



Computational prediction of immunodominant antigenic regions & potential protective epitopes for dengue vaccination

Karthikeyan Muthusamy, Krishnasamy Gopinath & Dharmalingam Nandhini

Department of Bioinformatics, Science Campus, Alagappa University, Karaikudi, India

Received March 2, 2014

Background & objectives: Epitope-based vaccines (EVs) are specific, safe and easy to produce. However, vaccine failure has been frequently reported due to variation within epitopic regions. Therefore, development of vaccines based on conserved epitopes may prevent such vaccine failure. This study was undertaken to identify highly conserved antigenic regions in the four dengue serotypes to produce an epitope-based dengue vaccine.

Methods: Polyprotein sequences of all four dengue serotypes were collected and aligned using MAFFT multiple sequence alignment plugin with Geneious Pro v6.1. Consensus sequences of the polyproteins for all four dengue serotypes were designed and screened against experimentally proven epitopes to predict potential antigenic regions that are conserved among all four dengue serotypes.

Results: The antigenic region VDRGWGNGCGLFGKG was 100 per cent conserved in the consensus polyprotein sequences of all four dengue serotypes. Fifteen experimentally proven epitopes were identical to the immunodominant antigenic region.

Interpretation & conclusions: Computationally predicted antigenic regions may be considered for use in the development of EVs for protection against dengue virus. Such vaccines would be expected to provide protection against dengue infections caused by all dengue serotypes because these would contain antigenic regions highly conserved across those serotypes. Therefore, the immunodominant antigenic region (VDRGWGNGCGLFGKG) and 15 potential epitopes may be considered for use in dengue vaccines.

Key words Antigenic region - dengue infection - epitope-based vaccine - sequence analysis

Dengue viruses (DENVs) are the aetiologic agent for dengue fever and can sometimes cause fatal haemorrhagic fever/shock syndrome. It is estimated that 50-100 million dengue infections occur annually in over 100 endemic countries¹. In India, 201771 infected cases and 920 deaths were reported from 2007 to 2013 and in 2015 alone, 99913 infected cases were reported².

Vaccination may be the most viable option to prevent dengue infection, but unfortunately, it has been largely unsuccessful in preventing outbreaks. The ideal dengue vaccine must induce long-term immunity against all the four dengue serotypes. Even after more than 60 yr of research, a licensed vaccine against dengue is still elusive. At present, only 'candidate' dengue vaccines

are available. Candidate dengue vaccine development has been based on a variety of technologies including live attenuated virus, recombinant proteins, chimeric vaccines, recombinant viral vectors and DNA vaccines³. Epitopes are the regions of an antigen that are bound by antigen-specific adaptive immune membrane receptors on lymphocytes or by secreted antibodies. Epitopes provide valuable information for disease prevention, diagnosis, treatment⁴, and also represent a new strategy for prophylactic and therapeutic induction of pathogen-specific immunity⁵. Epitope-based vaccines (EVs) can induce both cellular (T-cell epitopes) and humoral (B-cell epitopes) immune responses. T-cell epitopes bind to major histocompatibility complex (MHC) and interact with the T-cell receptor, stimulating a T-cell response. B-cell epitopes interact with antibodies. Recognition of the appropriate peptide-MHC complexes by the antigen-specific receptor of T-lymphocytes leads to cell proliferation and a cascade of cellular immune responses⁶. MHC-II molecules bind longer peptides (15-20 amino acids), whereas MHC-I molecules bind shorter peptides (9 amino acids or less)⁷. Prediction of such epitope binding is critical. Hence, various computational algorithms and methods have been used for the prediction of epitopes⁸ and mostly conserved epitopes should be considered for vaccine development⁹.

The identification of immunodominant antigenic regions has significantly advanced the development of peptide-based vaccines. Potential advantages of the epitope-based approach include increased safety and the opportunity to rationally engineer epitopes for increased potency and breadth¹⁰. Antigenic regions induce a more specific immune response to a disease compared to other vaccination methods¹¹. Computational approaches in predicting epitopes are an alternative to the expensive experimental discovery. However, computational methods may not predict functional protective epitopes due to variation among epitope regions across strains or serotypes. Advantages of computational prediction of highly conserved epitopes for use in vaccines are the time saved and the safety of such vaccines¹⁰.

This study was aimed to identify highly conserved antigenic regions in the four dengue serotypes (DENV1, DENV2, DENV3, DENV4) with the goal to develop an epitope-based dengue vaccine. Experimentally proven epitopes were used to identify potential immunodominant antigenic regions. These antigenic regions were identified by screening experimentally

proven epitopes against consensus sequences from all four dengue polyproteins.

Material & Methods

Computational methods were carried out in the Pharmacogenomics and Computer Aided Drug Design (CADD) Laboratory, Department of Bioinformatics, Alagappa University, Karaikudi, Tamil Nadu, India, during March-November 2013.

Sequence collection: To identify the conserved antigenic regions, all available DENV whole genomes for the four dengue serotypes were retrieved from the NCBI Nucleotide database¹². Genomic regions, encoding for the polyproteins, were identified from GenBank (www.ncbi.nlm.nih.gov/genbank/). The respective protein sequences were retrieved from the NCBI protein database (www.ncbi.nlm.nih.gov/protein) and saved in FASTA format (www.ncbi.nlm.nih.gov/blast/fasta.shtml) These four sets of protein sequences were used for the generation of consensus sequences from multiple sequences.

Multiple sequence alignment: Geneious Pro v6.1 (Biomatters Ltd, New Zealand)¹³ was used for multiple sequence alignment (MSA) and consensus sequence generation from all four sets of polyprotein sequences. MSA was generated for polyproteins of all four serotypes using MAFFT Plugin¹⁴. FFT-NS1 algorithm¹⁵, BLOSUM 62 was selected and Gap open penalty - 1.53 and Offset value 0.123 was set to generate MSA. MSA of four dengue serotype with 100 per cent identical sites (amino acid residue matching in all the sequences) was used for the generation of consensus sequence.

Collection of experimentally proved positive epitopes: The positive epitopes were collected from Immune Epitopes Database (IEDB)¹⁶ which contains both positive and negative experimental results, to predict potential immunodominant antigenic regions in all serotypes. Experimentally proven positive epitope sequences were collected as these were previously shown to be recognized in the context of an immune response against the pathogen. The IEDB includes epitopes from all four serotypes (DENV1, DENV2, DENV3 and DENV4). The information associated with positive epitopes was exported as an excel file from the results page. This information included Epitope ID, peptide sequence, source antigen and source organism. Epitope sequences were collected from the exported excel file and saved in FASTA format. Dengue type, Epitope ID and position in polyprotein were added to the FASTA description line.

Identification of potential immunodominant antigenic region: The IEDB analysis resource was used to predict the antigenic region¹⁷ conserved in all dengue serotypes. The epitope conservancy analysis tool was used to calculate the degree of conservancy of an epitope within a given protein sequence set at different degrees of sequence identity¹⁸. The positive epitopes of each serotype were screened against its respective consensus sequence. Epitopes having 100 per cent identity with its target consensus sequence were collected and saved in a single FASTA file. These collected sequences were again screened against all consensus sequences to predict antigenic regions conserved in all dengue serotypes. Epitopes having 100 per cent identity with four consensus sequences were considered highly conserved, and thus ideal candidates for EVs useful against all serotypes during dengue viral infection.

Results

Sequence collection and generation of consensus sequence: All available whole genomes were retrieved from the NCBI nucleotide database. Whole genome information for the four types was collected in GenBank format. All four serotypes were differentiated, and protein sequences were collected from coding DNA sequences row with the sequences saved in FASTA format. Collected sequences were saved as separate files with FASTA extension. MSAs were generated for each serotype, and these MSAs were used to generate consensus sequences for further prediction of a conserved antigenic region among all serotypes.

Collection of experimentally proved positive epitopes: The IEDB includes both positive and negative epitopes; however, only positive epitopes (DENV1: 168, DENV2: 645, DENV3: 119 and DENV4: 79) were used to screen for potential antigenic regions.

Identification of immunodominant antigenic regions: To predict a conserved antigenic region among all four serotypes, epitopes from each dengue serotype (DENV1-289, DENV2-1290, DENV3-315 and DENV 4-95) were screened against their respective consensus sequence to find the degree of similarity between these sequences. Epitope sequences having 100 per cent identity with the target sequences were collected. A sum of 87 epitopes (DENV1-2, DENV2-55, DENV3-6 and DENV4-24) with 100 per cent identity to the target sequence were identified. These 87 sequences were combined together in a single FASTA file. These sequences were screened against all four serotypes to find the conserved immunodominant antigenic region.

The antigenic region VDRGWGNGCGLFGKG was found to have no genetic variation and was conserved in all of the dengue polyproteins. This region was considered immunodominant as it was conserved in all four dengue serotypes. It is expected that this region will not allow antigenic resistance. Among these 87 sequences, only 15 epitopes had 100 per cent identity with this immunodominant antigenic region. These 15 epitopes were from the envelope glycoprotein and were experimentally tested against DENV with positive responses. Since this region was conserved among all four serotypes, it was expected to be potentially protective against all four dengue types. The Table provides information about predicted 15 potential immunodominant antigenic regions along with the experimental details obtained from the IEDB for the predicted epitopes¹⁹⁻²².

Discussion

Traditional vaccines are based on the intact pathogen, either inactivated or live attenuated, or proteins²³. EVs are able to immunize using the minimal structure, consisting of a well-defined antigen, with respect to its antigenicity and immunogenicity. However, in clinical usage, some EVs have failed to protect against pathogens. Therefore, development of a vaccine based on conserved epitopes may avoid such antigenic resistance²⁴.

Computational approaches in predicting T-cell epitopes and analysis of sequence conservation have been widely used for the epitope-based vaccination against various pathogens such as HIV-1, Vaccinia virus, DENV, West Nile virus, bacteria and protozoan parasites²⁵. It is a powerful alternative strategy for the experimental discovery of candidate epitopes which is less expensive²⁶. The major concern with the experimental approach for screening of positive epitopes is that these may be too narrow in scope and may not be effective against other pathogens or in different hosts. This challenge for EVs may be overcome by identifying epitopes conserved across multiple strains/serotype of an organism to protect against a genetically diverse community²⁷. The *in silico* prediction of potential immunodominant antigenic regions is another approach that can greatly reduce the time and effort involved in screening potential epitopes²⁸. Thus, this approach can be used as a template for the analysis of other pathogens, providing a novel and generalized approach to the formulation of EVs that are effective against a broad diversity of pathogens²⁷.

Table. Details concerning predicted 15 potential immunodominant antigenic regions that have 100 per cent identity with all four serotypes

Epitope number	IEDB epitope ID	Source organism	Epitope sequence	Position in polypeptide	Model system	Detection method
1	23250	YFV and dengue fever virus	GWGNGCGLF	381-389	Transgenic mice	Peptide-HLA-I binding assay ¹⁹
2	6309	DENV2 virus env gp	CGLFGK	385-390	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
3	9978	DENV2 virus env gp	DRGWGN	378-383	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
4	18866	DENV2 virus env gp	GCGLFG	384-389	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
5	20850	DENV2 virus env gp	GLFGKG	386-391	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
6	21431	DENV2 virus env gp	GNGCGL	382-387	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
7	23249	DENV2 virus env gp	GWGNGC	380-385	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
8	43963	DENV2 virus env gp	NGCGLF	383-388	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
9	54016	DENV2 virus env gp	RGWGNG	379-384	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
10	68057	DENV2 virus env gp	VDRGWG	377-382	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
11	72483	DENV2 virus env gp	WNGCGC	381-386	Human	Solid-phase binding assay & anti-peptide immunoassay ²⁰
12	72484	DENV2 non-structural-1 and env gp	WNGCGLFG	381-389	BALB/c mice and mice	ELISA and immunoblot assays ²¹
13	9979	DENV2 virus env gp	DRGWGNGC	378-385	Human, mouse and rabbit	ELISA ²²
14	43964	DENV2 virus env gp	NGCGLFGK	383-390	Human, mouse and rabbit	ELISA ²²
15	78288	DENV2 virus non-structural-1 and env gp	VDRGWGNGC	377-385	BALB/c mice	ELISA and immunoblot assays ²¹

YEV, yellow fever virus; IEDB, Immune Epitopes Database; HLB, human leucocyte antigen; DENV, dengue viruses; env gp: Envelope glycoprotein

In this study, an immunodominant antigenic region VDRGWGNGCGLFGKG was found to be conserved (100%) in all dengue polypeptide sequences. Conserved regions in the viral protein sequences are the prime target for EV formulation²⁹. The use of this antigenic region in EVs would confer broader protection across multiple strains³⁰ because these antigenic regions are conserved throughout the evolution. In this study 15 experimentally proven epitopes with 100 per cent identity to this immunodominant antigenic region were identified. Since the immunodominant antigenic region is conserved in all dengue serotypes, it is anticipated that identified peptides from DENV1 and DENV2 would also elicit immune responses against

all other serotypes. The antigenicity of the identified polypeptide was previously verified using the ELISA method¹⁹⁻²². Hence, these 15 polypeptides may be considered as potentially protective epitopes for use in dengue vaccination and are expected to overcome the problem of the shift of antigenic sites due to drug resistance and may also have the potential to trigger an immune response against all four dengue serotypes.

Acknowledgment

Authors acknowledge the infrastructure facilities provided by the Department of Bioinformatics, Alagappa University, Postgraduate Diploma in Structural Pharmacogenomics (funded by UGC, Government of India).

Conflicts of Interest: None.

References

1. *Dengue and severe dengue*. Geneva: World Health Organization; 2014. Available from: <http://www.who.int/mediacentre/factsheets/fs117/en/>, accessed on January 8, 2014.
2. *Dengue*. Delhi: Directorate of National Vector Borne Disease Control Programme (NVBDCP); Available from: <http://www.nvbdc.gov.in/den-cd.html>, accessed on January 13, 2016.
3. Whitehead SS, Blaney JE, Durbin AP, Murphy BR. Prospects for a dengue virus vaccine. *Nat Rev Microbiol* 2007; 5 : 518-28.
4. Goldsby RA, Kindt TJ, Kuby J, Osborne BA. *Immunology*. 5th ed. New York: W. H. Freeman; 2002.
5. Khan AM, Miotto O, Heiny AT, Salmon J, Srinivasan KN, Nascimento EJ, *et al.* A systematic bioinformatics approach for selection of epitope-based vaccine targets. *Cell Immunol* 2006; 244 : 141-7.
6. Rothbard JB, Geftter ML. Interactions between immunogenic peptides and MHC proteins. *Annu Rev Immunol* 1991; 9 : 527-65.
7. McFarland BJ, Beeson C. Binding interactions between peptides and proteins of the class II major histocompatibility complex. *Med Res Rev* 2002; 22 : 168-203.
8. Bremel RD, Homan EJ. An integrated approach to epitope analysis I: dimensional reduction, visualization and prediction of MHC binding using amino acid principal components and regression approaches. *Immunome Res* 2010; 6 : 7.
9. Good MF, Berzofsky JA, Miller LH. The T cell response to the malaria circumsporozoite protein: an immunological approach to vaccine development. *Annu Rev Immunol* 1988; 6 : 663-88.
10. Sette A, Keogh E, Ishioka G, Sidney J, Tangri S, Livingston B, *et al.* Epitope identification and vaccine design for cancer immunotherapy. *Curr Opin Investig Drugs* 2002; 3 : 132-9.
11. Toussaint NC, Kohlbacher O. Towards *in silico* design of epitope-based vaccines. *Expert Opin Drug Discov* 2009; 4 : 1047-60.
12. *Nucleotide*. Bethesda, MD: National Library of Medicine (US); 1999. Available from: <https://www.ncbi.nlm.nih.gov/nucleotide>, accessed on July 15, 2013.
13. Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, *et al.* *Geneious 5.4* [Software Created by Biomatters Ltd., Auckland, New Zealand], 2011. Available from: <http://www.geneious.com>, accessed on October 9, 2014.
14. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013; 30 : 772-80.
15. Feng DF, Doolittle RF. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J Mol Evol* 1987; 25 : 351-60.
16. Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, *et al.* The immune epitope database 2.0. *Nucleic Acids Res* 2010; 38 : D854-62.
17. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, *et al.* Immune epitope database analysis resource. *Nucleic Acids Res* 2012; 40 : W525-30.
18. Bui HH, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics* 2007; 8 : 361.
19. Lund O, Nascimento EJ, Maciel M Jr., Nielsen M, Larsen MV, Lundegaard C, *et al.* Human leukocyte antigen (HLA) class I restricted epitope discovery in yellow fever and dengue viruses: importance of HLA binding strength. *PLoS One* 2011; 6 : e26494.
20. Innis BL, Thirawuth V, Hemachudha C. Identification of continuous epitopes of the envelope glycoprotein of dengue type 2 virus. *Am J Trop Med Hyg* 1989; 40 : 676-87.
21. Falconar AK. Monoclonal antibodies that bind to common epitopes on the dengue virus type 2 nonstructural-1 and envelope glycoproteins display weak neutralizing activity and differentiated responses to virulent strains: implications for pathogenesis and vaccines. *Clin Vaccine Immunol* 2008; 15 : 549-61.
22. Aaskov JG, Geysen HM, Mason TJ. Serologically defined linear epitopes in the envelope protein of dengue 2 (Jamaica strain 1409). *Arch Virol* 1989; 105 : 209-21.
23. Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. *Open Biol* 2013; 3 : 120139.
24. Claesson MH. Why current peptide-based cancer vaccines fail: lessons from the three Es. *Immunotherapy* 2009; 1 : 513-6.
25. Desai DV, Kulkarni-Kale U. T-cell epitope prediction methods: an overview. *Methods Mol Biol* 2014; 1184 : 333-64.
26. Sánchez-Burgos G, Ramos-Castañeda J, Cedillo-Rivera R, Dumonteil E. Immunogenicity of novel dengue virus epitopes identified by bioinformatic analysis. *Virus Res* 2010; 153 : 113-20.
27. Bremel RD, Homan EJ. An integrated approach to epitope analysis II: a system for proteomic-scale prediction of immunological characteristics. *Immunome Res* 2010; 6 : 8.
28. De Groot AS, Berzofsky JA. From genome to vaccine – New immunoinformatics tools for vaccine design. *Methods* 2004; 34 : 425-8.
29. Tong JC, Ren EC. Immunoinformatics: current trends and future directions. *Drug Discov Today* 2009; 14 : 684-9.
30. Staneková Z, Varecková E. Conserved epitopes of influenza A virus inducing protective immunity and their prospects for universal vaccine development. *Virology* 2010; 7 : 351.

Reprint requests: Dr Karthikeyan Muthusamy, Department of Bioinformatics, Science Campus, Alagappa University, Karaikudi 630 004, Tamil Nadu, India
e-mail: mkbioinformatics@gmail.com