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 countries. In India, 201771 infected cases and 920 deaths were reported from 2007 to 2013 and in 2015 alone, 99913 infected cases were reported.

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 vaccines. 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, treatment, and also represent a new strategy for prophylactic and therapeutic induction of pathogen-specific immunity. 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 responses. MHC-II molecules bind longer peptides (15-20 amino acids), whereas MHC-I molecules bind shorter peptides (9 amino acids or less). Prediction of such epitope binding is critical. Hence, various computational algorithms and methods have been used for the prediction of epitopes and mostly conserved epitopes should be considered for vaccine development.

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 breadth. Antigenic regions induce a more specific immune response to a disease compared to other vaccination methods. 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 epitopic 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 vaccines.

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 database. 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) 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 Plugin. 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) 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 region\textsuperscript{17} 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 identity\textsuperscript{18}. 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 DENV4-95) 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 VDRWGNGCGLFGKG 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 epitopes\textsuperscript{19-22}.

Discussion

Traditional vaccines are based on the intact pathogen, either inactivated or live attenuated, or proteins\textsuperscript{23}. 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 resistance\textsuperscript{24}.

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 parasites\textsuperscript{25}. It is a powerful alternative strategy for the experimental discovery of candidate epitopes which is less expensive\textsuperscript{26}. 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 community\textsuperscript{27}. The \textit{in silico} prediction of potential immunodominant antigenic regions is another approach that can greatly reduce the time and effort involved in screening potential epitopes\textsuperscript{28}. 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 pathogens\textsuperscript{27}. 
In this study, an immunodominant antigenic region VDRGWNGCGLFGKG was found to be conserved (100%) in all dengue polyprotein sequences. Conserved regions in the viral protein sequences are the prime target for EV formulation\(^\text{29}\). The use of this antigenic region in EVs would confer broader protection across multiple strains\(^\text{30}\) 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 method\(^\text{19-22}\). 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).

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

<table>
<thead>
<tr>
<th>Epitope number</th>
<th>IEDB epitope ID</th>
<th>Source organism</th>
<th>Epitope sequence</th>
<th>Position in polyprotein</th>
<th>Model system</th>
<th>Detection method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23250</td>
<td>YFV and dengue fever virus</td>
<td>GWGNGCGLF</td>
<td>381-389</td>
<td>Transgenic mice</td>
<td>Peptide-HLA-I binding assay(^\text{19})</td>
</tr>
<tr>
<td>2</td>
<td>6309</td>
<td>DENV2 virus env gp</td>
<td>CGLFGK</td>
<td>385-390</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>3</td>
<td>9978</td>
<td>DENV2 virus env gp</td>
<td>DRGWGN</td>
<td>378-383</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>4</td>
<td>18866</td>
<td>DENV2 virus env gp</td>
<td>GCGLFG</td>
<td>384-389</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>5</td>
<td>20850</td>
<td>DENV2 virus env gp</td>
<td>GLFGKG</td>
<td>386-391</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>6</td>
<td>21431</td>
<td>DENV2 virus env gp</td>
<td>GNGCGL</td>
<td>382-387</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>7</td>
<td>23249</td>
<td>DENV2 virus env gp</td>
<td>GWGNGC</td>
<td>380-385</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>8</td>
<td>43963</td>
<td>DENV2 virus env gp</td>
<td>NCGGLF</td>
<td>383-388</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>9</td>
<td>54016</td>
<td>DENV2 virus env gp</td>
<td>RGWGN</td>
<td>379-384</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>10</td>
<td>68057</td>
<td>DENV2 virus env gp</td>
<td>VDRGWG</td>
<td>377-382</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>11</td>
<td>72483</td>
<td>DENV2 virus env gp</td>
<td>WGNGCG</td>
<td>381-386</td>
<td>Human</td>
<td>Solid-phase binding assay &amp; anti-peptide immunoassay(^\text{20})</td>
</tr>
<tr>
<td>12</td>
<td>72484</td>
<td>DENV2 non-structural-1 and env gp</td>
<td>WGNGCGGLFG</td>
<td>381-389</td>
<td>BALB/c mice and mice</td>
<td>ELISA and immunoblot assays(^\text{21})</td>
</tr>
<tr>
<td>13</td>
<td>9979</td>
<td>DENV2 virus env gp</td>
<td>DRGWGN</td>
<td>378-385</td>
<td>Human, mouse and rabbit</td>
<td>ELISA(^\text{22})</td>
</tr>
<tr>
<td>14</td>
<td>43964</td>
<td>DENV2 virus env gp</td>
<td>NCGGLFG</td>
<td>383-390</td>
<td>Human, mouse and rabbit</td>
<td>ELISA(^\text{22})</td>
</tr>
<tr>
<td>15</td>
<td>78288</td>
<td>DENV2 virus non-structural-1 and env gp</td>
<td>VDRGWNGC</td>
<td>377-385</td>
<td>BALB/c mice</td>
<td>ELISA and immunoblot assays(^\text{31})</td>
</tr>
</tbody>
</table>

YEV, yellow fever virus; IEDB, Immune Epitopes Database; HLB, human leucocyte antigen; DENV, dengue viruses; env gp: Envelope glycoprotein
Conflicts of Interest: None.

References


