Kermanshah University of Medical Sciences

Journal of Kermanshah University of Medical Science

Journal homepage: Htpp://journals.kums.ac.ir/ojs/index.php/jkums

Original Article

In silico Comparison of Structural Features and Predicted Epitopes of Envelope Protein of Zika Virus with the Homologous Proteins of Two Closely Related Viruses

Samira Ebrahimi[1], Hassan Mohabatkar[1]*, Mandana Behbahani[1]

1. Dept. of Biotechnology, Faculty of Advanced Sciences and Technologies, University of Isfahan, Isfahan, Iran

Article Info

Keywords: Zika virus, Flavivirus, Envelope protein, Bioinformatics

***Corresponding Author**:
Department of Biotechnology, Faculty of Advanced Sciences and Technologies, University of Isfahan, Isfahan, Iran; Phone: +983137934391
Fax: +983137932342
Email: h.mohabatkar@ast.ui.ac.ir

Received: 31 July, 2017
Accepted: 07 November, 2017

J Kermanshah Univ Med Sci. 2017; 21(3): 103-107

Abstract

Introduction: Of recent, Zika virus (ZIKV) has spread worldwide apart from its original geographical zone (Federated States of Micronesia). Investigations prove a relationship between ZIKV and microcephaly in newborn babies. Research on the essential proteins of this virus may help in preventing its epidemic.

Methods: Herein, the in silico study for envelope protein of ZIKV and two other viruses, that is, dengue virus (DENV) and Spondweni virus (SPOV), was carried out. In the present study, the essential structural properties of the envelope protein of ZIKV were predicted using the bioinformatics tools such as PseAAC, GOR IV, BCPREDS, MHCPred, DiANNA, GlycoEP, and MEGA.

Results: Envelope protein of ZIKV shares certain common features with DENV and SPOV. But our results show that ZIKV has similar structural features as SPOV in comparison to DENV.

Conclusion: By understanding the differences and similarities in these viruses, we may find an appropriate treatment for ZIKV. Also, finding the effectively shared epitopes between the similar viruses could help in finding a common vaccine between these viruses.

Introduction

Zika virus (ZIKV) belongs to the family of Flaviviridae and the genus Flavivirus. ZIKV is a small, enveloped virus, which has a single-stranded positive-sense RNA genome [1, 2]. This virus is a mosquito-borne flavivirus, which causes zoonosis. Zoonoses are naturally transmissible diseases from vertebrates to humans [2, 3]. Of recent, the world is facing a public health emergency due to the dramatic rise in the number of newborn babies with microcephaly. In countries that face an increased number of newborn babies with microcephaly, the spread of ZIKV is also observed [4]. As ZIKV and microcephaly in neonates seem to be associated [2], it is essential to explore the structural properties of the virus for appropriate treatment. The results of the serological surveys in Nigeria in 1954 reported that ZIKV infected humans [5]. Furthermore, several cases were reported in other African countries [6], Asian countries, several islands of the Pacific region since 2007, and in the United States since about early 2015 [7, 8]. Spread of ZIKV infection on Yap Island (2007) and in French Polynesia (2013–2014), with further spread to the Cook Islands, New Caledonia, and Easter Island, revealed that this arbovirus could spread outside its usual geographical zone and might cause large outbreaks [2]. Several mosquito species are found to be the vectors of ZIKV, including *Aedes africanus*, *Aedes apicoargenteus*, and *Aedes aegypti* [9].

Envelope protein is crucial in the biology of flaviviruses and it is the dominant antigen that plays a significant role in the immunological responses in the infected hosts [10]. This protein is essential during the initial attachment of the viral particle to the host cells. Envelope protein of flaviviruses includes three domains. Domain I organizes the whole structure. Domain II has four disulfide bridges. Domain III incorporates in the receptor–ligand interaction [11]. Discovering the structure of the envelope protein may help in understanding the biology and mechanism of the virus–cell interactions of these viruses. The present study was performed to predict certain medically important structural properties of the envelope protein of ZIKV and to compare this protein with the same protein of two closely related viruses, that is, Spondweni virus (SPOV) and dengue virus (DENV). The present *in silico* investigation provided a wide range of reliable results, which may prove to be economic and may save time.

Materials and methods

The reference sequence of the envelope protein of ZIKV (Accession Number: YP_009227198.1) with 500 amino acids length, that of DENV (Accession Number: NP_722460.2) with 495 amino acids length, and of SPOV (Accession Number: YP_009227187.1) with 505 amino acids length were fetched from the National Center for Biotechnology Information (NCBI) [12].

Amino Acid Composition

The amino acid compositions of the envelope proteins were calculated by PseAAC server. PseAAC can demonstrate a sequence in a separate model without missing the order of information. It is a mixture of a set of separate sequence correlation factors and the 20 components of the conventional amino acid composition [13].

Prediction of Secondary Structure

The 3-D structure of the ZIKV envelop protein is available at the Protein Data Bank (PDB) with PDB ID: 5JHM. In the present study, the percentage of each of the secondary structures was required for comparison between the viruses; therefore, the secondary structures of the proteins were predicted. Secondary structure prediction was performed by GOR (Garnier-Osguthorpe-Robson) IV. GOR provided two outputs: the first one reveals the number of amino acids in several secondary structures, whereas the second one demonstrates the likelihood values for each secondary structure at each amino acid position [14].

Prediction of B-Cell Epitopes

Knowledge of the epitopes represents a crucial step in the rational design of novel vaccines. Linear B-cell epitopes were predicted by BCPREDS, a server predicting the linear B-cell epitopes that explore the subsequence Kernel and amino acid pair antigenicity scale [15].

Prediction of T-Cell Epitopes

The T-cell epitopes were predicted by MHCPred, [16] a Common Gateway Interface server, written in Perl, which operates under Microsoft Windows NT [17].

Prediction of Disulfide Connectivity

The disulfide connectivity was predicted using the DiANNA server, which is a multistep program and a novel neural network-based approach. Initially, PSIPRED is run to predict the protein's secondary structure, and then PSIBLAST is run against the nonredundant SwissProt to obtain a multiple alignment of the input sequence [18].

Prediction of Glycosylation Sites

The glycosylation sites in a protein from its amino acid sequence were predicted using the GlycoEP server, in which the datasets were obtained from SwissProt. Protein sequences were in FASTA format, and C, N, and O types of glycosylation were chosen [19].

Prediction of Evolutionary Distance Between Envelope Proteins

Comparative analysis of the molecular sequences is necessary in order to reach the evolutionary history of species [20]. In the present study, MEGA 6.0 software was used to find the evolutionary distance between the envelope protein of ZIKV with DENV and SPOV. Evolutionary distance was constructed by using the neighbor joining algorithm.

Results

Basic Local Alignment Search Tool, which was available at NCBI revealed 72% identity between the ZIKV and SPOV envelope proteins and 59% identity between the ZIKV and DENV envelope proteins.

Amino Acid Composition

In all three envelope proteins, glycine and threonine were the most prevalent amino acids. Following these amino acids, the percentage of lysine, alanine, and serine in all three proteins was high and the least residue was tryptophan.

Prediction of Secondary Structure

Table 1 presents the results of the secondary structure prediction. The percentage of all types of the predicted secondary structures was not significantly different. The maximum percentage of the random coils was predicted in ZIKV.

Table 1. Predicted percentage of the secondary structures of envelope protein of the three viruses.

Virus Name	No. of Residues in Alpha Helices	No. of Residues in Extended Strand	No. of Residues in Coils	Alpha Helix %	Extended % Strand	Random Coil %
ZIKV	72	172	256	14.4	34.4	51.2
DENV	99	144	252	20	29.09	50.91
SPOV	102	170	227	20.2	34.85	44.95

Prediction of B-Cell Epitopes

Table 2 presents the results of the linear B-cell epitope prediction; it indicates eight regions as linear B-cell epitopes for the envelope protein of ZIKV. Corresponding regions were predicted for two other proteins, for example, our results revealed epitopes that were conserved between the three envelope proteins.

Table 2. Predicted B-cell epitope of envelope protein of the viruses.

No.	ZIKV	DENV	SPOV
1	156TDEDRAKVEVTPNSPRAEAT	153NETTEHGTTATITPQAPTSE	349DTNSMASTGRLITANPVVTE
2	324VEVQYAGTDGPCKIPVQMAV	389LSWFKKGSSIGKMFEATARG	120TCVKKLTGKSIQPENLEYRV
3	216DIPLPWHAGADTGTPHWNNK	364PVNIEAEPPFGESYIVVGAG	145ASQHGGMINNDTNHQHDKEN
4	374LDPPFGDSYIVIGVGDKKIT	243AHAKKQEVVVLGSQEGAMHT	95GYVDRGWGNGCGLFGKGSIV
5	118KFTCSKKMTGKSIQPENLEY	63AKISNTTTDSRCPTQGEATL	379IDPPFGDSYIIVGTGTTKIT
6	66SDMASDSRCPTQGEAYLDKQ	219PWTSGASTSQETWNRQDLLV	223SLPWHTGATSNNHHWNNKEA
7	94RTLVDRGWGNGCGLFGKGSL	99RGWGNGCGLFGKGSLITCAK	66SEMASDSRCPTQGEAYLDKM
8	345MQTLTPVGRLITANPVITES	320VLVQVKYEGTDAPCKIPFSS	426GDTAWDFGSVGGMFNSVGKF
9		25LEHGSCVTTMAKDKPTLDIE	273ALEAESDGHKATIYSGHLKC
10			1IRCIGIGNRDFIEGMSGGTW

Prediction of T-Cell Epitopes

Table 3 presents the results of T-cell epitope prediction. The predicted T-cell epitopes with the lowest predicted IC50 (inhibitory concentration), reveal the best binding affinities. The results of DRB0101 revealed the lowest IC50 for the envelope proteins of ZIKV and SPOV and A0203 revealed the lowest IC50 for DENV. The data presented in Table 3 are important due to the difference between the effect of different flaviviruses on MHC-I and MHC-II molecules and the role that their epitopes play in developing vaccines [21]. Working on such potent epitopes could lead us to multivalent flavivirus vaccines [22].

Table 3: MHC binding peptides of envelope protein of the viruses with the best-predicted binding affinities for each allele.

No.	Alleles	ZIKV		DENV		SPOV	
		Peptides	IC50(nm)	Peptides	IC50(nm)	Peptides	IC50(nm)
1	A0101	191GLDFSDLYY199	6.75	360DKEKPVNIE386	15.92	378EIDPPFGDS386	16.98
2	A0201	370MMLELDPPF378	9.55	135LKYSVIVTV143	14.45	59YCYEANISE67	10.07
3	A0202	486ALGGVMIFL494	1.75	206WLVHKQWFL214	2.39	89QFVCKRGYV97	5.94
4	A0203	44ELVTTTVSN52	1.36	474SLSMTCIAV482	1.32	477GLNARGGTV485	1.29
5	A0206	447AAFKSLFGG455	6.17	443AYGVLFSGV451	8.45	453AFKALFGGM461	15.85
6	A0301	328YAGTDGPCK336	11.67	128KIVQYENLK136	6.64	5GIGNRDFIE13	9.62
7	A1101	304CTAAFTFTK312	8.20	344GVTQNGRLI352	8.85	483GTVAMSFMG491	8.28
8	A3101	181SLGLDCEPR189	80.54	406ARGARRMAI414	79.43	172ASAPRVEVE180	93.97
9	A6801	308FTFTKVPAE316	12.05	211QWFLDLPLP219	6.52	462SWFTQLLIG470	10.57
10	A6802	142SVHGSQHSG150	17.86	431SVGKLIHQI439	18.07	434SVGGMFNSV442	13.84
11	A3501	281FSGHLKCRL289	125.6	279FAGKLKCRL287	89.33	412FEATMRGAK420	153.11
12	DRB0101	64SISDMASDS72	0.86	49EVTNPAVLR57	2.33	45LVTTTASNM53	0.40
13	DRB0401	338PVQMAVDMQ346	22.18	314ETQHGTVLV322	52.72	252KKQTAVVLG260	32.66
14	DRB0701	259GAVHTALAG267	15.45	444YGVLFSGVS452	13.27	137YRVLVSVHA145	6.73
15	H2Db	4IGVSNRDFV12	5.93	4VGIGNRDFV12	10.5	4IGIGNRDFI12	2.35
16	H2Kb	169SPRAEATL176	2.74	376SYIVVGAG383	2.26	58SYCYEANI65	2.20
17	H2Kk	132PENLEYRI139	2.97	84EEQDTNFV91	4.91	490MGIGAMLI497	2.82

Prediction of Disulfide Connectivity

DiANNA program revealed five disulfide bridges in the three envelope proteins (Table 4). Some of these disulfide bridges such as position 30-116 are conserved in different members of the flaviviruses including SPOV, DENV, AHFV [23], and ZIKV.

Prediction of Glycosylation Sites

Results of the glycosylation sites are presented in Table 5.

Table 4. Prediction of disulfide bonds in the envelope proteins.

No.	Cysteine Position		
	ZIKV	DENV	SOPV
1	3-105	30-116	3-105
2	30-92	60-92	30-92
3	30-116	60-116	30-116
4	60-92	74-105	74-105
5	60-116	92-121	74-121

Table 5. Predicted C, N, and O glycosylation sites in the envelope protein of the three viruses.

	ZIKV	DENV	SPOV
C-Linked Glycosylation	--------	--------	--------
N-Linked Glycosylation	--------	67	64
O-Linked Glycosylation	40,47,66	32,40,66,68,69,70	40,47,48,66

Prediction of Evolutionary Distance Between the Envelope Proteins

Results of the molecular evolutionary genetic analysis of the envelope proteins are presented in Table 6.

Table 6. Evolutionary distance between the envelope proteins of ZIKV, DENV, and SPOV.

Virus Name	DENV	SPOV
ZIKV	0.520	0.327

Discussion

Our results revealed that the percentage of amino acids with small side chains such as glycine and alanine in the envelope protein of all three viruses was higher [24, 25]. Among the three envelope proteins, ZIKV revealed the maximum percentage of random coils and the least percentage of alpha helices. In addition, our results demonstrated that ZIKV did not possess a complicated secondary structure. As the random coil scores well in the epitope prediction, the study on epitopes of envelope proteins is remarkable in the present investigation.

Finding epitopes by the experimental methods is expensive and time consuming. Therefore, the epitope prediction using bioinformatics methods can save the expense of peptide synthesis and the experimental working time [26, 27]. Epitopes were predicted for different groups of flaviviruses such as DENV [28],

Japanese encephalitis virus [29], and Yellow fever virus [30]. The best vaccine is the one that secures human beings against most of these viruses. Our epitope prediction results revealed certain shared epitopes between ZIKV, DENV, and SPOV. Finding the conserved epitopes between ZIKV and other flaviviruses could help in developing appropriate vaccines.

The presence of a disulfide bridge is essential in antigenicity, infectivity, and protective activities of different proteins [31]. In the present investigation, all of the three envelope proteins had five disulfide bridges in the nearby positions. In a research on the epitope models of tick-borne encephalitis virus envelope glycoprotein, it was reported that the disulfide bridges were crucial in the conformation, and consequently, in the activity of the epitopes [32].

Glycosylation is one of the most abundant and essential posttranslational modification of proteins. Glycoproteins are entailed in various biological functions like cell–cell interactions, protein folding, cell recognition, and host–pathogen interactions [19]. Interestingly, in Flaviviridae, a carbohydrate attached to a viral protein may have different functions or may cause different structures. It is indicated that a single N-linked glycosylation site in the premembrane and envelope proteins of a flavivirus called JEV (has high homology with ZIKV) is crucial for folding of the protein. It makes the single N-linked glycosylation site pivotal in the cytotoxicity and production of JEV particles [33]. Therefore, the different glycosylation sites in ZIKV envelope protein could emphasize its importance. Studies on the evolutionary distance reported that ZIKV is close to SPOV. This confirms our results because ZIKV has similar structural features as SPOV in comparison to DENV.

Conclusion

Due to insufficient information on the properties of ZIKV and the global concern about microcephaly in newborn babies, this study was performed. Epitopic regions, secondary structure, and other structural characteristics of the envelope protein of ZIKV and two closely related viruses was predicted and compared. Herein, the series of user-friendly web-servers were used to provide the essential information for probable treatment of ZIKV. Our results reveal that ZIKV shares several similar properties with the other two viruses. Understanding the structural and functional features of ZIKV and comparing them with similar known viruses creates opportunities for finding effective treatment. In the present study, finding the effectively shared epitopes between the similar viruses could help in finding a common vaccine between these viruses. The *in vitro* and *in vivo* immunological tests should be performed in order to validate the suitability of the epitopes for vaccine development.

Acknowledgments

We acknowledge the support by University of Isfahan for this study.

References

1. Torresi J, Tapia-Conyer R, Margolis H. Preparing for dengue vaccine introduction: recommendations from the 1st Dengue v2V International Meeting. PLoS Negl Trop Dis. 2013;7(9):e2261.
2. Calvet G, Aguiar RS, Melo AS, Sampaio SA, de Filippis I, Fabri A, et al. Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. Lancet Infect Dis. 2016;16(6):653-660.
3. World Health Organization, 2008. Zoonoses and Veterinary Public Health (cited 25 September 2008). Available from http://www.who.int/zoonoses/en.
4. Mlakar J, Korva M, Tul N, Popović M, Poljšak-Prijatelj M, Mraz J, Kolenc M, et al. Zika virus associated with microcephaly. N Engl J Med. 2016;374(10):951-8.
5. Fagbami AH. Zika virus infections in Nigeria: virological and seroepidemiological investigations in Oyo State. J Hyg (Lond). 1979;83(2):213-9.
6. Robin Y, Mouchet J. [Serological and entomological study on yellow fever in Sierra Leone]. Bull Soc Pathol Exot Filiales. 1975; 68(3):249-58.
7. McCrae AW, Kirya BG. Yellow fever and Zika virus epizootics and enzootics in Uganda. Trans R Soc Trop Med Hyg. 1982;76(4):552-62.
8. Buathong R, Hermann L, Thaisomboonsuk B, Rutvisuttinunt W, Klungthong C, Chinnawirotpisan P, et al. Detection of Zika virus infection in Thailand, 2012-2014. Am J Trop Med Hyg. 2015;93(2):380-3.
9. Marchette NJ, Garcia R, Rudnick A. Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia. Am J Trop Med Hyg. 1969;18(3):411-5.
10. Wu KP, Wu CW, Tsao YP, Kuo TW, Lou YC, Lin CW, et al. Structural Basis of a Flavivirus Recognized by Its Neutralizing Antibody solution structure of the domain III of the japanese encephalitis virus envelope protein. J Biol Chem. 2003;278(46):46007-13.
11. Pattnaik P. Kyasanur forest disease: an epidemiological view in India. Rev Med Virol. 2006;16(3):151-65.
12. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Res. 2008;36(Web Server issue):W5-9.
13. Shen HB, Chou KC. PseAAC: a flexible web server for generating various kinds of protein pseudo amino acid composition. Anal Biochem. 2008;373(2):386-8.
14. Garnier J, Gibrat JF, Robson B, Doolittle R. GOR secondary structure prediction method version IV. Methods Enzymol. 1996;266:540-53.
15. Chen J, Liu H, Yang J, Chou KC. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino acids. 2007;33(3):423-8.
16. Guan P, Doytchinova IA, Zygouri C, Flower DR. MHCPred: bringing a quantitative dimension to the online prediction of MHC binding. Appl Bioinformatics. 2003;2(1):63-6.
17. Guan P, Hattotuwagama CK, Doytchinova IA, Flower DR. MHCPred 2.0: an updated quantitative T-cell epitope prediction server. Appl Bioinformatics. 2006;5(1):55-61.

18. Ferrè F, Clote P. DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res. 2005;33(suppl 2):W230-2.
19. Chauhan JS, Rao A, Raghava GP. In silico platform for prediction of N-, O-and C-glycosites in eukaryotic protein sequences. PloS one. 2013;8(6):e67008.
20. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30(12):2725-9.
21. Othman S, Rahman NA, Yusof R. All serotypes of dengue virus induce HLA-A2 major histocompatibility complex class I promoter activity in human liver cells. Trans R Soc Trop Med Hyg. 2010;104(12):806-8
22. Hughes HR, Crill WD, Davis BS, Chang GJ. A West Nile virus CD4 T cell epitope improves the immunogenicity of dengue virus serotype 2 vaccines. Virology. 2012;424(2):129-37.
23. Mohabatkar H. Computer-based comparison of structural features of envelope protein of Alkhurma hemorrhagic fever virus with the homologous proteins of two closest viruses. Protein Pept Lett. 2011;18(6):559-67.
24. Ni H, Burns NJ, Chang GJ, Zhang MJ, Wills MR, Trent DW, et al. Comparison of nucleotide and deduced amino acid sequence of the 5′ non-coding region and structural protein genes of the wild-type Japanese encephalitis virus strain SA14 and its attenuated vaccine derivatives. J Gen Virol. 1994;75(6):1505-10.
25. Rechkina EA, Denisova GF, Masalova OV, Lideman LF, Denisov DA, Lesnova EI, et al. [Epitope mapping of antigenic determinants of hepatitis C virus proteins by phage display]. Mol Biol (Mosk). 2006;40(2):357-68
26. Bian H, Reidhaar-Olson JF, Hammer J. The use of bioinformatics for identifying class II-restricted T-cell epitopes. Methods. 2003;29(3):299-309.
27. Mohabatkar H. Prediction of epitopes and structural properties of Iranian HPV-16 E6 by bioinformatics methods. Asian Pac J Cancer Prev. 2007;8(4):602-6.
28. Widman DG, Baric RS. Dengue virus envelope protein domain I/II hinge: a key target for dengue virus vaccine design?. Expert Rev Vaccines. 2015;14(1):5-8
29. Kulkarni R, Sapkal G, Mahishi L, Shil P, Gore MM. Design and characterization of polytope construct with multiple B and T H epitopes of Japanese encephalitis virus. Virus Res. 2012;166(1-2):77-86.
30. Bonaldo MC, Garratt RC, Caufour PS, Freire MS, Rodrigues MM, Nussenzweig RS, et al. Surface expression of an immunodominant malaria protein B cell epitope by yellow fever virus. J Mol Biol. 2002;315(4):873-85.
31. Vorovich MF, Timofeev AV, Mal'dov DG, Khaustov VI, El'bert LB. [Changes in biological properties of tick-borne encephalitis virus during cleavage of disulfide bonds in protein E]. Vopr Virusol. 1989;34(6):698-701.
32. Guirakhoo F, Heinz FX, Kunz C. Epitope model of tick-borne encephalitis virus envelope glycoprotein E: analysis of structural properties, role of carbohydrate side chain, and conformational changes occurring at acidic pH. Virology. 1989;169(1):90-9.
33. Zai J, Mei L, Wang C, Cao S, Fu ZF, Chen H, et al. N-glycosylation of the premembrane protein of Japanese encephalitis virus is critical for folding of the envelope protein and assembly of virus-like particles. Acta Virol. 2013;57(1):27-33.