A comprehensive analysis of predicted HLA binding peptides of JE viral proteins specific to north Indian isolates

Pawan Sharma1, 2, Komal Saxena2, Sanjay Mishra1 & Ajay Kumar2*

1School of Biotechnology, IFTM University, Lodhipur Rajput, Delhi Road (NH-24), Moradabad, 244001, U.P., India; 2Institute of Biomedical Education and Research, Manglayatan University, Aligarh, 202145, India; Ajay Kumar – E mail: ajaymtech@gmail.com; Phone: +91-9412883081; *Corresponding author

Received April 14, 2014; Revised May 13, 2014; Accepted May 24, 2014; Published June 30, 2014

Abstract:
Japanese encephalitis (JE), a viral disease has significantly increased worldwide especially, in the developing region due to challenges in immunization, vector control and lack of appropriate treatment methods. An effective, yet an expensive heat-killed vaccine is available for the disease. Therefore, the design and development of short peptide vaccine candidate is promising. We used immune-informatics methods to perform a comprehensive analysis of the entire JEV proteome of north Indian isolate to identify the conserved peptides binding known specific HLA alleles among the documented JEV genotypes 1, 2, 3, 4 and 5. The prediction analysis identified 102 class I (using propred I) and 118 class II (using propred) binding peptides at 4% threshold value. These predicted HLA allele binding peptides were further analyzed for potential conserved region using IEDB (an immune epitope database and analysis resource). This analysis shows that 78.81% of class II (in genotype 2) and 76.47% of HLA I (in genotype 3) bound peptides are conserved. The peptides IPIVSASL, KGAQLAAAL, LAVFLICVL and FRTLFGGMS, VFLICVLTV, are top ranking with potential super antigenic property by binding to all HLA allele members of B7 and DR4 super-types, respectively. This data finds application in the design and development of short peptide vaccine candidates and diagnostic agents for JE following adequate validation and verification.

Keywords: Immunoinformatics, epitope, short peptide, supertype, vaccine.

Background:
Japanese encephalitis (JE) is a major viral disease of human beings in developing countries. JE virus causes membrane inflammation of the brain and leads to deleterious effects on Central Nervous System (CNS). JE virus is a single stranded RNA virus that belongs to the genus Flavivirus of family flaviviridae [1]. Flavivirus genus comprises of several other human pathogenic viruses such as Saint Louis virus, Dengue virus, Yellow fever virus and West Nile virus. Positive sense RNA genome of JE virus has 11,000 nucleotides, which encode a single polypeptide of 3432 amino acids. The virus has three structural proteins: Capsid, precursor membrane protein and envelope protein and seven nonstructural proteins: NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5 [2].

JE is a mosquito born viral disease, Culex tritaeniorhynchus and C. visnui mosquitoes are major vectors for JE virus transmission in human beings. JE virus affects 50,000 people in Asia with an annual fatality rate of 5-35% [3, 4, 5]. The most severe known epidemic of JE was reported in the Gorakhpur, northern region
of India in 2005, which affected 5,737 lives and 1344 deaths [6]. JE disease is characterized by several primary and secondary clinical symptoms such as brain membrane inflammation, continuous fever cause irreversible neuron damage, psychiatric and neurological disorder with limb paralysis etc. [4, 5, 7]. JE virus has five known genotypes, which are distributed in various worldwide geographical areas Table 1 (see supplementary material) [8, 9, 10, 11, 12].

Nakayama JE virus strain is widely used in JE vaccine that belongs to genotype III. Genotype III is the most widely distributed genotype and it is the only genotype isolated from the Indian subcontinent. Furthermore, The JE disease burden is increasing day by day in developing countries due to the impracticality of immunization, vector control methods and lack of therapeutic treatment [2, 13, 14]. As a result, vaccination is the only way to prevent JE [15]. In present scenario, a number of vaccines have been developed in several countries, but only inactivated mouse brain derived Nakayama strain vaccine is the most commercially used vaccine [16, 17]. Now-a-days, Nakayama strain vaccine has been replaced by Vero cell derived JE vaccine (IXIARO) which can effectively boost the immunity [18, 19]. There are various drawbacks of this vaccine such as vaccine production shortage, high cost and neurological adverse effects especially in low-income countries, which increase the disease burden of JE with time [17, 20, 21, 22].

Among all available JE vaccines, an epitope vaccine is more potent than killed, attenuated and cell cultured derived vaccines, gives better immunity and devoid of adverse effects of entire viral proteins [23, 24]. The majority of available current vaccines have involvement of only structural proteins but non-structural proteins cannot be ignored. As reported earlier, nonstructural proteins are produced in live virus forms, show a good immune response [25] and can work as a major target for human anti JEV specific T cells produced during natural infections [26, 27].

The development of epitopes based vaccines generally requires the knowledge of the adaptive immune system. T_{H} cells and T_{c} cells can recognize antigen when bounded with MHC class II and I molecule, respectively [28, 29]. Major Histocompatibility Complex (MHC) which is also known as Human Leukocyte Antigen (HLA) in humans is a membrane glycoprotein and extremely polymorphic in nature. These HLA molecules can bind to a spectrum of antigenic linear epitopes derived from antigen processing, which initiate an immune response, but HLA binding does not assure to generate T-cell immune response alone. The peptide binding specificity varies for different HLA alleles in a combinatorial manner among ethnic populations. It has been reported that the majority of alleles can be covered within few HLA supertypes, where different members of a supertype bind similar peptides; these similar peptides are called super antigens. Recently, nine major HLA class I supertypes (HLA- A1, A2, A3, A24, B7, B27, B44, B58, and B62 and seven HLA class II supertypes (main DR, DR4, DRB3, main DQ, DQ7, main DP, and DP2) have been determined by comparing peptide-binding data [30, 31]. Peptides exhibiting super antigenic property by binding to a maximum number of HLA alleles or HLA supertypes with their conserved nature can surmount the problem of HLA allele’s population coverage and chance of antigen escape related to antigenic drift or shift. Therefore, the present study is designed for a comprehensive analysis of predicted HLA binding peptides of JE viral proteins specific to north Indian isolates.

Methodology:
The complete genome and protein sequences of JEV of north Indian origin strain (Accession No.ABU94628) were obtained from sequence database NCBI (http://www.ncbi.nlm.nih.gov/entrez). DDBJ database was used to calculate the number of adenine, cytosine, guanine and thymine bases in the genome. The physiochemical properties of all viral proteins such as iteration of amino acids within proteins, their molecular weight and pI value of predicted epitopes were analyzed using proteomics analysis platform of ExPasy [32]. In addition, the variation and conservation of envelope protein residues in all five genotypes, were done by using a protein variability server at 0.46 threshold value of Simpson diversity. The envelope protein of SA14-14-2 strain (PDB ID- 3PS4) was taken as a base structure to map the variable and conserved regions in genotypes 1,2,3,4 and 5 [33]. The flowchart of methodology has been represented in (Figure 1).

**Figure 1:** Flowchart of methodology employed in comprehensive analysis of predicted HLA binding peptides of JE viral proteins.

**Screening of T cell epitopes**
All the structural and non-structural proteins of JEV (Accession No.ABU94628) were analyzed for screening of possible dominant T cell epitopes using immunoinformatics tools such as Propred I and Propred. Propred I and Propred were used at 4% threshold for binding analysis of all possible peptides to 47 class I and 51 class II HLA alleles respectively. These tools are highly valuable to recognize antigenic HLA binding peptides [34, 35].

**Predicted T cell epitopes worldwide conserved region study**
All the predicted T cell epitopes of JEV north Indian origin strain, were undergone for worldwide conserved region study among JEV genotypes 1, 2, 3, 4 and 5. Before conserved region...
study it is necessary to retrieve all proteins sequences of all genotypes from NCBI database. Therefore maximum 5 sequences of each JEV protein were retrieved from the NCBI database randomly for genotypes 1, 2, 3, 4 and 5.

The predicted T cell epitopes of each protein of JEV strain along with 1 to 5 same protein sequences of a single genotype were taken to Immune Epitope Database and Analysis Resource (IEDB) conservancy tool [36]. This cycle was repeated for all five genotypes for all proteins of JEV strain.

The nanomer T cell epitopes having 70 - 100% conserved region with a maximum single and double mutation were selected while discarded the epitopes having less than 70% conserved region with more than two mutations. After conserved region analysis, isoelectric point (pl) value of predicted peptides was calculated for all mutated and conserved epitopes [37].

Figure 2: Plot shows variable region, having variability score more than 0.46 threshold value in envelope protein by Simpson variability method. Peaks marked with stars are depicting the variable residues (129, 222, 327 and 369).

Figure 3: Variability and conserved region of envelope protein shown by red and blue gradient respectively onto three dimensional base structure of envelope protein (3P54).

Result & Discussion:

The result of the study indicated that JEV genome comprises 10976 base pairs with GC content 51.35 %. The GC content was found to be 2.7% higher than AT contents. The genome translates into a polyprotein that afterwards separated into structural and non-structural proteins. The structural envelope protein has the highest molecular weight 52975.81 kDa. All protein physiological properties such as molecular weight, amino acid number and frequency amino acid were listed in Table 2 (see supplementary material). Frequency of amino acid in protein is directly associated with the pl value and their binding with HLA alleles. The study of envelope protein variability among all genotypes, the amino acid sequence positions 129, 222, 327 and 369 were observed with high Simpson variability (Figure 2) which was also shown in 3D mapped structure (Figure 3).

Total 118 HLA class II binding T cell epitopes were extracted by propred tool Table 3 (see supplementary material). The highest number of T cell epitopes was represented by envelope protein comprising 28.81% of all predicted HLA class II epitopes. Envelope protein predicted epitopes such as LVTVNPFWA, VGLYTVNP, FRTLFGGMS, LKGAQRLAA and FNSIGKAVH were found to be potential binders of 20-50 HLA II alleles.

Envelope protein followed by NS1 and NS2A proteins representing 12.71% and 11.86% of the predicted HLA II epitopes respectively. In case of non-structural proteins (NS1, NS2A, NS2B, NS4A, NS4B and NS5) LWGDGVES, FVHNDVEAW, FGITSTRVW, YVELVAAAQ, FMLAGLMAV, VFLICVLT, LLMVVFLIP, LVFLGCWGQ, LTAAATLTL, and VVLTLPLLH were predicted as the most potential binders in term of binding score, conserved nature and the HLA II alleles coverage.

In case of HLA class I binding T cell epitopes, total 102 epitopes were extracted using propred I (Table 3). Again, the highest number of T cell epitopes was represented by envelope protein comprising 23.51% of all the predicted HLA I epitopes. Envelope protein epitopes such as QALAGAIVV, GHGTTVIYL, KGAAQRALAA and TTLKGAQRL are the potential binders to range of 20 - 50 HLA I alleles. In case of non-structural proteins (NS1, NS2A, NS2B, NS4A, NS4B, NS5) KSILFAPEL, YLPETPRSL, LMFAIVGGL, LKKENAVDL, AVVGGLAE1, LALLMVVL1, AVLGLALLV, LMFAMIVGGL, IAGTLILIAL, LAVFLICVL, KATGSASSL, FMWLGARYL peptides were predicted as best binders in the term of binding score, conserved region and the HLA I alleles coverage.

The conserved region analysis of total 118 predicted HLA class II binding epitopes, 29 epitopes were found to be 100% conserved in all genotypes. The 118 predicted HLA II peptides showed 72 % conserved nature with genotype I, 78.81% with genotype II, 75 % with genotype III, 54% with genotype IV and 39.83% with genotype V (Figure 4). Predicted HLA II binding epitopes were found highly conserved in genotype II (78.81%). Similarly, the conserved region analysis of total 102 predicted HLA class I binding epitopes, 21 epitopes were found 100% conserved in all genotypes. The 102 predicted HLA I peptides showed 70.58% conserved nature with genotype I, 75.49% with genotype II, 76.47% with genotype III, 62.47% with genotype IV and 52% with genotype V (Figure 4). Predicted HLA I binding epitopes were found highly conserved in genotype III (76.47%). LMTNHNTDI, MINIEASQ, LVTVNPFWA, IP1YSVASC, and VLTLATFFL epitopes were found as common binders for HLA class I and II alleles. LVTVNPFWA epitope of envelope protein

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 10(6): 334-341 (2014) 336 © 2014 Biomedical Informatics

open access
BIOINFORMATION
open access

was found best binder in term of the HLA allele coverage with 100% conserve nature in all genotypes.

As discussed earlier, the concept of HLA supertype has a profound role in the understanding of T cell epitope selection, degeneration and discrimination during T cell mediated immune response [30]. In the HLA supertype analysis, IEDB web server was also used to check binding of best epitopes with also those HLA alleles, which are not included in propped server. For an example, DR4 HLA II supertype members such as DRB1*0401, 0405 and 0802 are not available in propped server. Findings revealed that LVTNPFFVA, IPIVSVASL, KGAQRLAAL, LAVFLICVL epitope bonding to all members of B7 HLA I supertype (B*0702, B*3501, B*5101, B*5102, B*5301, B*5401) but these peptides also show selective bonding to some members but not all members of the other HLA I supertypes. FRTLFGGMS, VFLICVLTV epitope were binding to all members of DR4 HLA II supertype (DRB1*0401, 0405 and 0802) but not all members of the other HLA II supertypes. Therefore LVTNPFFVA, IPIVSVASL, KGAQRLAAL, LAVFLICVL, FRTLFGGMS and VFLICVLTV epitope were also showing their super antigenic property. These predicted potential novel epitopes are sufficient to work as vaccine rather than using whole proteins as vaccine candidates because it has been confirmed few epitopes can represent complete antigenicity of any protein [23]. Similar to this study, epitope based vaccines have given promising result against several highly infectious diseases such as HIV1, HIV and Tuberculosis [24, 38, 39]. Thus in the present study, propped I and propped server were used for screening of best T cell epitopes from proteome of JEV north Indian isolate followed by worldwide conserved region analysis in all genotypes (1,2,3,4 and 5). The predicted epitopes were nanomers and could be used as vaccine candidates and diagnostic reagents for JEV.

Identification of class I and class II HLA specific JE viral peptides at 4% threshold value by using Propred I and Propred, respectively. We report the presence of 29 class II and 21 class I specific conserved peptides in all known genotypes. The HLA specific predicated are seem to be highly conserved in genotypes 2 and 3, while limited in 1, 4 and 5. We further found that the peptides IPIVSVASL, KGAQRLAAL, LAVFLICVL and FRTLFGGMS, VFLICVLTV, are top ranking with potential super antigenic property by binding to all HLA allele members of B7 and DR4 super-types, respectively. This data finds application in the design and development of short peptide vaccine candidates and diagnostic agents for JE following adequate validation and verification.

Acknowledgement:
The present paper is concerned with a piece of Ph.D. program of the author Pawan Sharma (M.Tech, Anna University, Chennai, India). The authors are grateful to Prof. (Dr.) S.K. Maheshwari and Dr. Brijesh Sharma for time-to-time suggestions and inspiration enabling them to successfully accomplish the write-up of this manuscript. The authors are also thankful to Director and all faculty members of the Institute of Biomedical Education and Research, Mangalayatan University Aligarh, U.P, India for providing necessary facilities and encouragement.

References:
[1] Endy TP & Nisalak A, Curr Top Microbiol Immunol. 2002 267: 11 [PMID: 12082986]
[2] Schiöler KL et al. Cochrane Database Syst Rev. 2007 18: CD004263 [PMID: 17636750]
[3] Kumar R et al. Indian J Med Res. 1993 97: 9 [PMID: 8387460]
[4] Schneider RJ et al. Southeast Asian J Trop Med Public Health. 1974 5: 560 [PMID: 4376275]
[5] Tsai TF et al. Vaccines 1994 2: 671 [DOI: 10.1371/journal.pntd.0001952]
[6] Parida M et al. Emerg Infect Dis. 2006 12: 1427 [PMID: 17073905]
[7] Solomon T & Vaughn DW, Curr Top Microbiol Immunol. 2002 267: 171 [PMID: 12082989]
[8] Chen WR et al. J Gen Virol. 1990 71: 2915 [PMID: 2273991]
[9] Chen WR et al. Am J Trop Med Hyg. 1992 47: 61 [PMID: 1322071]
[10] Williams DT, J Gen Virol. 2000 81: 2471 [PMID: 10993935]
[11] Ni H & Barrett AD, J Gen Virol. 1995 76: 401 [PMID: 7844559]
[12] Uchil PD & Satchidanandam V, Am J Trop Med Hyg. 2001 65: 242 [PMID: 11561712]
[13] Hoke CH et al. Infect Dis. 1992 165: 631 [PMID: 1313068]
[14] Igarashi A, Curr Top Microbiol Immunol. 2002 267: 139 [PMID: 12082987]
[15] Tsai TF, Vaccine 2000 18: 1 [PMID: 10821969]
[16] Barrett AD, Biologicals 1997 25: 27 [PMID: 9167006]
[17] Monath TP, Curr Top Microbiol Immunol. 2002 267: 105 [PMID: 12082985]
[18] Erra EO et al. Vaccine 2013 32: 119 [PMID: 24176496]
[19] Erra EO et al. Clin Infect Dis. 2012 55: 825 [PMID: 22696017]
[20] http://www.cdc.gov/mmwr/preview/mmwrhtml/00020596.htm
[21] Shlim DR & Solomon T, Clin Infect Dis. 2002 35: 183 [PMID: 12087525]
[22] Okabe N, Uirusu 2005 55: 303 [PMID: 16557017]
[23] De Groot AS et al. Immunol Cell Biol. 2002 80: 255 [PMID: 12067413]
[24] Sharma P & Kumar A, J Proteomics Bioinform. 2010 3: 275 [DOI:10.4172/jpb.1000151]
[25] Konishi E et al. Am J Trop Med Hyg. 1995 53: 278 [PMID: 7573713]
[26] Desai A et al. Microbiol Immunol. 1995 39: 269 [PMID: 7651240]
[27] Kumar P et al. Arch Virol. 2003 148: 1569 [PMID: 12898332]
[28] Reche PA & Reinherz EL, Nucleic Acids Res. 2005 33: W138 [PMID: 15980443]
[29] Zhang GL et al. Nucleic Acids Res. 2005 33: W172 [PMID: 15980449]
[30] Kangueane P et al. Front Biosci. 2005 1: 879 [PMID: 15569626]
[31] Greenbaum J et al. Immunogenetics. 2011 63: 325 [PMID: 21305276]
[32] http://web.expasy.org/protparam/
[33] http://imed.med.ucm.es/PVS/
[34] Singh H & Raghava GP, Bioinformatics 2003 19: 1009 [PMID: 12761064]
[35] Singh H & Raghava GPS, Bioinformatics 2001 17: 1236 [PMID: 11751237]
[36] http://tools.immuneepitope.org/tools/conserved_region/iedb_input.
[37] Khan MA et al. Cell Immunol. 2006 244: 141 [PMID: 17434154]
[38] Jardine J et al. Science 2013 340: 711 [PMID: 23539181]
[39] Feng G et al. PLoS ONE. 2013 8: e61135 [PMID: 23637790]

Edited by P Kangueane

Citation: Sharma et al. Bioinformation 10(6): 334-341 (2014)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.
**BIOINFORMATION**

**Supplementary material:**

**Table 1:** All JEV genotypes distribution in worldwide geographical regions

| No | Genotype   | Geographical regions of isolation                                      |
|----|------------|-----------------------------------------------------------------------|
| 1  | Genotype I | Cambodia, Northern Thailand and Korea                                  |
| 2  | Genotype II| Southern Thailand, Indonesia, Malaysia and Australia                   |
| 3  | Genotype III| Asian countries such as Japan, China, India, Taiwan and Sri Lanka      |
| 4  | Genotype IV| Indonesia                                                             |
| 5  | Genotype V | Malaysia                                                              |

**Table 2:** Dataset statistics and features of the JEV Genome and Proteome such as molecular weight, amino acid number and percentage of highly and least iterative amino acid residues in individual proteins.

| Bases   | No.  | Proteins   | aa number | Mol. Wt. | Highly iterative | % iterative of | Least iterative | % of iterative |
|---------|------|------------|-----------|----------|-----------------|---------------|----------------|---------------|
| Total bp| 10976| Capsid     | 118       | 12940.5  | L               | 12.7          | H/Q            | 0.8           |
| A       | 3019 | Flavi Prep | 75        | 9846.46  | D               | 10.7          | F/S            | 1.3           |
| C       | 2505 | Flavi M    | 71        | 7925.08  | L               | 14.1          | D/H/M          | 1.4           |
| G       | 3132 | Flavi glycoprot | 298     | 32281.16 | G               | 9.4           | W              | 1.7           |
| T       | 2320 | Flavi glycoprot | 99     | 10636.01 | V               | 12.1          | Q/W            | 1.0           |
|         |      | Flavi E stem | 97      | 10058.64 | G               | 16.5          | H              | 2.1           |
|         |      | NS1        | 355       | 40340.29 | E               | 9.0           | M/Q            | 1.7           |
|         |      | NS2A       | 216       | 23265.84 | L               | 16.7          | C/H            | 0.9           |
|         |      | NS2B       | 128       | 13968.14 | A               | 12.5          | C/H            | 0.9           |
|         |      | NS4A       | 144       | 15616.72 | L               | 16.0          | C/N/W/Y        | 0.7           |
|         |      | NS4B       | 248       | 26174.31 | L               | 14.1          | C              | 0.8           |
|         |      | NS5        | 649       | 74365.25 | E               | 8.5           | C              | 1.4           |

**Table 3:** Top ranking predicted HLA binding peptides of JE viral proteins specific to north Indian isolates

| S.N0. | Epitope position | Predicted T cell Epitopes | pI Value | HLA Alleles, Class | JEV Genotypes |
|-------|------------------|--------------------------|----------|--------------------|---------------|
| Capsid|                  |                          |          |                    |               |
| 1     | 41-49            | FVLALITFF                | 5.52     | 35, II             | FVLALI(TG3,VG4,LA,G5)T(SG4,AG5)FF |
| 2     | 46-54            | ITFFKFTAL                | 8.75     | 40, I              | I(TG3D,GADLG5T(SG4AG5)FF |
| 3     | 52-60            | TALAPTKAL                | 8.41     | 33, I              | TALA(SG5)PTKAL |
| Flavi prep|                |                          |          |                    |               |
| 4     | 7-15             | LMITNNTDI                | 3.80     | 24, II             | LMT(A9)I(VG2,G3,G4)NNTDI(VG5) |
| 5     | 2-10             | FQGKLLMTI                | 8.75     | 47, I              | FQGKLL(V5G5)MT(A9)I(VG2,G3,G4) |
| 6     | 3-11             | QKGLLMTN                 | 8.75     | 46, I              | QKGLL(V5G5)MT(A9)I(VG2,G3,G4)N |
| 7     | 4-12             | GKLLMNTN                 | 8.75     | 46, I              | GKL(V5G5)MT(A9)I(VG2,G3,G4)NN |
| Flavi M|                |                          |          |                    |               |
| 8     | 62-70            | LLLLVAPAY                | 5.52     | 38, II             | LLLLVAPAY*    |
| 9     | 61-69            | ILLVVAPA                 | 5.52     | 45, II             | ILLLVAPA*     |
| 10    | 54-62            | GQRVVFHTIL               | 9.75     | 45, I              | GQ(PG5)RVVFHTIL |
| Flavi glycoprot|            |                          |          |                    |               |
| 11    | 45-53            | MINIEASQL                | 4.00     | 11, II             | MINIEAS(VG3,TG9)QL |
| 12    | 43-51            | VRMINIEAS                | 5.97     | 46, II             | VRMINIEAS(VG3,TG9) |
| 13    | 45-53            | MINIEASQL                | 4.00     | 18, I              | MINIEAS(V5G7)GQ |
| 14    | 172-180          | NAPSITKL                 | 8.75     | 34, I              | NAPS(TG0)ITKL |
| 15    | 264-272          | QALAGAIVY                | 5.52     | 27, I              | QAGALAIVY*    |
| 16    | 170-178          | TPNAPSITL                | 5.19     | 33, I              | TPNAPS(TG0)ITL |
| Flavi glycoprot c|       |                          |          |                    |               |
| 17    | 23-31            | VVIELSYSG                | 4.00     | 35, II             | VVIELS(TG3,G2,LA,G4,G5YS(TG5)G |
| 18    | 38-46            | IPIVSVASL                | 5.52     | 43, II             | IPIV(SG5)SVASL |
| 19    | 55-63            | LVTVNPFFVA               | 5.52     | 48, II             | LVTVNPFFVA*   |
| 20    | 58-66            | VNPFFVAASS               | 5.49     | 34, II             | VNPFFVAASS(TG2,G3,G4,G5)SS(TG5) |
| 21    | 52-60            | VGRVLVTNVP               | 9.72     | 21, II             | VGRVLVTNVP*   |
| 22    | 38-46            | IPIVSVASL                | 5.52     | 43, I              | IPIV(SG5)SVASL |
| 23    | 55-63            | LVTVNPFFVA               | 5.52     | 48, I              | LVTVNPFFVA*   |
| 24    | 19-27            | GHGTVVIEL                | 5.24     | 25, I              | GHGTVVIEL*    |
Flavi E Stem

25. 46-54  FRTLFGGMS  9.75  29, II  FRTLFGGMS*
26. 88-96  LVFLATNVH  6.74  49, II  LV(LGLG)VFLATNVH
27. 60-68  LMGALLLWLM  5.52  32, II  LMGA(VG4)LLLWLM
28. 8-16  LKGAQRQLAA  11.00  34, II  LKGAQRQLAA*
29. 79-87  LAFATGGV  5.52  11, II  LAFATGGV*
30. 31-39  FNSIGKAVH  8.76  26, II  FNSIGKAVH*
31. 50-58  FGGMSWITQ  5.52  15, II  FGGMSWITQ*
32. 83-91  ATGGLVLFL  5.57  35, I  A(VG4)GTLGTV(TG4)LV(LGLG)FL
33. 81-89  FLATGGVLV  5.52  33, I  FLA(VG4)GLGTV(TG4)LV(LGLG)FL
34. 6-14  TTLKGAQRQL  11.00  33, I  TTLKGAQRQL*
35. 9-17  KGAQRQLAL  11.00  29, I  KGAQRQLAL*

Flavi N51
36. 230-238  LWGDGVEES  3.57  10, II  LWGDGVEES*
37. 56-64  VRSVTRELH  9.58  34, II  V(RG2)RSVTRELH
38. 19-27  FVHNDVEAW  4.35  10, II  FVHNDVEAW*
39. 159-167  FGITSTRWV  9.75  10, II  FGITSTRWV*
40. 119-127  KSILFAPEL  6.00  36, I  KSI(LG4)LSFAPEL
41. 33-41  YLPETPRSL  6.00  34, I  YLPETPRK(GG4)SG4LS
42. 77-85  LLKENAVDL  4.37  33, I  LL(LG4)KLENADL

Flavi N52A
43. 72-80  FKIQPAPLFL  8.75  13, II  FKIQPAPLFL*
44. 46-54  YVVLVAAMAF  5.52  20, II  YVVLVAAMAF*
45. 13-21  LRRKWTRAL  12.30  10, II  LRRKWTRAL*
46. 25-33  AVLGLLVL  5.57  15, I  AVLGLLVL*
47. 83-91  VLTLATFL  5.49  13, I  VLT(VG4)LATFL

Flavi N52B
48. 96-104  VLRMSCIGL  8.22  31, II  V(LG4)LMSCIGL
49. 41-49  VVSGBKATDM*  5.81  15, II  VVSGBKATDM*
50. 30-38  FMLAGLMAV*  5.52  32, II  FMLAGLMAV*
51. 10-18  LMAIVGGL  5.52  14, II  LMAIVGGL*
52. 120-128  WLTLLTTR  11.17  14, II  WLTLLTTRK*
53. 117-125  FYGWLTLLT  8.59  24, II  FYGWLTLLT*
54. 10-18  LMAIVGGL  5.52  25, I  LMAIVGGL*
55. 13-21  GIVGGLAL  4.00  31, I  GIVGGL(ML4G)

Flavi N54A
56. 66-74  LLMQRKGI  11.00  39, II  LLM(KG4)QRKGI
57. 132-140  VFICLVLTV  5.49  46, II  VFICLVLTV*
58. 54-62  IVAITVMGTG  5.52  40, II  IV(VG4)TVMGTG
59. 109-117  LLLLMMVLP  5.52  43, II  LLLLMMVLP*
60. 76-84  KMGLGALVL*  8.75  21, I  KMGLGALVL*
61. 107-115  IALLLMVVL  5.52  29, I  I(LG4)ALLLMVVL
62. 101-109  IAGTLLIAL  5.52  28, I  IAGTLLI(LG4)LAL
63. 130-138  LAVFLICVL  5.52  27, I  LAVFLICVL*
64. 98-106  GTKIAGTL  8.75  12, I  GTKIAGTL*

Flavi N54B
65. 36-44  LRPATAWAL  9.75  19, II  LRPATAWAL*
66. 97-105  LVFLGCGWQ  5.52  37, II  LVFLGCGWQ*
67. 83-91  VLTLATFFL  5.49  13, II  VLT(VG4)LATFFL
68. 201-209  LVAAATGTL  5.52  39, II  LVAAATGTL*
69. 122-130  YGMYLPGQW*  5.52  16, II  YGMYLPGQW*
70. 50-58  VVLTPLKH  8.73  31, II  VVLTPLKH*
71. 154-162  MVATDVPEL  3.67  9, I  MVATDVPEL*
72. 48-56  STVVLTPLL  5.24  19, I  STVVLTPLL*
73. 92-100  DLTGLVLFL  3.80  13, I  DLTGLVLFL*
74. 232-240  GSYLAGGSI  5.52  11, I  GSYLAGGSI*
75. 126-134  LGPKWQAEAL  4.00  14, I  LGPKWQAEAL*
76. 89-97  TDLDDLTVGL  3.56  11, I  TDLDDLTVGL*

Flavi N55
77. 470-478  VMKDRGSRIV  8.72  8, II  VMKDRGSRIV(LG4G)V(LG4)
78. 71-79  VNGVKKLMS  8.72  14, II  VNGVKKLMS*
79. 62-70  KATGSASSL  8.75  23, I  KATGSASSL*
Interpretation for Table 3:

| 80. | 226-234 | FMWLGARYL | 8.75 | 20, I | FMWLGARYL* |

Bold amino acid in T cell epitopes column indicates the anchor residues.
In JEV genotype column, bold amino acid residues with superscript genotypes as G1, G2, G3, G4 and G5 in brackets ( ) show varied amino acid in genotypes.
D as superscript in genotype column indicated change in pI value of peptide in superscript genotype.
Bold and * peptides shown in Genotype column are fully conserved peptides in all genotypes.
Interpretation example for JEV genotype column: A(V_{G4}D)TGGV(T_{G4})LV(L_{G4,G5})FL, here original peptide is ATGGVFL. Alanine (A) first amino acid is replaced by Valine (V) in genotype 4 with change pI. Fifth amino acid Valine (V) is replaced by Threonine (T) in genotype 4. Seventh amino acid of original peptide, Valine is replaced by Leucine in genotype 4 and Genotype 5.