Original article:

IN SILICO CD4+, CD8+ T-CELL AND B-CELL IMMUNITY ASSOCIATED IMMUNOGENIC EPITOPE PREDICTION AND HLA DISTRIBUTION ANALYSIS OF ZIKA VIRUS

Essam Mohammed Janahi[a*], Anupam Dhasmana[b,c], Vandana Srivastava[d],
Aditya Narayan Sarangi[e], Sana Raza[c,f], Jamal M. Arif[g], Madan Lal Bramha Bhatt[c],
Mohtashim Lohani[c,h], Mohammed Yahya Areeshi[h], Anand Murari Saxena[d], Shafiul Haque[h,i*]

[a] Department of Biology, College of Science, University of Bahrain, P.O. Box 32038, Kingdom of Bahrain
[b] Research Cell, Amity University Lucknow Campus, Lucknow-226028, UP, India
[c] Department of Radiotherapy, King George Medical University, Lucknow-226003, UP, India
[d] Department of Zoology, Lucknow University, Lucknow-226007, UP, India
[e] Biomedical Informatics Centre, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow-226014, UP, India
[f] Department of Biosciences, Integral University, Lucknow-226026, UP, India
[g] Department of Biochemistry, University of Hail, Hail-2440, Saudi Arabia
[h] Research and Scientific Studies Unit, College of Nursing and Allied Health Sciences, Jazan University, Jazan-45142, Saudi Arabia
[i] Department of Biosciences, Jamia Millia Islamia (A Central University), New Delhi-110025, India
* Corresponding authors: Dr. Shafiul Haque, Research and Scientific Studies Unit, College of Nursing and Allied Health Sciences, Jazan University, Jazan-45142, Saudi Arabia, Phone & Fax No.: +966-1-73174383, E-mail: shafiul.haque@hotmail.com
Dr. Essam Mohammed Janahi, Department of Biology, College of Science, University of Bahrain, P.O. Box 32038, Kingdom of Bahrain, Phone: +973-1743 7425, Fax: +973-17449158, Email: emohammed@uob.edu.bh

ABSTRACT

Zika virus (ZIKV) is a mosquito-borne flavivirus distributed all over Africa, South America and Asia. The infection with the virus may cause acute febrile sickness that clinically resembles dengue fever, yet there is no vaccine, no satisfactory treatment, and no means of evaluating the risk of the disease or prognosis in the infected people. In the present study, the efficacy of the host's immune response in reducing the risk of infectious diseases was taken into account to carry out immuno-informatics driven epitope screening strategy of vaccine candidates against ZIKV. In this study, HLA distribution analysis was done to ensure the coverage of the vast majority of the population. Systematic screening of effective dominant immunogens was done with the help of Immune Epitope & ABCPred databases. The outcomes suggested that the predicted epitopes may be protective immunogens with highly conserved sequences and bear potential to induce both protective neutralizing antibodies, T & B cell responses. A total of 25 CD4+ and 16 CD8+ peptides were screened for T-cell mediated immunity. The predicted

epitope "TGLDFSDLYYLTMNNKHWLV" was selected as a highly immunogenic epitope for humoral immunity. These peptides were further screened as non-toxic, immunogenic and non-mutated residues of envelop viral protein. The predicted epitope could work as suitable candidate(s) for peptide based vaccine development. Further, experimental validation of these epitopes is warranted to ensure the potential of B- and T-cells stimulation for their efficient use as vaccine candidates, and as diagnostic agents against ZIKV.

INTRODUCTION

Zika virus (ZIKV), a pathogenic vector borne flavivirus is a member of the Flaviviridae family, transmitted by the Aedes mosquitoes and is known to cause Zika fever. This virus was first identified in the blood of a sentinel Rhesus monkey during a study on Yellow fever in the Zika forest of Uganda in 1947. This was followed by subsequent isolation of this virus from a pool of Aedes mosquitoes of African region (Dick et al., 1952). Since then, several outbreaks of Zika have been reported in humans in Africa and Asia. Currently, it has been reported in Brazil, the Pacific region, and is also known to spread widely in Central and South America.

The Zika fever illness is generally characterized by some of the symptoms which are similar to other arbovirus infections like Dengue and Chikungunya. These symptoms include fever, malaise, conjunctivitis, rashes on skin, anorexia, headache as well as pain in the muscles and joints (Simpson, 1964). In case of pregnant women, Zika virus may spread from the mother to the fetus and is believed to be associated with microcephaly, a serious birth defect characterized by smaller and poorly developed brain in the newborns (Torjesen, 2016). Pregnant women are therefore advised to avoid traveling to areas prone to incidence of Zika. Zika virus has 25-30 nm nucleocapsid, which is surrounded by a lipid bilayer containing envelope proteins. It has a single stranded RNA genome with 10,141 base pairs accountable for encoding of nearly 3380 amino acids (Fonseca et al., 2014). The virus has two flanking 5' and 3' non-coding regions (NCR) and a long open reading frame (ORF) "59-C-prM-E-NS1-NS2A-NS2BNS3-NS4A-NS4B-NS5-39" coding for a polyprotein that cleaves into its envelope (E), capsid (C), precursor membrane (prM), and some other non-structural (NS) proteins (Chambers et al., 1990; Kuno et al., 2007). Zika virus is maintained through a sylvatic transmission cycle. The envelope protein is the major surface protein of the virus, and is involved in various aspects of the viral cycle and is believed to mediate binding and membrane fusion (Lindenbach et al., 2003). The non-structural protein, NS5 has RNA-dependent RNA polymerase (RdRP) that acts at its C-terminal, while the N-terminal is involved in RNA capping (Lindenbach and Rice, 2003). The 3' non-coding region is made up of about 428 nucleotides having 27 folding patterns (Kuno et al., 2007), which are believed to be involved in the processes like translation, stabilization of the genome, packaging of RNA, and recognition of the cellular or the viral factors (Lindenbach and Rice, 2003).

Due to lack of rapid diagnostic techniques, effective treatment and availability of novel vaccines against the Zika virus has been a public health emergency of the international concern declared by the World Health Organization (Salaam-Blyther, 2016). Hence, it is necessary to develop a substantial amount of immunity in humans against Zika virus to prevent the widespread outbreak of the infection and its associated effects. Taking a step towards this development, we screened the immunogenic epitopes of the envelope protein

of Zika virus using an *in silico* approach. This might be helpful in developing peptide based vaccines against the Zika virus.

MATERIALS AND METHODS

Sequence retrieval

The 504 length amino acid sequence of viral envelope (E) protein (GenBank: AIC06934.1 & RCSB PDB ID: 5IZ7), involved in the host cell binding and fusion activity was retrieved from UniprotKB Database (www.uniprot.org). The retrieved sequence (A0A060H177_9FLAV E protein) was a further subjected to immunogenicity assessment and epitope prediction.

Immunogenicity prediction of the viral protein

The VaxiJen V2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) was used for the evaluation of immunogenicity of the retrieved protein sequence. This server works on Auto Cross Covariance (ACC) algorithm that predicts protective antigens, tumor antigens, and subunit vaccines with the precision level of 70 to 89 % for the discrimination between antigens and non-antigens (Doytchinova and Flower, 2007).

Epitope prediction

(a) B-cell epitope prediction

B-cell epitopes were predicted using an ABCPred online server (www.imtech.res.in/abcpred) by selecting the window length of 20 amino acids, which were recommended for humaoral immunity. This server is based on information - processing algorithms inspired by the biological nervous system. The server ranks the epitopes as per their respective scores. The higher score of the peptides means a greater probability of becoming most suitable immunogenic epitope (Saha et al., 2006).

(b) CD4+ and CD8+ epitope prediction

The envelop protein sequence was analyzed for the screening of the possible dominant T-cell epitopes using immuno-informatics (Immune Epitope Database) tool IEDB online server (www.iedb.org) (Bui et al., 2006). Human leukocyte antigen (HLA)-I and HLA-II binding resources present in the IEDB, were used for the binding analysis of all the possible peptides considering the presence of all the HLAs in the database. The peptide length was set to be 10 and 15 for HLA-I and HLA-II, respectively. Default IEDB recommended prediction method combines the predictions from ANN, SMM, and CombLib algorithms. The epitopes were predicted on the basis of lowest percentile rank and high binding affinity.

Antigenicity of the identified epitopes

The immunogenicity of both B-cell and T-cell predicted epitopes was evaluated using a VexiJen V2.0 online server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html)

Determination of conserved regions & toxicity of the identified epitopes

The immunogenic epitopes were checked for the conserved regions and further subjected to ToxinPred severe (Gupta et al., 2013) for the differentiation of toxic or non-toxic peptides. The ToxinPred server is based on Support Vector Machine (SVM) and Quantitative Matrix based algorithm, and generates quantitative matrix on the basis of probability or frequency of amino acid at a particular position.

Determination of physico-chemical properties of the identified epitopes

The Peptide Property Calculator (https://www.genscript.com) server was used to determine the best solvent for the predicted peptide based on its amino acid sequence.

Population coverage analysis

Due to the dependence of major histocompatibility complex (MHC) of T-cell response, the peptides with wide range of HLA binding

specificities result in wider population coverage in terms of geographical expansion. The population coverage rate of the predicted epitopes was calculated by employing the IEDB population coverage tool (http://tools.immuneepitope.org/tools/population/iedb_input)(Bui et al., 2006). The predicted epitopes with its all binding HLA alleles for the worldwide distribution were tabulated. The schematic representation of the entire methodology of *in silico* CD4+, CD8+ T-cell & B-cell epitope prediction and HLA distribution of Zika virus is given in Figure 1.
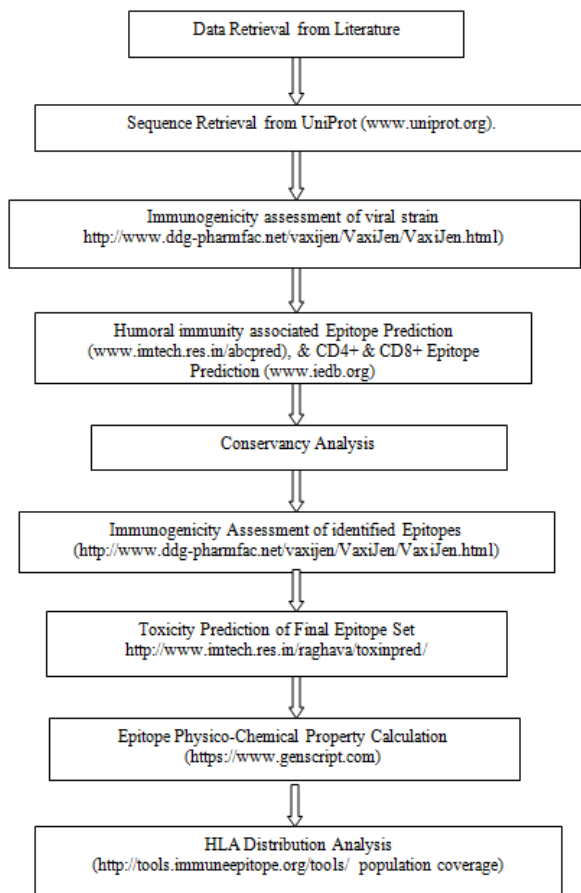

Figure 1: Schematic representation of the entire *in silico* strategy of CD4+, CD8+ T-cell and B-cell epitope prediction and HLA distribution of Zika virus

RESULTS

Protein immunogenicity and epitope identification

The immunogenicity of the viral envelope protein was ensured using VaxiJen V2.0 online server, keeping the threshold at 0.4.

The results obtained suggest that the viral sequence is a probable antigen with a score of 0.6178. The prediction and identification of B-cell epitopes in target antigens, being a key step in the epitope based vaccine development, was performed by using ABCPred server. Based on the artificial neural network (ANN) method, the epitopes predicted by ABCPred were in the descending order of their scores, depicting that the top most has the best binding affinity. Therefore, the first (top) epitope 'TGLDFSDLYYLTMNN-KHWLV' with the score 0.90 at the position 194 to 213 showing the highest probability to be an antigenic determinant, was selected for further study.

MHC-II restricted CD4+ T-cells activation plays an important role in initiating and upholding a proficient antibody response or cytotoxic T-lymphocytes (CTL) response. Further, MHC-I restricted CD8+ CTLs have a vital role in combating the viral infection. Thus, the comprehending principles of T-cell activation and epitope-based vaccine design, direct the identification of helper T-cell epitopes and CTL epitopes. The helper T-cell epitopes and CTL epitopes in the 'Env' protein of Zika virus were identified using the IEDB online tool. The immunogenic peptides were ranked in accordance with their consensus percentile rank from SWISS-PROT program. The lowest percentile rank demonstrated as the best binders. By selecting the first (lowest percentile rank) epitope for each of HLA class-II and HLA class-I gave 63 helper T-cell epitopes (Supplementary Information: Table S1) and 76 CTL epitopes (Supplementary Information: Table S2).

Conserved regions and immunogenicity of the identified epitopes

Peptides with strong antigenicity are always better candidates to be B- and T-cell epitopes than those with weak antigenicity. Hence, the assessment of immunogenicity of the identified peptides was performed using the VaxiJen V2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). This tool predicts the antigenicity of a peptide

on the basis of amino acid properties and their positions in the peptide. The results demonstrated that the identified B-cell epitope is immunogenic in nature with score of 1.2758, and among all the identified T-cell epitopes, a total of 16 HLA class-II and 25 HLA class-I epitopes were found to be conserved as well as immunogenic (Supplementary Information: Table S3).

Toxicity and solvency analysis

The final epitope set (B- and T-cell epitopes) was further tested for the toxicity analysis and medium solvency. All the predicted immunogenic epitopes were non-toxic and were classified as acidic, basic or neutral (Table 1). Acidic peptides were initially dissolved in water, if not, then < 50 µl NH_4OH was added and diluted to 1 ml with deionized water. For basic peptides, 10 % and higher solutions of acetic acid was tried; in case not dissolved, then < 50 µl TFA was added and diluted up to 1 ml with deionized water to solubilize the peptide. Neutral peptides required organic solvents like acetonitrile, methanol, or isopropanol.

Population coverage analysis

HLA allocation varies among diverse ethnic groups spread over different geographical regions of the world. Therefore, while designing a potent vaccine, population coverage must be taken into account to cover maximum possible populations. In this study, all alleles in IEDB database were identified as best binders with the predicted epitopes and were used to establish the population coverage for these epitopes (Table 2). A significant percentage of the specific area population coverage of all the identified immunogenic epitopes was found all over the world.

DISCUSSION

Vaccination appears to be the most optimal and successful strategy to reduce the menace of dreadful infectious diseases worldwide. Currently, the immune-informatics approach is the most promising approach for the development of peptide based vaccines. Due to lack of effective treatment for the microcephaly caused by Zika virus, there is an urgent need for a cost effective vaccine against this deadly virus. Earlier studies have already demonstrated that application of potential epitope(s) instead of whole protein as a vaccine has been confirmed to represent the complete antigenicity of any protein (DeGroot et al., 2002). Also, in the recent past, the epitope based vaccines have given promising results against highly infectious diseases such as H1N1, HIV and Tuberculosis (Correia et al., 2014). These vaccines may have a lower chance of eliciting immune response against self-antigens, thereby avoiding autoimmune responses (Aziz et al., 2008). The traditional molecular immunology techniques used to search for immunogenic peptides to be used as vaccine candidates are extensively time consuming, labour intensive, and expensive. However, as a remedy, in recent years, various computational approaches have been proposed and tested by many researchers to reduce the time, cost, and labour involved (Davies and Flower, 2007; Shi et al., 2015). On the similar lines, the suggested *in silico* strategy might prove as a boon for the experimental biologists in rapid screening and identification of probable vaccine candidates. T-lymphocytes play a central role in the generation of a protective immune response in many microbial infections (Esser et al., 2003), hence the prediction and screening of T-cell epitopes can be exploited as a potential vaccine candidate is a critical requirement for the designing of epitope-based vaccines.

Table 1: Prediction of most immunogenic, non-mutant and non-toxic B-cell and T-cell epitopes of the envelope protein of Zika virus (the threshold immunogenicity of the peptides was 0.4 and the total predicted range of VaxiJen score was from 2.0901 to 0.4505)

No.	Peptides/Epitopes		Start	End	Percentile Rank	Immunogenicity		Tox-icity	Attrib-ute
						Thresh-old	VaxiJen Score*		
	B-cell epitopes								
1	TGLDFSDLYYLTMNNKHWLV		194	213	0.90	0.4	1.2758	Non-Toxic	neutral
	T-cell epitopes								
	HLA Class-I								
1	SLMCLALGGV	HLA-A*02:01, HLA-A*02:06	485	494	0.2	0.4	1.7363	Non-Toxic	Neutral
2	LLMWLGLNTK	HLA-A*03:01	471	480	0.25	0.4	1.8391	Non-Toxic	Basic
3	EATLGGFGSL	HLA-A*25:01	177	186	1.7	0.4	1.0328	Non-Toxic	Acidic
4	ELDPPFGDSY	HLA-A*26:01	377	386	0.9	0.4	0.6953	Non-Toxic	Acidic
5	NMAEVRSYCY	HLA-A*29:02, HLA-B*15:01	52	61	0.3	0.4	1.2831	Non-Toxic	Neutral
6	KAFEATVRGA	HLA-A*30:01	409	418	1.35	0.4	0.9016	Non-Toxic	Basic
7	RTGLDFSDLY	HLA-A*30:02	193	202	0.25	0.4	1.8346	Non-Toxic	Acidic
8	RLSSGHLKCR	HLA-A*31:01, HLA-A*74:01	283	292	0.7	0.4	2.0901	Non-Toxic	Basic
9	DMQTLTPVGR	HLA-A*33:03, HLA-C*04:01	348	357	0.7	0.4	0.8543	Non-Toxic	Neutral
10	TTVSNMAEVR	HLA-A*02:01, HLA-A*02:06	48	57	0.2	0.4	0.7175	Non-Toxic	Neutral
11	TTTVSNMAEV	HLA-A*03:01	47	56	0.3	0.4	0.6972	Non-Toxic	Acidic
12	TPHWNNKEAL	HLA-A*25:01	233	242	0.35	0.4	1.5710	Non-Toxic	Basic
13	LRLKGVSYSL	HLA-A*26:01	298	307	0.35	0.4	1.1682	Non-Toxic	Basic
14	VDIELVTTTV	HLA-A*29:02, HLA-B*15:01	41	50	0.6	0.4	1.3702	Non-Toxic	Acidic
15	KMMLELDPPF	HLA-A*30:01	373	382	0.2	0.4	1.0318	Non-Toxic	Acidic
16	AETLHGTVTV	HLA-A*02:01, HLA-A*02:06	319	328	0.3	0.4	0.5422	Non-Toxic	Neutral
17	HHWHRSGSTI	HLA-A*03:01	398	407	0.3	0.4	0.4776	Non-Toxic	Basic
18	RMAVLGDTAW	HLA-A*25:01	420	429	0.2	0.4	0.8599	Non-Toxic	Neutral
19	YCYEASISDM	HLA-A*26:01	59	68	0.2	0.4	1.0191	Non-Toxic	Acidic
20	AHAKRQTVVV	HLA-A*29:02, HLA-B*15:01	248	257	0.3	0.4	0.7216	Non-Toxic	Basic
21	SRCPTQGEAY	HLA-A*30:01	72	81	0.9	0.4	0.6146	Non-Toxic	Neutral
22	FGDSYIVIGV	HLA-A*30:02	382	391	0.2	0.4	0.6413	Non-Toxic	Acidic
23	DAHAKRQTVV	HLA-A*31:01, HLA-A*74:01	247	256	0.5	0.4	0.6538	Non-Toxic	Basic
24	ESTENSKMML	HLA-A*33:03, HLA-C*04:01	367	376	0.5	0.4	0.4843	Non-Toxic	Acidic
25	AVDMQTLTPV	HLA-A*68:01	346	355	0.4	0.4	0.7836	Non-Toxic	Acidic

No.	Peptides/Epitopes		Start	End	Percentile Rank	Immunogenicity		Toxicity	Attribute
						Threshold	Vaxi-Jen Score*		
	B-cell epitopes								
1	TGLDFSDLYYLTMNNKHWLV		194	213	0.90	0.4	1.2758	Non-Toxic	neutral
	T-cell epitopes								
	HLA Class-II								
1	RTGLD-FSDLYYLTMN	HLA-DPA1*01/DPB1*04:01	193	207	0.61	0.4	1.4016	Non-Toxic	Acidic
2	PRTGLD-FSDLYYLTM	HLA-DPA1*01:03/DPB1*02:01	192	206	0.05	0.4	1.4164	Non-Toxic	Acidic
3	WLVHKEW-FHDIPLPW	HLA-DQA1*01:01/DQB1*05:01	211	225	0.52	0.4	0.6855	Non-Toxic	Basic
4	NTKNGSISLM-CLALG	HLA-DQA1*01:02/DQB1*06:02	478	492	1.15	0.4	1.9904	Non-Toxic	Basic
5	AVHTALA-GALEAEMD	HLA-DQA1*03:01/DQB1*03:02, HLA-DQA1*04:01/DQB1*04:02, HLA-DQA1*05:01/DQB1*02:01	264	278	0.79	0.4	0.4323	Non-Toxic	Acidic
6	EGAVHTALA-GALEAE	HLA-DQA1*05:01/DQB1*03:01	262	276	0.93	0.4	0.4632	Non-Toxic	Acidic
7	QPENLEYRIM-LSVHG	HLA-DRB1*01:02, HLA-DRB1*04:08, HLA-DRB1*04:10, HLA-DRB1*04:21, HLA-DRB1*04:26, HLA-DRB1*11:20, HLA-DRB1*11:28, HLA-DRB1*13:05, HLA-DRB5*01:05	131	145	0.23	0.4	0.4589	Non-Toxic	Neutral
8	EPRTGLD-FSDLYYLT	HLA-DRB1*03:01	191	205	0.72	0.4	1.1568	Non-Toxic	Acidic
9	RLKMDKLRL-KGVSYS	HLA-DRB1*03:05, HLA-DRB1*04:02, HLA-DRB1*08:04, HLA-DRB1*08:06, HLA-DRB1*08:13, HLA-DRB1*11:02, HLA-DRB1*11:07, HLA-DRB1*11:14, HLA-DRB1*11:21, HLA-DRB1*13:01, HLA-DRB1*13:04, HLA-DRB1*13:22, HLA-DRB1*13:23, HLA-DRB1*13:27	292	306	0.67	0.4	0.7441	Non-Toxic	Basic
10	GTLLMWLGLNTKNGS	HLA-DRB1*03:06, HLA-DRB1*03:07, HLA-DRB1*03:08,	469	483	0.18	0.4	1.2662	Non-Toxic	Basic
11	DFSDLYYLT-MNNKHW	HLA-DRB1*04:01	197	211	0.11	0.4	1.1408	Non-Toxic	Neutral
12	LALGGVLIFLSTAVS	HLA-DRB1*04:04, HLA-DRB1*04:23, HLA-DRB1*08:02, HLA-DRB1*11:01, HLA-DRB1*11:04, HLA-DRB1*11:06, HLA-DRB1*13:07, HLA-DRB1*13:11	489	503	0.3	0.4	0.4505	Non-Toxic	Neutral
13	FSDLYYLT-MNNKHWL	HLA-DRB1*04:05	198	212	0.3	0.4	0.8870	Non-Toxic	Basic
14	PRAEAT-LGGFGSLGL	HLA-DRB1*07:03, HLA-DRB1*15:02, HLA-DRB1*15:06	174	188	0.58	0.4	1.4679	Non-Toxic	Neutral
15	CLALGGVLIFLSTAV	HLA-DRB1*15:01	488	502	0.07	0.4	0.5558	Non-Toxic	Neutral
16	ENSKMM-LELDPPFGD	HLA-DRB3*01:01	370	384	0.01	0.4	0.7744	Non-Toxic	Acidic

Table 2: Prediction of population coverage rate (%) of MHC Class-I and MHC Class-II

S.No.	MHC Class-1				MHC Class-II		
	Population / Area	Coverage [a]	Average hit [b]	PC90 [c]	Coverage [a]	Average hit [b]	PC90 [c]
1	East Asia	91.33%	1.88	1.05	98.78%	3.29	2.05
2	Northeast Asia	82.73%	1.46	0.58	99.78%	4.13	2.74
3	South Asia	91.88%	1.94	1.07	99.93%	4.72	3.26
4	Southeast Asia	82.82%	1.44	0.58	98.75%	3.18	2.01
5	Southwest Asia	90.46%	1.83	1.02	99.62%	4.02	**2.57**
6	Europe	98.67%	2.85	1.88	99.99%	5.36	3.98
7	East Africa	95.00%	2.18	1.23	100.00%	5.25	4.01
8	West Africa	93.74%	2.08	1.15	99.97%	5.00	3.63
9	Central Africa	94.33%	2.14	1.19	99.94%	4.57	3.20
10	North Africa	95.49%	2.24	1.26	99.85%	4.44	3.06
11	South Africa	96.61%	2.32	1.35	45.98%	0.49	0.19
12	West Indies	88.48%	1.57	0.87	99.41%	3.46	2.24
13	North America	96.52%	2.35	1.35	100.00%	5.92	4.54
14	Central America	6.25%	0.06	0.11	99.99%	4.26	3.16
15	South America	88.06%	1.65	0.84	100.00%	5.13	3.83
16	Oceania	74.87%	1.17	0.40	99.95%	4.88	3.41
Average (Standard deviation)		83.71 **(22.26%)**	1.71 **(0.57)**	0.91 **(0.35)**	96.37% **(13.02%)**	4.26 **(1.22)**	2.99 **(1.01)**

Recently, a similar approach was used for the identification of highly immunogenic peptides of Env protein of Zika virus (Badawi et al., 2016) and suggested highly recommended T cell epitopes MMLELDPPF as therapeutic peptide vaccine to interact with MHC class I along with three other epitopes MAVLGDTAW, KEWFHDIPL and DTAW-DFGSV. Also, three epitopes showed high affinity to interact with MHC class II alleles i.e., FKSLFGGMS, LITANPVIT and VHTALA-GAL. Alam et al. (2016) proposed MHC class-I potential peptides YRIMLSVHG, VLIFLSTAV and MMLELDPPF, GLD-FSDLYY as potent peptide epitopes for CD4+ and CD8+ T-cells, respectively. The current study was aimed to identify B-cell, T-cell highly immunogenic, non toxic and non variable peptides which have maximum world population coverage area.

The sequence of the viral envelope protein was used for the evaluation of the viral immunogenicity to predict the vaccine candidates against the immune-responses of B- and T-cells. The B-cell epitope and the best/highest percentile score generated a single epitope 'TGLDFSDLYYLTMNNKHWLV' at a position of 194 to 213 with immunogenic score 1.2758. The putative CD4+ & CD8+ T-cell core epitopic sequences for all the HLA super types were predicted. We also checked our predicted immunogenic epitopes for the conserved regions and further subjected them to ToxinPred server for the differentiation of toxic or non-toxic peptides. The assessment of the immunogenicity and toxicity properties of the putative epitopes using various online servers showed that the predicted epitopes were immunogenic in nature, and no sequence variability was detected in the region containing the potential epitopes.

The epitope NTKNGSISLMCLALG at position 478-492 for HLA Class-II and the epitope RLSSGHLKCR at position 283-292 for HLA Class-I was found to have the highest antigenic score of 1.9904 and 2.0901, respectively, and both were found to be basic in nature (Table 1).

The putative epitopes have excellent worldwide population coverage as depicted in Table 2, and may provide worldwide immune protection (83.71 % of the average MHC

Class-I coverage and 96.37 % of the average MHC Class-II coverage) against ZIKV.

The application of *in silico* approaches by immunologists has helped to speed up the *in vitro, in vivo* and *ex vivo* studies by cutting down the large epitope dataset into a smaller one that needs to be confirmed by the experimental analysis. The predictions made by the present study on the primary amino acid sequence of Zika virus should help other researchers to experimentally verify the immuno-dominant epitopes fundamental to the understanding of the pathobiology of Zika virus leading to effective vaccine development.

The data from the current work on B- and T-cell peptide epitopes for 'Env' protein of Zika virus might prove a milestone in the development of vaccines to combat Zika virus induced infection. The suggested *in silico* strategy might prove as a boon for the experimental biologists in rapid screening and identification of probable vaccine candidates. The proposed non–toxic epitopes would be a relevant representative of a large proportion of the human population. This work indicates that the proposed epitopes prompt the future vaccine development against the Zika virus. However, the potential of the screened epitopes by experimental validation is warranted.

Supplementary Information

Table S1. Predicted immunogenic peptides for the HLA class-II

Table S2. Predicted immunogenic peptides for the HLA class-I

Table S3. Immunogenicity assessment of the identified peptides using the VaxiJen

Statement of conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

REFERENCES

Alam A, Ali S, Ahmad S, Malik MZ, Ishrat R. From ZikV genome to vaccine: in silico approach for the epitope-based peptide vaccine against Zika virus envelope glycoprotein. Immunology. 2016;149:386-99.

Azizi A, Anderson DE, Torres JV, Ogrel A, Ghorbani M, Soare C, et al. Induction of a broad cross-subtype-specific HIV-1 immune responses by a novel multivalent HIV-1 peptide vaccine in cynomolgus macaques. J Immunol. 2008;180:2174-86.

Badawi MM, Osman MM, Alla AAF, Ahmedani AM, Abdalla MH, Gasemelseed MM, et al. Highly conserved epitopes of ZIKA envelope glycoprotein may act as a novel peptide vaccine with high coverage: Immunoinformatics approach. Am J Biomed Res. 2016;4:46-60.

Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 2006;7:153.

Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organization, expression, and replication. Annu Rev Microbiol. 1990;44:649-88.

Correia BE, Bates JT, Loomis RJ, Baneyx G, Carrico C, Jardine JG, et al. Proof of principle for epitope-focused vaccine design. Nature. 2014;507:201-6.

Davies MN, Flower DR. Harnessing bioinformatics to discover new vaccines. Drug Discov Today. 2007;12: 389-95.

DeGroot AS, Sbai H, Aubin CS, McMurry J, Martin W. Immuno-informatics: Mining genomes for vaccine components. Immunol Cell Biol. 2002;80:255-69.

Dick GW, Kitchen SF, Haddow AJ. Zika virus isolations and serological specificity. Trans R Soc Trop Med Hyg. 1952;46:509-20.

Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8:4.

Esser MT, Marchese RD, Kierstead LS, Tussey LG, Wang F, Chirmule N. Memory T cells and vaccines. Vaccine. 2003;21:419-30.

Fonseca K, Meatherall B, Zarra D, Drebot M, Mac-Donald J, Pabbaraju K, et al. First case of zika virus infection in a returning canadian traveller. Am J Trop Med Hyg. 2014;91:1035-8.

Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R. In silico approach for predicting toxicity of peptides and proteins. PLoS ONE. 2013;8:e73957.

Kuno G, Chang JJ. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Arch Virol. 2007;152:687-96.

Lindenbach BD, Rice CM. Molecular biology of flaviviruses. Adv Virus Res. 2003;59:23-61.

Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins. 2006;65:40-8.

Salaam-Blyther T. Zika virus: global health considerations. CRS INSIGHT. April 14, 2016 (IN10433).

Shi J, Zhang J, Li S, Sun J, Teng Y, Wu M, et al. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: An in silico approach applied to emerging infectious diseases. PLoS ONE. 2015;10:e0144475.

Simpson DI. Zika virus infection in man. Trans R Soc Trop Med Hyg. 1964;58:335-8.

Torjesen I. Zika virus outbreaks prompt warnings to pregnant women. BMJ. 2016;352:i500.