Results

3.1. Prediction of HLA I binding epitopes

By using ProPred1, 23 antigenic epitopes for HLA class I were predicted in the consensus polyprotein sequence of Zika virus (Table 1). Epitopes were highly conserved across the different strains of Zika virus. Notably, the promiscuous epitope NS5$^{2592–2600}$ was found to bind to 26 of the 47 HLA class I alleles.

Epitope C$^{25–34}$ was predicted to bind to 18 HLA class I alleles. However, its antigenicity score was 0.2434. Hence, that epitope was not considered for further immunogenicity analysis. Epitope prM$^{211–219}$ was predicted to bind to 7 alleles and its antigenicity score was significant i.e. 1.0359.

The overall antigenic prediction of epitope E$^{777–785}$ was the highest among the envelope epitope sequences. However, it was predicted to bind to only 7 HLA class I alleles. Epitopes E$^{706–714}$ and E$^{757–765}$ were predicted to bind 12 alleles. However, their antigenicity scores were below 0.5, indicating that they were non-antigenic sequences.

NS1$^{897–905}$ was predicted to bind 18 alleles. However, its antigenicity score was determined to be significantly low i.e. −0.4375. Similarly, NS1$^{1017–1025}$ was predicted to bind 14 alleles but its antigenicity was −0.2001. NS1$^{1059–1067}$ was found to bind to 13 alleles and its antigenicity was significant i.e. 0.8225. Moreover, this epitope was found to be conserved among all the polyprotein sequences in our study. Hence, NS1$^{1059–1067}$ was the most suitable epitope identified in the NS1 of polyprotein.

Similarly, NS2A$^{1339–1347}$, NS2A$^{1355–1363}$ and NS2A$^{1156–1164}$ were identified as the most significant epitopes in the NS2A. NS2B$^{1402–1410}$ and NS2B$^{1381–1389}$ were predicted to bind 14 and 10 HLA class I alleles respectively and their antigenicity score was found to be significant i.e. 0.85.

Epitopes NS3$^{2018–2026}$, NS3$^{1622–1630}$ and NS3$^{1688–1696}$ were predicted to bind 20, 9 and 6 number of alleles and their antigenicity scores were found out to be 1.1351, 1.6507 and 1.5064. Remarkably, all these three epitopes were found to be completely conserved among all the polyprotein sequences included in our study.

Similarly, NS4A$^{2199–2207}$, NS4A$^{2177–2185}$ and NS4A$^{2220–2238}$ were predicted to bind 7, 9 and 8 number of HLA class I alleles and their antigenicity scores were found to be 1.9567, 1.2576 and 0.8174, respectively. However, the immunogenicity score of epitope NS4A$^{2177–2185}$ was found to be significantly low i.e. −0.00704. Interestingly, the other two epitopes were found to be completely conserved among all the polyprotein sequences included in the study, indicating that they may have potential evolutionary significance.

NS4B$^{2459–2467}$ was predicted to bind 10 HLA class I alleles and its antigenicity score was significant i.e. 0.7985.

Epitopes NS5$^{2592–2600}$ and NS5$^{2520–2528}$ were predicted to bind 26 and 12 number of alleles respectively and their antigenicity scores were determined to be 0.8097 and 0.5710 respectively. However, the immunogenicity score of NS5$^{2592–2600}$ was found to be significantly low i.e. −0.01296. Epitopes NS5$^{3065–3073}$, NS5$^{2638–2644}$ and NS5$^{2898–2906}$ were predicted to bind 7, 8 and 9 alleles respectively and their antigenic scores

Table 1

T-cell class I MHC-specific predicted epitopes of the Zika virus polyprotein and their number of alleles, antigenicity prediction score and immunogenicity score.

Part of polyprotein	aa position	Epitope sequence	No. of alleles	Antigenicity score	Immunogenicity score
prM	211	RRSRRAVTL	7	1.0359	0.14313
Envelope protein	307	GGTWVDVVL	10	0.5728	0.35553
	462	NSPRAEATL	10	0.7956	0.23653
	640	QTLTPVGRL	11	1.0334	0.12236
	777	MCLALGGVL	7	1.3683	0.10959
NS1	1059	KGPWHSEEL	13	0.8225	0.23806
NS2A	1156	SLGVLVILL	13	0.5583	0.18644
	1339	KNLPFVMAL	12	1.3984	0.01276
	1355	VDPINVVGL	11	0.9097	0.21752
NS2B	1381	TAVGLICAL	10	0.8523	0.13934
	1402	GPMAAVGLL	14	0.8531	0.08145
NS3	1622	DGDIGAVAL	9	1.6507	0.26865
	1688	KKQLTVLDL	6	1.5064	0.0315
	2018	AAIEGEFKL	20	1.1351	0.244
NS4A	2177	LGLLGTVSL	9	1.2576	−0.00704
	2199	KMGFGMVTL	7	1.9567	0.05402
	2220	EPARIACVLIVVFLLLVVL	8	0.8174	0.5603
NS4B	2459	WGWGEAGAL	10	0.7985	0.29179
NS5	2520	RGGGTGETL	12	0.571	0.22198
	2592	QPYGKVIDL	26	0.8097	−0.01296
	2638	SYGWNIVRL	8	1.1058	0.41795
	2898	VSSWLWKEL	9	0.6107	0.2434
	3065	RFDLENEAL	7	1.2378	0.19481

were found to be 1.2378, 1.1058 and 0.6107 respectively. Notably, the immunogenic score of epitope NS5$^{2638–2644}$ was found to be significant as compared to other epitopes i.e. 0.41795.

3.2. Prediction of HLA II binding epitopes

By using ProPred, 48 antigenic epitopes for HLA class II were predicted in the consensus polyprotein sequence of Zika virus (Table 2). Epitopes were found to be conserved among the sequences to a large extent.

Epitope C$^{91–100}$ was predicted to bind to 9 HLA class II alleles and its antigenicity score was found to be significantly high i.e. 0.6639.

It was predicted that prM$^{158–166}$ will bind with 22 HLA class II alleles and its antigenicity score was found to be high i.e. 1.6065. Epitope prM$^{158–166}$ has been identified as an important epitope in the prM region of the consensus sequence.

Seven important promiscuous HLA class II binding epitopes were detected in the envelope of the consensus sequence. E$^{403–411}$ was predicted to bind 25 HLA II alleles and its antigenic score was found to be 1.2397. It was predicted that E$^{764–772}$ will bind 10 alleles but its antigenic score was significantly high as compared to E$^{403–411}$ i.e. 2.6277.

NS1$^{1004–1012}$ was predicted to bind 21 HLA class II alleles and its antigenic score was found to be 1.0216. NS1$^{1124–1132}$ and NS1$^{965–973}$ were predicted to bind 12 and 11 alleles but their antigenic score were found to be 2.4357 and 1.6908, respectively.

Seven important promiscuous HLA II binding epitopes were detected in the NS2A of the consensus sequence. It was predicted that NS2A$^{1157–1169}$ will bind 25 HLA II alleles and its

Table 2

T-cell class II MHC-specific predicted epitopes of the Zika virus polyprotein and their number of alleles and antigenicity prediction score.

Part of polyprotein	aa position	Epitope sequence	No. of alleles	Antigenicity score
Protein C	91	MLRIINARKE	9	0.6639
prM	158	IQIMDLGHM	22	1.6065
Envelope protein	403	LVTCAKFAC	25	1.2397
	470	LGGFGSLGL	11	2.0219
	545	VVVLGSQEG	17	1.1365
	588	LRLKGVSYS	19	1.5390
	676	YIVIGVGEK	13	2.0591
	764	WLGLNTKNG	10	2.6277
	765	LGLNTKNGS	14	2.0667
NS1	878	VQLTVVVGS	17	0.7907
	878	VQLTVVVGSVKNPM	13	0.5533
	965	VREDYSLEC	11	1.6908
	982	VKGKEAVHS	16	0.9250
	1 004	WRLKRAHLI	21	1.0216
	1 045	LSHHNTREG	10	1.0129
	1 124	WYGMEIRPR	12	2.4357
NS2A	1 157	LGVLVILLMVQEG	25	0.6307
	1 178	IIISTSMAV	17	0.5011
	1 241	FRANWTPRE	11	2.7633
	1 282	WLAIRAMVVPRT	11	0.8872
	1 285	IRAMVVPRT	14	0.9024
	1 326	FMLLSLKGK	10	2.0071
	1 343	FVMALGLTAVRLVDPINVVGLLLLTRSGK	10	1.1286
Serine protease NS2B	1 383	VGLICALAG	14	0.9110
	1 411	IVSYVVSGK	11	1.0866
Serine protease NS3	1 597	VQLLAVPPG	34	0.5392
	1 663	YVSAITQGR	13	1.1295
	1 722	VILAPTRVV	18	0.6400
	1 762	LMCHATFTS	25	0.6299
	2 034	FVELMKRGD	10	1.1065
	2 046	WLAYQVASA	13	0.6893
NS4A	2 174	IMLLGLLGT	17	1.0058
	2 185	LGIFFVLMRNKGIGKM	23	0.7587
	2 202	FGMVTLGAS	13	1.0230
	2 229	IVVFLLLVVLIP	12	0.5274
NS4B	2 331	YNNYSLMAM	14	0.9352
	2 364	LLMIGCYSQ	22	1.0663
	2 442	VLLIAVAVSS	26	0.5388
	2 496	FRGSYLAGA	14	0.6322
NS5	2 536	LNQMSALEF	18	1.0863
	2 701	IKVLCPYTS	26	0.7501
	2 750	IKSVSTTSQ	13	0.8471
	2 924	VRSNAALGA	28	1.1428
	2 973	YNMMGKREK	23	1.0356
	3 159	LRRSEKVTN	13	1.0871
	3 238	LHLKDGRSI	13	1.2828
	3 286	LLYFHRRDLRLMANAICSS	12	0.7456
	3 372	LIGHRPRTT	12	1.5275

antigenic score is 0.6307. NS2A$^{1241–1249}$ and NS2A$^{1326–1334}$ were predicted to bind 11 and 10 alleles but their antigenic score were found to be 2.7633 and 2.0071, comparatively. Out of the seven promiscuous epitopes identified in NS2A region, none was completely conserved among all the polyprotein sequences.

NS2B$^{1383–1391}$ and NS2B$^{1411–1419}$ were predicted to bind 14 and 12 HLA class II alleles but their antigenic score was found to be 0.9110 and 1.0866 respectively. Moreover, both epitopes were completely conserved among all the polyprotein sequences included in the study.

It was predicted that NS3$^{1597–1605}$ and NS3$^{1762–1770}$ will bind 34 and 25 HLA class II alleles and their antigenic scores were determined to be 0.5392 and 0.6299 respectively. Moreover, both NS3$^{1597–1605}$ and NS3$^{1762–1770}$ were conserved in all the 54 polyprotein sequences of our study. NS3$^{1663–1671}$ and NS3$^{2034–2042}$ were predicted to bind 13 and 10 HLA II alleles but their antigenic scores were found to be 1.1295 and 1.1065 respectively.

NS4A$^{2185–2200}$ was predicted to bind 23 alleles while its antigenic score was found to be 0.7587. NS4A$^{2174–2182}$ was predicted to bind to 17 HLA class II alleles while its antigenic score was found to be 1.0058. Moreover, this epitope was conserved in all the polyproteins included in the study.

NS4B$^{2364–2372}$ was predicted to bind 22 HLA class II alleles and its antigenic score was found to be 1.0663. NS4B$^{2442–2451}$ was predicted to bind 26 alleles but its antigenic score was less comparatively i.e. 0.5388.

NS5$^{2924–2932}$ and NS5$^{2701–2709}$ were predicted to bind 28 and 26 HLA class II alleles and their antigenic scores were determined to be 1.1428 and 0.7501, respectively. In comparison, NS5$^{3372–3380}$ and NS5$^{3238–3246}$ were predicted to bind 12 and 13 alleles while their antigenic scores were found to be 1.5275 and 1.2828, respectively. Remarkably, NS5$^{2924–2932}$ epitope was found to be conserved among all the 54 polyprotein sequences included in our study.

Moreover it was concluded that none of the epitopes predicted in this study have been studied previously.

4. Discussion

Zika virus is mainly spread by the *Aedes aegypti* mosquito, and has been lately making its presence known throughout Central America and Latin America, but there are chances that it might spread to tropical and subtropical regions too [28]. Currently there is no FDA approved vaccine. Even more, there is no specific treatment apart from the recommended use of aspirin and acetaminophen to counteract the fever and muscle pain and preventive measures against mosquito bites [29].

The Genus Flavivirus consists of diverse and complex group of pathogens which are antigenically related. The genomes of these viruses comprise of total 10 proteins and the role of each of the protein in viral pathogenesis is not yet completely elucidated. Effective immunization treatment for some members exists while due to some immunobiology complexities, vaccines for other members are still to be made. Most of animal models are either immune to flavivirus or they do not completely represent all manifestations of disease. Some human data of Flavivirus do exist; however, it does not represent all forms of disease and its global variation in populations [30].

Epitope based vaccines are already showing hopeful results. This promising vaccine technology has allowed for the prevention and treatment of cancer, viral, bacterial and other diseases [31–36]. Numerous immune-bioinformatics methods and tools have now been developed to assist in the search for T-cell MHC binding epitopes. Design and development of a vaccine using T cell specific epitopes is considerably more favourable because they evoke a long-term immune response and dodge antigenic drift whereas antigen can effortlessly evade the antibody memory response [21]. Both the CD4+ and CD8+ T cells have essential role in antiviral immune response as well as clonal expansion of B cell. In this study, we used the full-length polyprotein sequence of the Zika virus in order to increase the coverage of the genome and to search for promiscuous epitopes in both the structural and non-structural proteins of this virus. At the time of writing, no Zika virus T or B-cell epitopes have been uploaded to the Immune Epitope Database Analysis Resource (IEDB); a manually curated repository of experimentally characterized immune epitopes. The IEDB can be used by scientists to help in the identification, characterization, mapping and evaluation of likely targets for vaccine, therapeutic and diagnostic nominees, and moreover to give us a broader knowledge of the pathogenesis and immunobiology of any new disease or epidemic. The current study is a first attempt which intended to screen novel and highly probable immunogenic epitopes for T cells across all the major proteins of ZIKV. Furthermore, these crucial data can be unified with data from supplementary databases (Pharmacogenomics, genomic, proteomic, or genomic), in doing so increasing the usefulness and wide-ranging scope of the analyses.

In our study, a greater number of epitopes were projected for MHC class II when compared to MHC class I. The results of this study are in conformity with a meta-analysis study that enumerated a greater number of Class II epitopes in the *Flavivirus* genus [37]. Out of 23 identified MHC 1 binding antigenic epitopes, 12 epitopes i.e. E$^{462–470}$, E$^{640–648}$, NS1$^{1059–1067}$, NS2A$^{1156–1164}$, NS2A$^{1355–1363}$, NS2B$^{1381–1389}$, NS3$^{1622–1630}$, NS3$^{1688–1696}$, NS3$^{2018–2026}$, NS4A$^{2177–2185}$, NS4A$^{2220–2238}$ and NS5$^{2520–2528}$, were completely conserved in intact form among all the polyprotein sequences included in the study. Epitope E$^{640–648}$ was predicted to bind to 11 HLA class I alleles and its antigenicity score was significant i.e. 1.0334. Moreover, its sequence was completely conserved in all the polyprotein sequences included in our study. Hence, E$^{640–648}$ was identified as the best envelope HLA 1 epitope in our study. Epitopes NS3$^{2018–2026}$ was predicted to bind 20 HLA Class 1 alleles and its antigenicity score were found out to be 1.1351. Amazingly, it was found to be completely conserved among all the polyprotein sequences included in our study, indicating that these epitopes can be important for developing universally applicable vaccines. Notably, the immunogenicity score of NS4A$^{2220–2238}$ was found to be the highest (i.e. 0.5603) among all the HLA class I binding epitopes determined in our study.

Out of 48 predicted HLA II epitopes, 18 epitopes were found to be completely conserved in all of the polyprotein sequences included in the study. These epitopes are C$^{92–99}$, E$^{470–478}$, E$^{588–596}$, NS1$^{878–886}$, NS1$^{878–891}$, NS1$^{1124–1132}$, NS2B$^{1383–1391}$, NS2B$^{1411–1419}$, NS3$^{1597–1605}$, NS3$^{1762–1770}$, NS3$^{2046–2054}$, NS4A$^{2174–2182}$, NS4A$^{2229–2240}$, NS4B$^{2331–2339}$, NS4B$^{2496–2504}$, NS5$^{2536–2544}$, NS5$^{2750–2758}$ and NS5$^{2924–2932}$. Notably, the antigenic epitope NS3$^{1597–1605}$ has been predicted to bind 34 out of 51 HLA class II alleles. Epitope C$^{92–99}$ was predicted to bind to 25 HLA class II alleles and its antigenicity score was found to be significantly high i.e. 1.8011. Moreover, as discussed, this

epitope was found to be completely conserved among all the polyprotein sequences included in our study. This indicates that this epitope can serve as an important part of universally applicable vaccines. $NS1^{1124–1132}$ was completely conserved among all the polyprotein sequences, indicating that this epitope can be useful for developing a universally applicable vaccine, especially considering its significant antigenic prediction.

Structures of all ten of flavivirus genus viral proteins are reported, though complete data of all the 10 proteins for a single virus have not been reported yet. The best overall epitope distribution is available for WNV and DENV (ten out of ten for both) and the highest number of epitopes for the whole genus have been obtained from NS3 and E proteins. The human epitope data collected from the patients of Japanese encephalitis, dengue hemorrhagic fever, dengue fever, yellow fever and West Nile fever indicates the presence of both B-cell and T-cell epitopes. [37].

According to a 2010 Meta-analysis study of all immune data in the Flavivirus genus [37], 1 200 epitopes were identified in that study and most of the epitopes belonged to the dengue virus group followed by WNV and YFW. The higher percentage of epitopes identified for dengue virus, WNV and YFW indicate their worldwide impact on mortality and morbidity in human population while smaller number of epitopes recognized for other viruses indicates the presence of established prophylaxes or their less dreadful impact on human population. All the epitopes reported up to date are peptidic in nature and there is objectively even scattering of B-cell and T-cell epitopes in the genus as a whole. T-cell epitopes have been recognized in six out of nine flavivirus and the largest numbers of T-cell epitopes reported are in WNV DENV and YFW. Both CD4 and CD8 epitopes are defined for flavivirus but it was observed that DENV viruses have predominantly CD8+ T-cell epitopes while other viruses (WNV and JEV) mostly have CD4 T-cell epitopes.

Not surprisingly the data on host distribution of epitope reactivity's indicates that most of the flavivirus epitopes are defined in either humans or mice. A large number of DENV epitopes were defined in humans, as expected but surprisingly very low numbers of epitopes for WNV, YFW and complete absence of epitopes for JEV. Speculation is that low number of epitopes for JEV is due to availability of JEV vaccine. Identification of epitopes in NHP still remains of great interest despite the fact that they are quite expensive and have limited availability but they can be used as natural hosts and have biological and immunological similarities with humans [37,38].

Due to lack of a suitable animal model, very small numbers of protective epitopes are reported for flavivirus: DENV, JEV, and TBEV [37–41]. Many animal models are used to study diverse characteristics of flavivirus infection but, the standard model used is murine model. Although mice natural resistance to infection caused by certain flavivirus is problematic as it causes hindrance in measuring protective immunity of animal, Humanized or susceptible mouse models are being developed which can mimic disease symptoms more closely related to humans. However, until then we mostly have to rely on extrapolation of clinical studies [37,42].

A large number of epitopes (both larger and smaller) are identified in humans for the period of the natural course of infection for Dengue, West Nile, and Japanese Encephalitis viruses respectively. However, the numbers of epitopes were higher for DENV [37]. The contemporary data available is, however, inadequate and cannot provide a solution to the questions related to the immunopathological aspects of these viruses. However, we can extrapolate the epitope findings of other members of the *Flavivirus* genus to the ZIKV [43,44].

One downside of our study is the lack of *in vitro* and *in vivo* studies to test whether these peptidic epitopes will elicit a strong and protective immune response in humans. Since these epitopes were predicted using an in-silico approach, experimental studies are a must before such epitopes are used in vaccine formulations.

Conflict of interest statement

The author declares that there is no conflict of interest.

References

[1] Calisher CH, Gould EA. Taxonomy of the virus family Flaviviridae. *Adv Virus Res* 2003; **59**(59): 1-20.

[2] Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V. Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLoS Negl Trop Dis* 2012; **6**(2): e1477.

[3] Kuno G, Chang GJ. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. *Arch Virol* 2007; **152**(4): 687-696.

[4] Hayes EB. Zika virus outside Africa. *Emerg Infect Dis* 2009; **15**(9): 1347-1350.

[5] Ioos S, Mallet HP, Goffart IL, Gauthier V, Cardoso T, Herida M. Current Zika virus epidemiology and recent epidemics. *Med Mal Infect* 2014; **44**(7): 302-307.

[6] Musso D, Nhan T, Robin E, Roche C, Bierlaire D, Zisou K, et al. Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. *Euro Surveill* 2014; **19**(14): 1-3.

[7] Scully C, Robinson A. Check before you travel: Zika virus-another emerging global health threat. *Br Dent J* 2016; **220**(5): 265-267.

[8] Lanciotti RS, Kosoy OL, Laven JJ, Velez JO, Lambert AJ, Johnson AJ, et al. Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg Infect Dis* 2008; **14**(8): 1232-1239.

[9] Rodriguez-Morales AJ, Bandeira AC, Franco-Paredes C. The expanding spectrum of modes of transmission of Zika virus: a global concern. *Ann Clin Microbiol Antimicrob* 2016; **15**(1): 1.

[10] Maestre AM, Caplivski D, Fernandez-Sesma A. Zika virus: more questions than answers. *EBio Med* 2016; **5**: 2-3.

[11] Mansuy JM, Dutertre M, Mengelle C, Fourcade C, Marchou B, Delobel P, et al. Zika virus: high infectious viral load in semen, a new sexually transmitted pathogen. *Lancet Infect Dis* 2016; **16**(405): 00138-00139.

[12] Kashima S, Slavov SN, Covas DT. Zika virus and its implication in transfusion safety. *Rev Bras Hematol Hemoter* 2016; **38**(1): 90-91.

[13] Focosi D, Maggi F, Pistello M. Zika virus: implications for public health. *Clin Infect Dis* 2016; **63**(2): 227-233.

[14] Lucchese G, Kanduc D. Zika virus and autoimmunity: from microcephaly to Guillain-Barré syndrome, and beyond. *Autoimmun Rev* 2016; **16**: 53-58.

[15] Arzuza-Ortega L, Polo A, Pérez-Tatis G, López-García H, Parra E, Pardo-Herrera LC. Fatal Zika virus infection in girl with sickle cell disease, Colombia. *Emerg Infect Dis* 2016; **22**(5); http://dx.doi.org/10.3201/eid2205.151934.

[16] Saxena SK, Elahi A, Gadugu S, Prasad AK. Zika virus outbreak: an overview of the experimental therapeutics and treatment. *Virusdisease* 2016; **27**(2): 111-115.

[17] Awasthi S. Zika virus: prospects for the development of vaccine and antiviral agents. *J Antivir Antiretrovir* 2016; **8**(1); http://dx.doi.org/10.4172/jaa.1000e130.

[18] Guy B, Lang J, Saville M, Jackson N. Vaccination against dengue: challenges and current developments. *Annu Rev Med* 2016; **67**(1): 387-404.

[19] Hari A, Ganguly A, Mu L, Davis SP, Stenner MD, Lam R, et al. Redirecting soluble antigen for MHC class I cross-presentation during phagocytosis. *Eur J Microbiol Immunol* 2015; **45**(2): 383-395.
[20] Oyarzun P, Ellis JJ, Gonzalez-Galarza FF, Jones AR, Middleton D, Boden M, et al. A bioinformatics tool for epitope-based vaccine design that accounts for human ethnic diversity: application to emerging infectious diseases. *Vaccine* 2015; **33**(10): 1267-1273.
[21] Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. *Open Biol* 2013; **3**(1): 806-811.
[22] Pickett BE, Greer DS, Zhang Y, Stewart L, Zhou L, Sun G, et al. Virus pathogen database and analysis resource (ViPR): a comprehensive bioinformatics database and analysis resource for the coronavirus research community. *Viruses* 2012; **4**(11): 3209-3226.
[23] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013; **30**(4): 772-780.
[24] Singh H, Raghava GP. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 2003; **19**(8): 1009-1014.
[25] Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 2001; **17**(12): 1236-1237.
[26] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007; **8**(1): 1-7.
[27] Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* 2015; **43**(D1): D405-D412.
[28] Mlakar J, Korva M, Tul N, Popović M, Poljšak-Prijatelj M, Mraz J, et al. Zika virus associated with microcephaly. *N Engl J Med* 2016; **374**(10): 951-958.
[29] Adhisivam B, Bhat BV. 17 vector-borne diseases. Recent advances in pediatrics-special volume 22-Immunology. *Infect Immun* 2013: 313.
[30] Gowen BB, Holbrook MR. Animal models of highly pathogenic RNA viral infections: hemorrhagic fever viruses. *Antiviral Res* 2008; **78**(1): 79-90.
[31] Oyarzun P, Kobe B. Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. *Hum Vaccin & Immunother* 2016; **12**(3): 763-767.
[32] Steward MW. The development of a mimotope-based synthetic peptide vaccine against respiratory syncytial virus. *Biologicals* 2001; **29**(3): 215-219.
[33] Almanzar G, Herndler-Brandstetter D, Chaparro SV, Jenewein B, Keller M, Grubeck-Loebenstein B. Immunodominant peptides from conserved influenza proteins–A tool for more efficient vaccination in the elderly? *Wien Med Wochenschr* 2007; **157**(5–6): 116-121.
[34] Amin N, Pupo M, Aguilar A, Camacho F, Alvarez M, Caballero Y, et al. Immunogenicity of NS4b dengue 3 virus mimotope presented to the immune system as multiple antigen peptide system. *ISRN Virol* 2013: 2013; http://dx.doi.org/10.5402/2013/924057.
[35] Hafner C, Wagner S, Jasinska J, Allwardt D, Scheiner O, Wolff K, et al. Epitope-specific antibody response to Mel-CAM induced by mimotope immunization. *J Investig Dermatol* 2005; **124**(1): 125-131.
[36] Olsen AW, Hansen PR, Holm A, Andersen P. Efficient protection against *Mycobacterium tuberculosis* by vaccination with a single subdominant epitope from the ESAT-6 antigen. *Eur J Microbiol Immunol* 2000; **30**(6): 1724-1732.
[37] Vaughan K, Greenbaum J, Blythe M, Peters B, Sette A. Meta-analysis of all immune epitope data in the *Flavivirus* genus: inventory of current immune epitope data status in the context of virus immunity and immunopathology. *Viral Immunol* 2010; **23**(3): 259-284.
[38] Koff WC, Burton DR, Johnson PR, Walker BD, King CR, Nabel GJ, et al. Accelerating next-generation vaccine development for global disease prevention. *Science* 2013; **340**(6136). 1232910.
[39] Bashyam HS, Green S, Rothman AL. Dengue virus-reactive CD8+ T cells display quantitative and qualitative differences in their response to variant epitopes of heterologous viral serotypes. *J Immunol* 2006; **176**(5): 2817-2824.
[40] Heinz FX, Berger R, Tuma W, Kunz C. A topological and functional model of epitopes on the structural glycoprotein of tick-borne encephalitis virus defined by monoclonal antibodies. *Virol ogy* 1983; **126**(2): 525-537.
[41] Allwinn R, Doerr H, Emmerich P, Schmitz H, Preiser W. Cross-reactivity in flavivirus serology: new implications of an old finding? *Immunol Med Microbiol* 2002; **190**(4): 199-202.
[42] Slifka MK. Vaccine-mediated immunity against dengue and the potential for long-term protection against disease. *Front Immunol* 2014; **5**(5). 195–195.
[43] Kollaritsch H, Paulke-Korinek M, Holzmann H, Hombach J, Bjorvatn B, Barrett A. Vaccines and vaccination against tick-borne encephalitis. *Expert Rev Vaccines* 2012; **11**(9): 1103-1119.
[44] Guy B, Lang J, Saville M, Jackson N. *Ann Rev Med* 2016; **67**: 387-404.