Top Down Computational Approach: A Vaccine Development Step to Find Novel Superantigenic HLA Binding Epitopes from Dengue Virus Proteome

Priti Sharma[1] · Pawan Sharma[2] · Sheeba[1] · Ajay Kumar[3]

Abstract

Dengue virus (DENV) is a major mosquito vector based human pathogenic flavivirus which is causing major threat worldwide, yet the availability of therapeutic treatment and several vaccines, still called for advance treatment and vaccine development. The present top down computational approach is a vaccine development step to find novel super antigenic HLA binding epitopes from DENV proteome. The approach used sequence based screening to find complete conserve and high population coverage, common epitopes among all DENV serotype. Propred and Immune Epitope Data Base were used for sequence based screening with recommended parameters. Among top 29 identified epitopes, five structural protein epitopes viz. ^{33}LQGRGPLKL^{41}, ^{249}VVVLGSQEG^{257}, ^{172}LVGIVTLYL^{180}, ^{146}MKILIGVVI^{154}, ^{72}YIIVGVEPG^{80} and one nonstructural protein epitope ^{18}LKNDIPMTG^{26} were showed high conserve nature and high population coverage from complete DENV proteome. Further structure based study involving docking and molecular dynamic simulation to confirm stable behavior of HLA allele–peptide complex to give potent cell mediated immune response. Docking of epitope ^{72}YIIVGVEPG^{80}–DRB1 0401 allele and epitope ^{33}LQGRGPLKL^{41}–B*5101 allele complexes showed the best binding energy of − 7.71 and − 7.20 kcal/mol, respectively and stable binding pattern over the time window during molecular dynamic simulation. This computational approach resulted novel epitopes which can be used in the design and development of short epitope based vaccines as well as diagnosis tools for dengue infection.

Keywords Computational · Dengue · Diagnosis · Epitope · Serotype · Vaccine

Introduction

Dengue virus (DENV) is a major *Aedes* mosquito vector based human pathogenic flavivirus which is causing worldwide high morbidity and mortality in recent times. Dengue (DENV) flavivirus belongs to family flaviviridae (Pulmanausahakul et al. 2010). DENV infection cause dengue fever with more severe forms Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS) with the help of vector *Aedes* (*Aedes albopictus* and *Aedes aegypti*) mosquitoes (Sanyaolu et al. 2017). Day by day, dengue becoming a major health problem because of some major factors such as population growth, urbanization and difficult mosquito vector control. All together, these factors increasing the cases of dengue infections all over the world (Gubler 2004). WHO reported the number of cases of dengue infection increased over 2.4 million to 4.2 million from 2010 to 2019 and the largest number of dengue cases ever reported globally in 2019 but therapeutic treatment and vaccine still under development. Therapeutic drugs refer Chimeric Yellow fever virus DENV Tetravalent Dengue Vaccine (CYD-TDV) and available conventional vaccines don't have potential to cope with this global DENV infection problem (WHO 2020). Dengue virus proteome comprises total ten proteins such as core, glycoprot and envelope as structural proteins and nonstructural (NS1, 2a, 2b, 3, 4a, 4b and 5) proteins (Guzman et al. 2010; Schiøler et al. 2007). The virus has four

✉ Ajay Kumar
ajaymtech@gmail.com

1 D. S. Degree College, Aligarh, Dr. B. R. Ambedkar University, Agra 282004, India

2 Institute of Engineering and Technology, Mangalayatan University, Aligarh 202145, India

3 Department of Biotechnology, Faculty of Engineering &Technology, Rama University, G.T. Road, Kanpur 209217, India

immunogenically related serotypes DEN-1, DEN-2, DEN-3, and DEN-4 (Pulmanausahakul et al. 2010).

A lot of research has been done to develop a successful dengue vaccine and therapeutic drugs which are yet ongoing. Still, there is no specific antiviral drug and vaccine available for dengue. CYD-TDV is mostly tested vaccine and licensed by several endemic countries. Apart this, several other vaccines such as tetravalent TV003/TV005 and DENVax vaccines also in phase IIIrd and IInd under vaccine development, but these vaccines have problems in efficacy which varied by age groups, serotype and severity of dengue disease (Vannice et al. 2016). Dengue infection with one serotype gives lifelong immunity to same serotype, but reinfection with hetro-serotype increased risk of severe dengue, which is a major problem in dengue vaccine development known as an Antibody Dependent Enhancement (ADE). During the course of infection, developed antibodies by one DENV serotype can recognize other serotype at time of reinfection but these antibodies are not specific to other serotype therefore these antibodies act as non neutralizing antibody (Pinheiro-Michelsen et al. 2020). During this, antigen and antibody complex is recognized by Fc receptors of phagocytic cells, which facilitate entry of different DENV serotype virus (Olsen et al. 2011).

Indeed, this ADE problem can be surmounted by identifying the potent peptide epitopes which are completely conserved among all DENV serotypes. This cross-reactivity and cross-protection is particularly significant in the case of DENV infection, so that immunity should be develop against all DENV serotype which require specific targeted common and conserve epitope regions from different geographical DENV serotypes (Williams et al. 2009; Kirsten et al. 2015). Epitope based vaccines are superior over traditional vaccine and also can overcome problems of side effects, specificity, ADE (Srivastava et al. 2020). The epitope alone does not sufficient to initiate an immune response; it has to bind to host Human Leucocyte Antigen (HLA). Due to high polymorphic nature of HLA in human, population coverage is a major issue in global vaccine development (Gupta and Kumar 2020a, b). The HLA molecules bind to a range of immunogenic sequential peptide epitopes through antigen processing and presentation, which induce ultimately helper T cell to release cytokines. These cytokines activate both Tc cells and B cells to give cell mediated and humoral immune response, respectively (Chaplin 2010). The epitope binding specificity to different alleles shows variations with different geographical populations. The combinations of few HLA alleles can represent the one geographical population called HLA Supertypes and the name of superantigen is given to those peptides which bind to several alleles of HLA supertype. Recently, Japanese encephalitis and West Nile virus epitope based vaccine development

studies also showed supertype analysis can help to find novel vaccine candidates (Sharma et al. 2014). Under the concept of HLA supertype, the screened conserved and superantigenic T Cell epitopes resolved the problem of population coverage (Kangueane et al. 2005).

The present top down computational approach is a vaccine development step to find novel super antigenic HLA binding epitopes which is an urgent need in present time which reduce time as well as experimental burden (Sharma et al. 2018). The similar vaccine development approaches have already given good results against many other infectious agents such as Chikungunya, Campylobacter vaccine development, recently (Saxena and Mishra 2020; Gupta and Kumar 2020a, b). The DENV proteome several fold sequence based screening resulted complete conserve, high population coverage, superantigenic epitopes. Consequently, these epitopes and alleles were structurally modeled for their docking study. The epitope peptide and respective HLA allele complexes having least binding energy were explored for molecular dynamic simulation to analyze binding stability. Thus, the current approach to epitope-based vaccine development is more dominant over traditional vaccines with respect to side effects, doses, cost and high production.

Methodology

The complete genome and protein sequence of Dengue Virus (DENV) was obtained from the NCBI sequence database, having accession Number AFZ40225 (Sayers et al. 2009). The genome and proteomics analysis of DENV were done by DDBJ database and ExPasy analysis tool, respectively. In addition, IEDB server was used for the variation and conservation study of epitopes for all DENV serotypes. The complete methodology employed in this study is presented in Fig. 1.

Screening of HLA Binding Epitopes

The whole proteome DENV with accession Number AFZ40225 was examined for screening of HLA binding epitopes by using Propred (Singh and Raghava 2001) and Propred I (Singh and Raghava 2003) sequence based algorithms. HLA Class II binding nanomer peptides were screened against 51 alleles with 4% threshold value by Propred, similarly HLA Class I binding nanomer peptides were screened against 47 alleles with same threshold value by Propred I. These predicted epitope peptides were again validated by Artificial Neural Network (ANN) and (SMM) based algorithms through IEDB server.
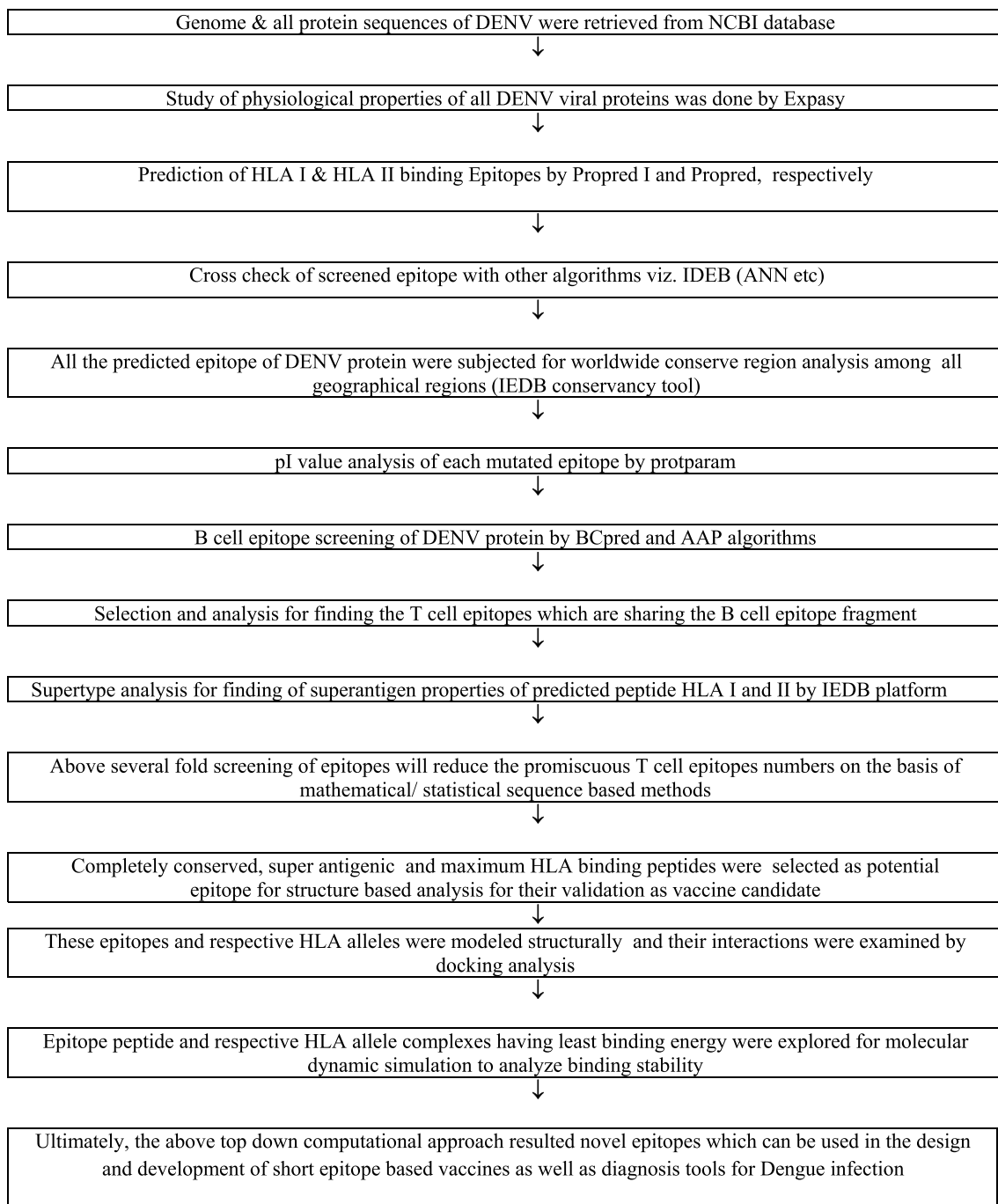
Genome & all protein sequences of DENV were retrieved from NCBI database

↓

Study of physiological properties of all DENV viral proteins was done by Expasy

↓

Prediction of HLA I & HLA II binding Epitopes by Propred I and Propred, respectively

↓

Cross check of screened epitope with other algorithms viz. IDEB (ANN etc)

↓

All the predicted epitope of DENV protein were subjected for worldwide conserve region analysis among all geographical regions (IEDB conservancy tool)

↓

pI value analysis of each mutated epitope by protparam

↓

B cell epitope screening of DENV protein by BCpred and AAP algorithms

↓

Selection and analysis for finding the T cell epitopes which are sharing the B cell epitope fragment

↓

Supertype analysis for finding of superantigen properties of predicted peptide HLA I and II by IEDB platform

↓

Above several fold screening of epitopes will reduce the promiscuous T cell epitopes numbers on the basis of mathematical/ statistical sequence based methods

↓

Completely conserved, super antigenic and maximum HLA binding peptides were selected as potential epitope for structure based analysis for their validation as vaccine candidate

↓

These epitopes and respective HLA alleles were modeled structurally and their interactions were examined by docking analysis

↓

Epitope peptide and respective HLA allele complexes having least binding energy were explored for molecular dynamic simulation to analyze binding stability

↓

Ultimately, the above top down computational approach resulted novel epitopes which can be used in the design and development of short epitope based vaccines as well as diagnosis tools for Dengue infection

Fig. 1 Flowsheet of top down computational approach for prediction of complete conserve, high population coverage and superantigenic epitopes to develop global DENV vaccine

Worldwide Conservancy Analysis of Screened T Cell Epitopes

All screened DENV T cell epitopes were undergone for worldwide conservancy study among different serotype. For conservancy study, NCBI Database was used to get the all protein sequences of all serotypes. Thus, maximum 4–5 sequences of each structural and non structural DENV protein were taken from database for all four DENV serotypes. The screened epitopes from each DENV strain protein were employed to Immune Epitope Database and Analysis Resource (IEDB) conservancy tool (Bui et al. 2007). This conservancy analysis was done against all four DENV serotype to identify complete conserve epitopes in all serotypes.

During this analysis, less than 70% conserve epitopes were screened out from further study and complete conserve, single and double mutated epitopes were forwarded for further analysis and also pI value calculation (Khan et al. 2006).

Cross Checking of T Cell Epitope Having Property of B Cell Epitope

The analysis first employed the BCPred method (EL-Manzalawy et al. 2008) to predict B cell epitopes from DENV proteome and further validation of results was done by an AAP method (Chen et al. 2007). This study useful to find those common T cell epitopes and B cell epitopes which share same peptide region and can have the ability to induce cell mediated as well as humoral immunity (Terry et al. 2015).

HLA Supertype Based Population Coverage Analysis for DENV

The HLA supertype concept gives strength in screening of high population coverage epitopes among a large pool of screened T cell epitopes. The potential conserve epitopes which shows binding with all HLA supertypes have high population coverage and become promising candidates for vaccine (Burrows et al. 2003). This analysis consists five–five HLA class I (A2, A3, A24, B7, B15) and HLA class II (DR, DR4, DR3, DP2, MainDP) supertypes. Class I HLA supertype have 23 alleles and HLA class II supertypes includes 14 alleles to maximize population coverage during analysis (Reche and Reinherz 2005).

This supertype analysis involved two known immunogenic peptides as positive controls to validate and compare with identified epitopes. These positive control peptides are [141]STLPETTVV[149] and [265]ILRGSVAHK[273] belongs to the Hepatitis core protein and H1N1 Nucleoprotein, respectively (Sharma et al. 2017; Singh et al. 2015a).

Modeling of Selected T-Cell Epitopes and Favored Alleles

Screened epitopes and favored alleles 3D models were created with the help of PEPstrMOD (Singh et al. 2015a, b) and Modeler 9.10 (Webb and Sali 2016) for structure based study, respectively.

Molecular Docking Analysis of Identified Epitopes with Respective HLA Alleles

Autodock 4.2 platform was used for the molecular binding study of epitopes and respective alleles (Morris et al. 2009) and also validated by Hex 8.0 (Ritchie et al. 2008). During docking study,.pdbqt file was generated by adding polar hydrogen atoms in place of water molecules from receptor and applying required charges (Kollman and Gasteiger charges etc.). Finally autodock tools were used to visualize and analyse the epitope and allele docked complexes. These binding analyses were further studied by Fast Fourier transformation based Hex 8.0 docking platform.

Molecular Dynamic (MD) Simulation Study of HLA Allele–Epitope Complexes

MD simulation of HLA allele and epitope complexes were done by using NAnoscale Molecular Dynamics (NAMD) (James et al. 2005) with Visual MD (VMD) (Humphrey et al. 1996). For MD simulations, epitope and allele complexes structural.PSF file was created through.PDB files with the help of a VMD PSF builder tool with consideration several force field parameters viz. equilibrium lengths, bond strengths and bonding interactions. NAMD produced trajectory.dcd file and source file (rmsd.tcl) also produced through VMD Tk console. The epitope allele complex root mean square deviation (RMSD) values were saved in form of the file rmsd.dat throughout the simulation process over the time window and RMSD plot was generated for the equilibrium stability of the complex during simulation.

Results

As per methodology all sequential steps results were shown in DENV vaccine development; to obtain best potential T cell epitopes as vaccine candidates against Dengue infection.

HLA Alleles Binding T Cell Epitopes of DENV

Top nanomer epitopes were predicted with 4% threshold value through a quantitative matrix of both Propred and Propred I algorithms, along with one more parameter also taken in consideration for selection of nanomer epitope such as epitope with high frequency to bind with HLA alleles (Sharma et al. 2017; Singh and Raghava 2003). Among top 29 identified epitopes, five structural protein epitopes viz. [33]LQGRGPLKL[41], [249]VVVLGSQEG[257], [172]LVGIVTLYL[180], [146]MKILIGVVI[154], [72]YIIVGVEPG[80] and one nonstructural protein epitope [18]LKNDIPMTG[26] were showed high conserve nature and population coverage from complete DENV proteome (Table 1). Additionally, with these above peptides, capsid protein represents QQLTKRFSL, LQGRGPLKL potential conserved epitopes.

Conservancy Analysis

The DENV conservancy study resulted 27 epitopes completely conserve among all predicted epitopes in all four serotypes. Among them, 15 HLA II binding epitopes and 11

Table 1 Top identified epitopes of DENV by Propred I and Propred with 4% threshold value

S. no.	Epitope position	Predicted T cell epitopes	PI value	HLA alleles, class	DENV serotypes
Capsid					
1.	22	QQLTKRFSL	11.0	43, I	**QQLTKRFSL***
2.	33	LQGRGPLKL	11.0	43, I	**LQGRGPLKL***
3.	45	LVAFLRFLT	9.75	45,II	**LVAFLRFLT***
Flavi glycoprot					
4.	249	VVVLGSQEG	4	26,II	**VVVLGSQEG***
Flavi E stem + E C					
5.	146	MKILIGVVI	8.50	20,II	**MKILIGVVI***
6.	172	LVGIVTLYL	5.52	43,II	**LVGIVTLYL***
7.	170	LVLVGIVTL	5.52	34,II	**LVLVGIVTL***
8.	97	MRGAKRMAI	12.01	39,II	**MRGAKRMAI***
9.	157	IGMNSRSTS	9.75	26,II	**IGMNSRSTS***
10.	72	YIIVGVEPG	4.0	20,II	**YIIVGVEPG***
11.	15	IVIRVQYEG	6.0	20,II	**IVIRVQYEG***
Flavi NS1					
12.	32	FQPESPSKL	6.0	36,I	FQPESPSK(AR^{D3})L
13.	102	LRPQPTELK	8.75	25,II	LR(T^{D3})PQPT(M^{D4})ELK
Flavi NS2A					
14.	108	TILELTDAL	3.67	20,I	**TILELTDAL***
15.	91	MMTTIGIVL	5.28	22,I	**MMTTIGIVL***
16.	104	TIPETILEL	3.79	21,I	**TIPETILEL***
17.	4	MALFLEEML	3.79	24,I	**MALFLEEML***
18.	97	IVLLSQSTI	5.52	26,II	**IVLLSQSTI***
Flavi NS2B					
19.	32	GLLTVCYVL	5.52	40,I	**GLLTVCYVL***
20.	41	TGRSADLEL	4.37	32,I	**TGRSADLEL***
21.	18	LKNDIPMTG	5.84	30,II	**LKNDIPMTG***
22.	36	VCYVLTGRS	8.19	26,II	**VCYVLTGRS***
23.	34	LTVCYVLTG	5.52	37,II	**LTVCYVLTG***
Flavi NS4A					
24.	58	ATVTGGIFL	5.57	22,I	**ATVTGGIFL***
25.	60	VTGGIFLFL	5.49	21,I	**VTGGIFLFL***
26.	51	LLLLTLLAT	5.52	32,II	**LLLLTLLAT***
27.	24	LAVLHTAEA	5.24	23,II	**LAVLHTAEA***
Flavi NS5					
28.	114	EPKEGTKKL	8.59	21,I	**EPKEGTKKL***
29.	201	YNMMGKREK	9.70	25,II	**YNMMGKREK***

Interpretation for Table 1

*Depict conserved epitope in among all serotype strains (DENV1, DENV2, DENV3 and DENV4)

In DENV Serotype column, green colour amino acid residues represent varied amino acid in different serotype strain

Interpretation example for DENV Serotype column: LR(T^{D3})PQPT(M^{D4})ELK, here original peptide is LRPQPTELK. Arginine (R) second amino acid is replaced by Threonine (T) in DENV3 serotype and Threonine (T) at sixth position is replaced by Methionine in DENV4 with change in pI

HLA I binding epitope were observed complete conserve. The screened HLA II binding peptides showed 76% conservancy with DENV2, 66% with DENV3, 61% with DENV1 and 57% with DENV4 (Fig. 2). DENV2 was observed highly conserved (76%) among all serotype. Similarly, screened HLA I binding peptides showed 78% conservancy with DENV2, 67% with DENV1, 63% with DENV3, and 61% with DENV4 (Fig. 2). Structural protein epitopes QQLTKRFSL, LQGRGPLKL, VVVLGSQEG, LVGIVTLYL, MKILIGVVI, YIIVGVEPG and nonstructural protein
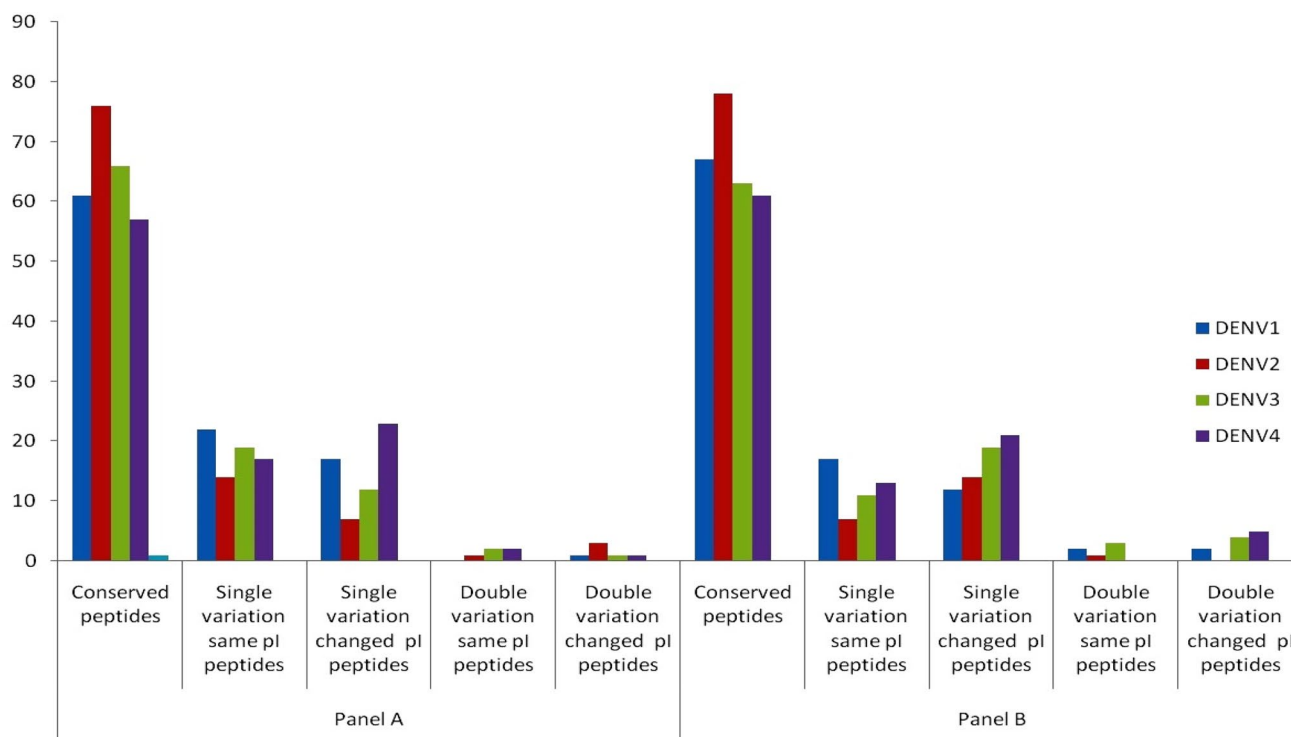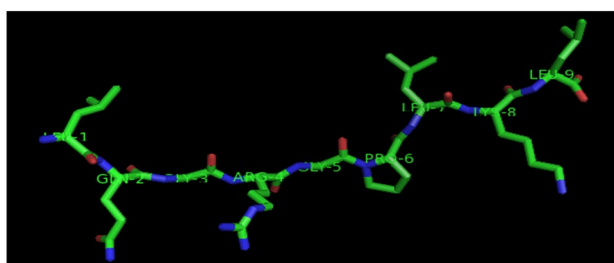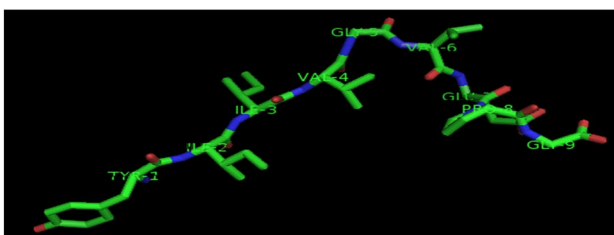
Fig. 2 A worldwide relative epitopes conservancy analysis of DENV serotypes for HLA alleles. Panel A representing analysis of HLA II binding peptides and Panel B representing analysis of HLA I binding peptides

a LQGRGPLKL

b YIIVGVEPG

Fig. 3 DENV epitope peptides model by PEPstrMOD viewed by PyMOL viewer. **a** YIIVGVEPG, **b** LQGRGPLKL (Singh et al. 2015a, b)

LKNDIPMTG, IVLLSQSTI, LLLLTLLAT, YNMMG-KREK peptide epitopes were observed top binder in case of conservancy and allele coverage (Fig. 3).

B Cell Epitopes with Property of T Cell Epitopes

DENV three T cell epitope have similarity with predicted B cell epitopes such as VVVLGSQEG, YIIVGVEPG and IVIR-VQYEG; but FQPESPSKL and LRPQPTELK T cell epitope show partially match with B cell epitope fragments.

B Cell Epitope Fragments Which Showed T Cell Epitopes Within as Highlighted Colored Text

DENV:

Seq 1: KNPHAKKQD**VVVLGSQEG**AM protein **Glyco-prot c envelope**

Seq 2: IEAEPPFGDS**YIIVGVEPG**Q protein **Flavi Eenvelope**

Seq 3: HGT**IVIRVQYEG**DGSPCKIP protein **Flavi Eenvelope**

Seq 4: **RPQPTELK**YSWKTWGKAKML protein **NS1 non structural**

Seq 5: TDNVHTWTEQYK**FQPESPSK** protein **NS1 non structural**

HLA Supertype Based Population Coverage Analysis for DENV

This findings revealed that epitopes QQLTKRFSL, LQGRG-PLKL, GLLTVCYVL and TGRSADLEL showed binding to all HLA alleles of supertype HLA class I A2, A3, A24 and B7 additionally these identified epitopes also show binding to several members of the other HLA class I supertypes as well (Table 2). Likewise, the epitopes VVVLGSQEG, MKILIGVVI, LVGIVTLYL, YIIVGVEPG, LKNDIPMTG and IVLLSQSTI were showed binding to almost all alleles DR, DR3, DP2, DR4 and Main DP HLA class II supertype; additionally, epitopes also presented binding to several members of other HLA class II supertypes (Table 3). These all identified HLA Class I and HLA class II binding epitopes showed acceptable result in comparison of both positive control peptides (STLPETTVV and ILRGSVAHK). Sharma et al. (2018) conducted similar supertype analysis for West Nile vaccine development which gave most promising vaccine candidates with similar sets of supertype alleles.

Therefore QQLTKRFSL, LQGRGPLKL VVVLGSQEG, MKILIGVVI, LVGIVTLYL, YIIVGVEPG, LKNDIPMTG and IVLLSQSTI epitopes were identified as superantigen peptides. These tabulated superantigen peptides and respective HLA alleles, were modeled for binding simulations study. In population coverage analysis by IEDB platform, these all conserved epitopes taken against most frequent supertype (IEDB recommended) set of alleles HLA class I and class II. This study reveals that mapped class I epitopes shows overall 79.12% of population coverage and mapped class II epitopes show overall 81.81% of population coverage (Table 4). The epitopes QQLTKRFSL and LQGRGPLKL showed 43.68% population coverage with HLA Class I set of alleles (HLA-A*02:01, HLA-A*02:02, HLA-A*02:06, HLA-A*02:03 and HLA-A*68:02) and MKILIGVVI, LVGIVTLYL, VVVLGSQEG and YIIVGVEPG epitopes showed 81.81%, 63.18%, 59.93% and 70.55% coverage with HLA class II set of alleles (HLA-DRB1*01:01, HLA-DRB1*15:01, HLA-DRB1*03:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*11:01, HLA-DRB1*12:01, HLA-DRB1*13:02, HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB4*01:01) respectively.

On the basis of above all, above sequence based results the most promiscuous T cell peptide epitopes viz. *DENV:*QQLTKRFSL, *DENV:*LQGRGPLKL *DENV:*VVVLGSQEG, *DENV:*MKILIGVVI, *DENV:*LVGIVTLYL, *DENV:*YIIVGVEPG, *DENV:*LKNDIPMTG and *DENV:*IVLLSQSTI were showed good population coverage, propred allele frequency, conservancy and super antigen property; which will undergo for structure based binding simulation analysis with most frequent and IEDB recommended set of HLA alleles viz. HLA I (A* 0101, A* 0201, A* 0301, B*0702, B*3501, B*5101,

B*5102 and B*5301) and HLA II (DRB1*0101, DRB1*0401, DRB1*0405 and DRB1*0301) as shown in Tables 2 and 3.

T-Cell Epitopes 3D Structure Modeling

The PEPstrMOD (Peptide 3D Prediction server) modeled structure of epitopes (LQGRGPLKL, QQLTKRFSL, LKNDIPMTG, VVVLGSQEG, YIIVGVEPG, IVLLSQSTI and LVGIVTLYL) with sequence length 9 residues.

Superantigenic DENV Epitopes and Respective HLA Alleles Binding Simulation Study

Structure based analysis by docking and Molecular Dynamic (MD) simulation is required to confirm peptide and favoured allele interaction and their complex stability to induce immune response (Peele et al. 2020). Here binding study of superantigen epitopes VVVLGSQEG, LVGIVTLYL, MKILIGVVI, YIIVGVEPG, LKNDIPMTG, IVLLSQSTI, QQLTKRFSL and LQGRGPLKL with HLA alleles, were done to analyse the docking pattern between them. Docking of epitope [72]YIIVGVEPG[80]–DRB1 0401 allele and epitope [33]LQGRG-PLKL[41]–B*5101 allele complexes showed the least binding energy of − 7.71 and − 7.20 kcal/mol, respectively.

Complex YIIVGVEPG–DRB1 0401 showed two Hydrogen bond viz. ARG146: NH and THR90: O (Fig. 4). Similarly complex of LQGRGPLKL–B*5101 showed three H bond viz. ASN77: NH, TYR99: OH and TYR9: OH (Fig. 5). Confirmation using Hex 8.0 also presented for these complexes to show their stability. Autodock 4.2 and Hex 8.0 result showed in Table 5.

Superantigen Epitope–HLA Allele Complexes Binding Simulation Study by NAMD

The superantigen epitope and HLA allele complexes with best binding energy viz. YIIVGVEPG–DRB1 0401 and LQGRGPLKL–B*5101 resulted by docking Autodock 4.2 were examined for stable binding pattern over the time window during molecular dynamic simulation (NAMD-VMD).

DENV epitope YIIVGVEPG–DRB1 0401 allele complex displayed the maximum RMSD value of 10.1 Å (Fig. 6) likewise LQGRGPLKL–B*5101 allele complex showed the maximum RMSD value of 8.4 Å (Fig. 7). Both complexes showed stable and equilibrium binding pattern as depicted in RMSD values.

Discussion

In present analysis, two superantigen epitopes viz. LQGRG-PLKL and YIIVGVEPG were found top DENV superantigen among all top screened superantigen epitopes with highest population coverage and conservancy. The study

Table 2 DENVepitopes having high affinity (less than 50 percentile value) with HLA class I Supertypes by IEDB ANN and SMM methods

DENV HLA Class I Supertype analysis

S. no.	Epitope	Supertype A2 (percentile value)				
		A*0201	A*0202	A*0203	A*0205	A*0206
1	QQLTKRFSL	4.5	4.2	7.6	1.6	1.6
2	LQGRGPLKL	6.7	4.9	7.8	3.1	3.8
4	GLLTVCYVL	0.99	1.7	4.0	3.1	2.4
5	TGRSADLEL	27	22	24	16	20
6	*HBV*- STLPETTVV	1.2	1.8	1.1	0.37	0.2
7	*H1N1*- ILRGSVAHK	14	12	6.8	12	16

DENV HLA Class I Supertype analysis

S. no.	Epitope	A3 Supertype				
		A*0301	A*1101	A*3101	A*3301	A*6801
1	QQLTKRFSL	18	16	6.5	7.5	24
2	LQGRGPLKL	15	27	26	47	48
4	GLLTVCYVL	36	36	17	29	52
5	TGRSADLEL	26	28	25	30	30
6	*HBV*- STLPETTVV	7.4	6.4	9.5	8.9	7.9
7	*H1N1*- ILRGSVAHK	0.02	0.62	0.44	2.5	2.3

DENV HLA Class I Supertype analysis

S. no.	Epitope	A24 Supertype				
		A*2301	A*2402	A*2403	A*2405	A*2407
1	QQLTKRFSL	4.0	4.6	4.2	4.6	4.5
2	LQGRGPLKL	4.3	5.7	5.4	5.7	4.4
4	GLLTVCYVL	8.7	11	9.4	11	8.7
5	TGRSADLEL	14	15	13	15	9.1
6	*HBV*- STLPETTVV	5.9	6.0	6.6	6.0	5.2
7	*H1N1*- ILRGSVAHK	22	21	22	21	18

DENV HLA Class I Supertype analysis

S. no.	Epitope	B7 Supertype				
		B*0702	B*3501	B*5101	B*5301	B*5102
1	QQLTKRFSL	3.0	7.6	13	9.0	14
2	LQGRGPLKL	14	12	14	8.5	13
4	GLLTVCYVL	29	36	24	31	30
5	TGRSADLEL	1.5	4.9	6.4	5.5	4.1
6	*HBV*- STLPETTVV	3.4	4.2	0.84	3.3	0.57
7	*H1N1*- ILRGSVAHK	14	30	51	43	54

DENV HLA Class I Supertype analysis

S. no.	Epitope	B15 Supertype				
		A*0101	B*1501	B1502	–	–
1	QQLTKRFSL	24	1.6	2.6		
2	LQGRGPLKL	21	1.4	3.8		
4	GLLTVCYVL	37	14	21		
5	TGRSADLEL	20	7.9	4.5		
6	*HBV*- STLPETTVV	3.8	3.4	2.3		
7	*H1N1*- ILRGSVAHK	20	5.7	9.6		

Table 2 (continued)

Less 50 value represents high affinity with allele

Table 3 DENV epitopes having high affinity (less than 50 percentile range) with HLA class II Supertypes by IEDB ANN and SMM methods

DENV HLA Class II Supertype analysis

S. no.	Epitope	DR Supertype (percentile range)				
		DRB1*0101	1*0701	1*0901	1*1101	1*1501
1	*DENV-* VVVLGSQEG	32–51	72–79	32–63	51–69	30–51
2	*DENV-* MKILIGVVI	15	5.90–35	21–51	20–31	4.10–8.80
3	*DENV-* LVGIVTLYL	6–20	20–42	56–61	16–29	1.50–1.80
4	*DENV-* YIIVGVEPG	7–54	22–64	16–61	33–73	30–73
5	*DENV –* LKNDIPMTG	19–66	61–88	37–60	44–65	31–51
6	*DENV-* IVLLSQSTI	15–23	21–27	35–51	18–39	9.20–33
7	*HBV-* STLPETTVV	15–90	34–76	20–98	53–67	50–92
8	*H1N1-* ILRGSVAHK	4.80–27	11–36	24–48	18–32	6.80–44

DENV HLA Class II Supertype analysis

S. no.	Epitope	DR4 Supertype			DR3 Supertype	
		DRB1*0401	1*0405	1*0802	DRB1*0301	3*0101
1	*DENV-* VVVLGSQEG	20–35	40–56	37–58	31–64	73–82
2	*DENV-* MKILIGVVI	33–50	18–37	6–21	24–56	41–49
3	*DENV-* LVGIVTLYL	7.40–24	1.30–6.60	18	9.30–50	34–64
4	*DENV-* YIIVGVEPG	16–44	19–47	5.70–57	19–84	52–63
5	*DENV –* LKNDIPMTG	7.80–31	27–58	28–46	1.40–29	24–47
6	*DENV-* IVLLSQSTI	8.20–28	4.40–18	7.10–22	18–49	59–76
7	*HBV-* STLPETTVV	23–76	33–83	25–83	19–59	71–82
8	*H1N1-* ILRGSVAHK	6.90–28	32–72	6.40–21	29–56	65–76

DENV HLA Class II Supertype analysis

S. no.	Epitope	DP2	Main DP Supertype		
		DPB1*0201	DPB1*0101	0402	0501
1	*DENV-* VVVLGSQEG	51–77	50–62	60–71	54–69
2	*DENV-* MKILIGVVI	19–36	35–68	25–51	39–70
3	*DENV-* LVGIVTLYL	13–17	39–67	23–45	53–76
4	*DENV-* YIIVGVEPG	16–45	18–48	23–49	26–70
5	*DENV –* LKNDIPMTG	41–83	28–73	27–73	26–74
6	*DENV-* IVLLSQSTI	15–28	24–35	27–36	41–45
7	*HBV-* STLPETTVV	38–72	17–77	26–79	35–65
8	*H1N1-* ILRGSVAHK	59–72	19–53	24–56	4.70–25

Less percentile value represents high affinity with allele

Table 4 DENV HLA class I (A) and II (B) binding epitopes population coverage calculation result

Population/area	Class I			Class II		
	Coverage[a]	Average hit[b]	PC90[c]	Coverage	Average hit[b]	PC90[c]
World	79.12%	5.75	0.48	81.81%	9.63	0.55

[a]Projected population coverage

[b]Average number of epitope hits/HLA combinations recognized by the population

[c]Minimum number of epitope hits/HLA combinations recognized by 90% of the population

Fig. 4 Docked peptide YIIVG-VEPG–DRB1 0401 allele complex showed two hydrogen bond interactions (green dot string) with THR90 and ARG146 amino acids by Autodock 4.2
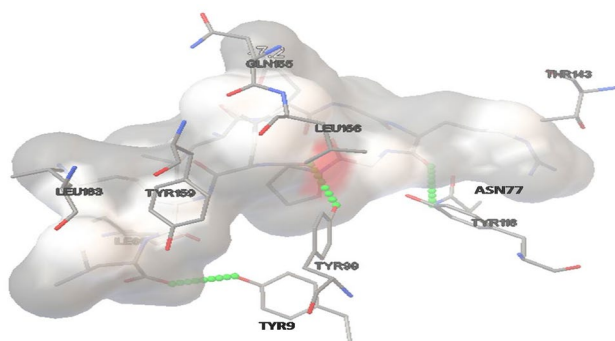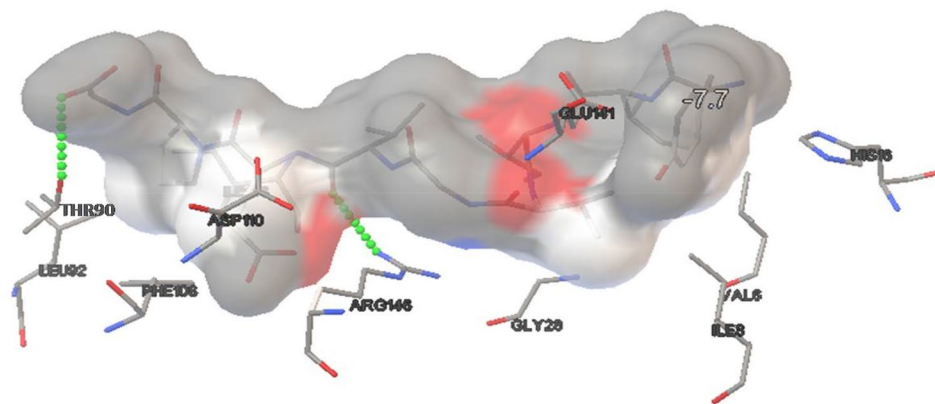

Fig. 5 Docked peptide LQGRGPLKL–B*5101 allele complex showed three hydrogen bond interactions (green dot string) with ASN 77, TYR99 and TYR9 amino acids by Autodock 4.2
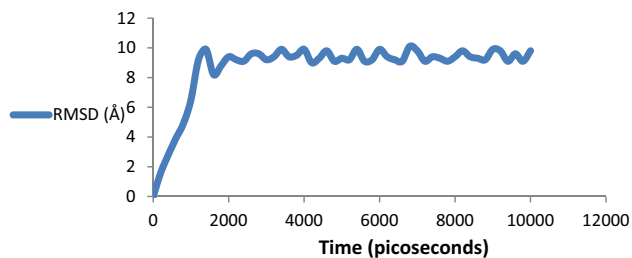

Fig. 6 NAMD-VMD simulation graph between RMSD value and time window (picoseconds) of YIIVGVEPG–DRB1 0401 complex, displaying equilibrium and stability with highest RMSD value (10.1 Å) at 6800 picoseconds

showed that LQGRGPLKL epitope binding to supertypes HLA class I all allele members viz. **A2** (A*0201, A*0202, A*0203, A*0205, A*0206), **A3**(A*0301, A*1101, A*3101, A*3301, A*6801), **A24** (A*2301, A*2402, A*2403, A*2405, A*2407), **B7** (B*5101, B*5102, B*5301, B*0702, B*3501) and **B15** (A*0101, B*1501, B*1502) and YIIVG-VEPG peptide epitope binding to HLA class II supertype, all allele members DR(DRB1*0101, 1*0701, 1*0901, 1*1101, 1*1501), **DR4** (DRB1*0401, 1*0405, 1*0802), **DR3** (DRB1*0301, 1*0101), **DP2** (DPB1*0201), **Main DP** (DPB1*0101, 1*0402, 1*0501) with recommended IEDB value (Tables 2 and 3). Both epitopes supertype analysis were compared with positive control peptides which gave acceptable values shown in Tables 2 and 3. As a result, epitopes LQGRGPLKL and YIIVGVEPG were identified as most potential superantigenic epitopes for DENV. Similar sequence based studies resulted superantigenic peptide epitope vaccine candidates from whole proteome to develop multi epitope vaccines against various pathogens, recently such as Japanese encephalitis and West Nile virus vaccine (Sharma et al. 2017, 2018). Screening potential epitopes from complete proteome of pathogen including structural and nonstructural proteins in top down manner result least number of potential epitopes which were tested with several parameters such as high frequency to bind HLA alleles, high population coverage and conservancy. These factors are essential to present good vaccine candidates and further these epitopes structure based study by docking and simulation will prove their interaction with alleles. The
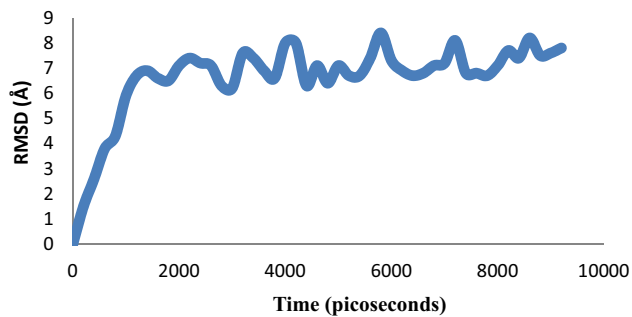
Table 5 Top DENV peptide epitopes with respective supertype alleles docking result by Hex 8.0 and Autodock 4.2

S. no	DENVepitope	Allele	Autodock 4.2		Hex 8.0	
			Binding energy	H bond	Hex energy	H bond
1.	VVVLGSQEG	DRB1_0405	− 5.60	2	− 501.70	1
2.	LVGIVTLYL	DRB1_0401	− 6.50	2	− 323.32	1
3.	YIIVGVEPG	DRB1_0401	− 7.71	2	− 421.72	1
4.	LQGRGPLKL	B*5101	− 7.20	3	− 248.27	1
5.	QQLTKRFSL	B*5101	− 7.19	4	− 221.67	2

Fig. 7 NAMD-VMD simulation graph between RMSD value and time window (picoseconds) of LQGRGPLKL–B*5101 complex, displaying equilibrium and stability with highest RMSD value (8.4 Å) at 5800 picoseconds

strong epitope and allele complex interaction with helper T cell result cell mediated and humoral immune responses (Chaplin 2010). These computational approaches gave the best platform for vaccine development by reducing random experimental time and cost (María et al. 2017). As the present sequence base study showed two epitopes LQGRG-PLKL and YIIVGVEPG as super antigen peptides which further under gone for structure based analysis by docking and Molecular Dynamic (MD) simulation to confirm peptide and favored allele interaction and their complex stability to induce immune response (Peele et al. 2020).

The present docking of DENV peptide YIIVGVEPG with allele DRB1 0401 showed good binding by both Hex 8.0 and Autodock 4.2 docking platforms. Binding energy of YIIVGVEPG–DRB1 0401 complex via Autodock 4.2 is − 7.71 kcal/mol and showed two hydrogen bonds viz. ARG146: NH and THR90: O amino acid residues. Molecular dynamic simulation NAMD-VMD analysis confirm that the complex YIIVGVEPG–DRB1 0401 showed stable binding pattern over the time window of 6800 picoseconds. Likewise LQGRGPLKL–B*5101 peptide allele complex also showed stable binding through Hex 8.0 and Autodock 4.2 tools. The complex showed good binding energy (− 7.20 kcal/mol) with three H-bonds viz. ASN77: NH, TYR99: OH and TYR9: OH.

Here, NAMD-VMD study further confirms that the LQGRGPLKL–B*5101 complex showed stable binding pattern over the time window of 5800 picoseconds. Finally the present analysis extracted two peptides LQGRGPLKL and YIIVGVEPG as the most potential superantigenic vaccine candidate for DENV. Along with this several other promising peptides can be utilize as diagnostic reagents as well as DENV vaccine candidates. Since long back computation top down vaccine development approaches resulted potential peptide candidates against many infectious diseases viz. H1N1, HIV Tuberculosis and JEV (Jardine et al. 2013; De Groot et al. 2013; Feng et al. 2013; Sharma and

Kumar 2010). Recently, also the similar peptide based vaccine development approaches already given good results for several other more infectious agents such as Chikungunya, Campylobacter vaccine development (Saxena and Mishra 2020; Gupta and Kumar 2020a, b). This approach has proven much advantage over conventional vaccine development procedure such as time, cost, side effects and no risk of the presence of infectious whole pathogens (Khalili et al. 2014; María et al. 2017). However, these peptide epitopes will be further examined as diagnostic reagents and potent vaccine candidates for DENV.

Conclusion

The present top down approach screened out two most promiscuous peptide epitopes LQGRGPLKL and YIIVGVEPG. These epitopes are superantigenic, novel conserved vaccine candidates and can also be good diagnostic reagents for DENV. These peptides can replace the use of whole viral proteins or whole attenuated viruses as vaccine candidates. These potential peptide epitopes could be further examined for vaccine and diagnostic reagents property.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Bui HH, Sidney J, Li W, Fusseder N, Sette A (2007) Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinform 8(1):361

Burrows SR, Elkington RA, Miles JJ, Green KJ, Walker S, Haryana SM, Moss DJ, Dunckley H, Burrows JM, Khanna R (2003) Promiscuous CTL recognition of viral epitopes on multiple human leukocyte antigens: biological validation of the proposed HLA A24 supertype. J Immunol 171:1407–1412

Chaplin DD (2010) Overview of the immune response. Allergy Clin Immunol 125(2 Suppl 2):S3–S23. https://doi.org/10.1016/j.jaci.2009.12.980

Chen J, Liu H, Yang J, Chou K (2007) Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids 33:423–428

De Groot AS, Ardito M, Terry F, Levitz L, Ross T, Moise L (2013) Low immunogenicity predicted for emerging avian-origin H7N9. Hum Vaccine Immunother 9(5):950–956

EL-Manzalawy Y, Dobbs D, Honavar V (2008) Predicting linear B-cell epitopes using string kernels. J MolRecognit 21:243–255

Feng G, Jiang Q, Xia M, Lu Y, Qiu W, Zhao D, Lu L, Peng G, Wang Y (2013) Enhanced immune response and protective effects of nano-chitosan-based DNA vaccine encoding T cell epitopes of Esat-6 and FL against mycobacterium tuberculosis infection. PLoS ONE 8(4):e61135

Gubler DJ (2004) The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? Comp Immunol Microbiol Infect Dis 27:319–330

Gupta N, Kumar A (2020a) Identification of potent vaccine candidates against *Campylobacter jejuni* using immunoinformatics approach. Int J Pept Res Ther 26:1303–1312

Gupta N, Kumar A (2020b) Designing an efficient multi-epitope vaccine against *Campylobacter jejuni* using immunoinformatics and reverse vaccinology approach. Microbial Pathog 147:104398

Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, Hunsperger E, Kroeger A, Margolis HS, Martínez E, Nathan MB, Pelegrino JL, Simmons C, Yoksan S, Peeling RW (2010) Dengue: a continuing global threat. Nat Rev Microbiol 8(12):S7–S16

Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. J Mol Gr 14:33–38

James CP, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802

Jardine et al (2013) Rational HIV immunogen design to target specific germline B cell receptors. Science 340:711–716

Kangueane P, Sakharkar MK, Rajaseger G, Bolisetty S, Sivasekari B, Zhao B, Ravichandran M, Shapshak P, Subbiah S (2005) A framework to sub-type HLA supertypes. Front Biosci 1(10):879–886

Khalili S, Jahangiri A, Borna H, Ahmadi Zanoos K, Amani J (2014) Computational vaccinology and epitope vaccine design by immunoinformatics. Acta Microbiol Immunol Hung 61(3):285–307

Khan AM, Miotto O, Heiny AT, Salmon J, Srinivasan KN, Nascimento EJ, Marques ET Jr, Brusic V, Tan TW, August JT (2006) A systematic bioinformatics approach for selection of epitope-based vaccine targets. Cell Immunol 244(2):141–147

Kirsten SV, John TR, Joachim H (2015) Next generation dengue vaccines: a review of the preclinical development pipeline. Vaccine 33:7091–7099

María RR, Arturo CJ, Alicia JA, Paulina MG, Gerardo AO (2017) The impact of bioinformatics on vaccine design and development. Vaccines. https://doi.org/10.5772/intechopen.69273

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 16:2785–2791

Olsen LR, Zhang GL, Reinherz EL, Brusic V (2011) FLAVIdB: a data mining system for knowledge discovery in flaviviruses with direct applications in immunology and vaccinology. Immunome Res 7(3):2

Peele KA, Durthi CP, Srihansa T, Krupanidhi S, Ayyagari VS, Babu DJ, Indira M, Reddy AR, Venkateswarulu TC (2020) Molecular docking and dynamic simulations for antiviral compounds against SARS-CoV-2: a computational study. Inform Med Unlocked 19:100345

Pinheiro-Michelsen JR, Souza RDSO, Santana IVR, da Silva PdS, Mendez EC, Luiz WB, Amorim JH (2020) Anti-dengue vaccines: from development to clinical trials. Front Immunol 11:1252. https://doi.org/10.3389/fimmu.2020.01252

Pulmanausahakul R, Khakpoor A, Smith DR (2010) The development of flavivirus vaccines. Afr J Biotechnol 9(4):409–415

Reche PA, Reinherz EL (2005) PEPVAC: a web server for multi-epitope vaccine development based on the prediction of supertypic MHC ligands. Nucleic Acids Res 33:138–142

Ritchie DW, Kozakov D, Vajda S (2008) Accelerating protein–protein docking correlations using a six-dimensional analytic FFT generating function. Bioinformatics 24:1865–1873

Sanyaolu A, Okorie C, Badaru O, Adetona K, Ahmed M et al (2017) Global epidemiology of dengue hemorrhagic fever: an update. J Human Virol Retrovirol 5(6):00179. https://doi.org/10.15406/jhvrv.2017.05.00179

Saxena P, Mishra S (2020) Study of the binding pattern of HLA class I alleles of Indian frequency and cTAP binding peptide for Chikungunya vaccine development. Int J Pept Res Ther 26:2437–2448

Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S et al (2009) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 37:D5–D15

Schiøler KL, Samuel M, Wai KL (2007) Vaccines for preventing Japanese encephalitis. Cochrane Database Syst Rev 18:3

Sharma P, Kumar A (2010) Immunoinformatics: screening of potential T cell antigenic determinants in proteome of H1N1 swine influenza virus for virus epitope vaccine design. J Proteom Bioinform 3:275–278

Sharma P, Saxena K, Mishra S, Kumar A (2014) A comprehensive analysis of predicted HLA binding peptides of JE viral proteins specific to north Indian isolates. Bioinformation 10(6):334–341

Sharma P, Srivastav S, Mishra S, Kumar A (2017) Sequence and structure based binding prediction study of HLA class I and cTAP binding peptides for Japanese encephalitis vaccine development. Int J Pept Res Ther 23:269–279. https://doi.org/10.1007/s10989-016-9558-0

Sharma P, Mishra S, Kumar A (2018) Analysis of promiscuous T cell epitopes for vaccine development against west nile virus using bioinformatics approaches. Int J Pept Res Ther 24:377–387. https://doi.org/10.1007/s10989-017-9624

Singh H, Raghava GPS (2001) Propred: prediction Of HLA-DR binding sites. Bioinformatics 17(12):1236–1237

Singh H, Raghava GP (2003) ProPred1: prediction of promiscuous MHC class-I binding sites. Bioinformatics 19(8):1009–1014

Singh A, Mitra M, Sampath G, Venugopal P, Rao JV, Krishnamurthy B, Gupta MK, Sri Krishna S, Sudhakar B, Rao NB, Kaushik Y, Gopinathan K, Hegde NR, Gore MM, Krishna Mohan V, Ella KM (2015a) A Japanese encephalitis vaccine from India induces durable and cross-protective immunity against temporally and spatially wide-ranging global field strains. J Infect Dis 212(5):715–725

Singh S, Singh H, Tuknait A, Chaudhary K, Singh B, Kumaran S, Raghava GPS (2015b) PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. Biol Direct 10:73

Srivastava S, Sharma SK, Srivastava V, Kumar A (2020) Proteomic exploration of *Listeria monocytogenes* for the purpose of vaccine designing using a reverse vaccinology approach. Int J Pept Res Ther. https://doi.org/10.1007/s10989-020-10128-1

Terry FE, Moise L, Martin RF, Torres M, Pilotte N, Williams SA, Groot ASD (2015) Time for T? Immunoinformatics addresses vaccine design for neglected tropical and emerging infectious diseases. Expert Rev Vaccines 14(1):21–35

Vannice KS, Durbin A, Hombach J (2016) Status of vaccine research and development of vaccines for dengue. Vaccine 34:2934–2938

Webb B, Sali A (2016) Comparative protein structure modeling using MODELLER. Curr Protoc Bioinform 54:5.6.1–5.6.37

WHO (2020) https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue

Williams KL, Zompi S, Beatty PR, Harris E (2009) A mouse model for studying dengue virus pathogenesis and immune response. Ann N Y Acad Sci 1171(Suppl 1):E12–E23