



OPEN

## Immunoinformatics design of a novel epitope-based vaccine candidate against dengue virus

Adewale Oluwaseun Fadaka<sup>1</sup>✉, Nicole Remaliah Samantha Sibuyi<sup>1</sup>, Darius Riziki Martin<sup>1</sup>, Mediline Goboza<sup>1</sup>, Ashwil Klein<sup>2</sup>, Abram Madimabe Madiehe<sup>1,3</sup> & Mervin Meyer<sup>1</sup>✉

Dengue poses a global health threat, which will persist without therapeutic intervention. Immunity induced by exposure to one serotype does not confer long-term protection against secondary infection with other serotypes and is potentially capable of enhancing this infection. Although vaccination is believed to induce durable and protective responses against all the dengue virus (DENV) serotypes in order to reduce the burden posed by this virus, the development of a safe and efficacious vaccine remains a challenge. Immunoinformatics and computational vaccinology have been utilized in studies of infectious diseases to provide insight into the host–pathogen interactions thus justifying their use in vaccine development. Since vaccination is the best bet to reduce the burden posed by DENV, this study is aimed at developing a multi-epitope based vaccines for dengue control. Combined approaches of reverse vaccinology and immunoinformatics were utilized to design multi-epitope based vaccine from the sequence of DENV. Specifically, BCPreds and IEDB servers were used to predict the B-cell and T-cell epitopes, respectively. Molecular docking was carried out using Schrödinger, PATCHDOCK and FIREDOCK. Codon optimization and in silico cloning were done using JCAT and SnapGene respectively. Finally, the efficiency and stability of the designed vaccines were assessed by an in silico immune simulation and molecular dynamic simulation, respectively. The predicted epitopes were prioritized using in-house criteria. Four candidate vaccines (DV-1–4) were designed using suitable adjuvant and linkers in addition to the shortlisted epitopes. The binding interactions of these vaccines against the receptors TLR-2, TLR-4, MHC-1 and MHC-2 show that these candidate vaccines perfectly fit into the binding domains of the receptors. In addition, DV-1 has a better binding energies of – 60.07, – 63.40, – 69.89 kcal/mol against MHC-1, TLR-2, and TLR-4, with respect to the other vaccines. All the designed vaccines were highly antigenic, soluble, non-allergenic, non-toxic, flexible, and topologically assessable. The immune simulation analysis showed that DV-1 may elicit specific immune response against dengue virus. Moreover, codon optimization and in silico cloning validated the expressions of all the designed vaccines in *E. coli*. Finally, the molecular dynamic study shows that DV-1 is stable with minimum RMSF against TLR4. Immunoinformatics tools are now applied to screen genomes of interest for possible vaccine target. The designed vaccine candidates may be further experimentally investigated as potential vaccines capable of providing definitive preventive measure against dengue virus infection.

Approximately 12.5% of the world population are affected by a group of pathogens known to cause neglected tropical diseases (NTDs) (CDC, 2020). This diverse group of tropical infections are common among the low-income countries in developing regions of the world. The adverse effects caused by these NTDs include but are not limited to impaired physical and cognitive development, and trapping the poor in a cycle of poverty and disease<sup>1</sup>. Apart from condemning affected people to live with disability and stigma, these NTDs deprive children from attending school, threaten job security, increase the financial burden on countries due to medical cost and in turn negatively affecting the economy of developing countries<sup>2–4</sup>. The ever growing list of NTDs includes,

<sup>1</sup>Department of Science and Innovation/Mintek Nanotechnology Innovation Centre, Biolabels Node, Department of Biotechnology, Faculty of Natural Sciences, University of the Western Cape, Bellville, South Africa. <sup>2</sup>Plant Omics Laboratory, Department of Biotechnology, Faculty of Natural Sciences, University of the Western Cape, Private Bag X17, Bellville 7535, Cape Town, South Africa. <sup>3</sup>Nanobiotechnology Research Group, Department of Biotechnology, Faculty of Natural Sciences, University of the Western Cape, Bellville, South Africa. ✉email: afadaka@uwc.ac.za; memeyer@uwc.ac.za

parasitic helminths, bacterial, protozoan, fungal, ectoparasitic and viral diseases. The top 10 hotspots where NTDs dominate include sub-Saharan Africa, Asia, Oceania and the Middle East<sup>5</sup>. For extremely pathogenic NTDs novel therapies are being sought. Although probably less pathogenic or less frequently encountered such as dengue fever, which is neglected despite an apparent increase in prevalence<sup>6</sup>.

The endemic nature of dengue virus has been implicated in more than 100 countries notably in regions of Africa, Americas, Eastern Mediterranean, South-East Asia and the Western Pacific. The largest number of global dengue infection cases ever reported was in 2019 affecting all regions whereas dengue virus transmission was discovered in Afghanistan for the first time. Due to the asymptomatic and mild nature of this disease, the actual number of cases are under-reported with many being misdiagnosed as other febrile illnesses<sup>7</sup>. Dengue is reported as the second most diagnosed cause of fever after malaria and also the dengue fever (DF) being the most rapidly spreading mosquito-borne disease<sup>8</sup>. The rate of infection of dengue virus by *Aedes* mosquito species (*Aedes aegypti* and *Aedes albopictus*) increases geometrically with annual infection of about 400 million and 22,000 deaths<sup>9</sup>. Some of the symptoms include high fever accompanied by severe headache, anorexia, abdominal discomfort, maculopapular rash, fatigue, muscle and joint pain, unpleasant metallic taste in the mouth, loss of appetite, vomiting and diarrhea<sup>10,11</sup>. The World Health Organization (WHO) classified DF into uncomplicated and severe DF. In addition, dengue transmission is attributed to four serotypes of the dengue virus (DENV 1–4), where, individuals can be infected by all the four serotypes in their lifetime.

There are several potential dengue vaccines at different stages of development in clinical trial with only one licensed for use. Most of these vaccines are primarily developed from the envelope proteins prM and E, which are expected to elicit protective immune responses in humans<sup>13,14</sup>. Nonetheless, it is essential to consider that the human immune response to DENV is dominated by highly cross-reactive antibodies equipped with neutralizing and enhancing activity<sup>15,16</sup>. The Dengvaxia manufactured by Sanofi Pasteur is a live, attenuated, and tetravalent recombinant vaccine called ChimeriVax<sup>12</sup>. This formulation was the first vaccine to be licensed against dengue in the year 2015 due to its promising results obtained in various clinical studies<sup>17</sup>. In May 2019, Dengvaxia was licensed and approved by the US Food and Drug Administration (FDA) in the United States in areas where dengue was prevalent. This vaccine is available in 19 countries, and its use is limited to for a particular age range and cannot be administered in Flavivirus-naïve individual<sup>18</sup>. Its formulation consists of chimeric viruses constructed by infectious clone technology. Individuals who receive the vaccine and have not been previously exposed to DV may likely develop severe dengue if they are infected after vaccination<sup>19,20</sup>. Although Dengvaxia was evaluated and licensed for human use in 20 countries, it does not contain the non-structural proteins of DENV which presents low protective efficacy and increases their risk of hospitalization<sup>21</sup>. In 2018 the safety issues from the phase III clinical trial of Dengvaxia was published<sup>22</sup>. Due to the increased rate of hospitalization occurring in seronegative vaccinated children between the age of 9–11, Dengvaxia was predicted as vaccine failures rather than as vaccine enhanced dengue disease<sup>23</sup>.

Other vaccine formulations against DENV include live attenuated tetravalent vaccine for dengue (LATV), TAK-003, tetravalent dengue (TDEN), dengue purified inactivated vaccine (DPIV), tetravalent DNA vaccine against dengue (TVDV), and genetic constructs of the tetravalent vaccine formulation (V180)<sup>24</sup>. The LATV was developed by the National Institute of Allergy and Infectious Diseases (NIAID) using recombinant DNA technology<sup>25</sup>. The formulation of this anti-dengue formulation contains the structural and non-structural DENV proteins and the trial study was proven to be safe<sup>26,27</sup>. LATV is currently in phase III clinical trial. The main challenge of LATV is that it requires a significant amount of DENV2 for a balanced seroconversion and the efficacy is yet unknown<sup>28</sup>. Tak-003 produced by Takeda Pharmaceutical Company Limited, is based on an attenuated virus and chimeric viruses constructed using recombinant DNA technology<sup>29</sup>. Currently in phase III clinical trial, Tak-003 in the was shown to be safe, immunogenic, and effective against all the four DENV serotypes with reduced cellular immune response<sup>30,31</sup>. Its efficacy was reported to be 80.2%. However, Tak-003 induces low levels of antibodies to DENV<sup>32</sup>.

V180 is another anti-dengue potential vaccine registered under the ClinicalTrials.gov is under the phase I clinical trial. It is composed of the recombinant forms of the DENV envelope glycoprotein<sup>33</sup>. Currently, V180 is being developed by Merck and the vaccine formulation is adjuvanted by ISCOMATRIX<sup>™</sup><sup>34</sup>. In preclinical study, V180 was proven to be safe and immunogenic due to the high quality of the recombinant proteins and ISCOMATRIX<sup>™</sup> adjuvant. Nonetheless, the immunogenic properties of V180 is dependent on adjuvants and three-dose immunization regimen is required<sup>35</sup>. TDEN, DPIV, and TVDV produced by the U.S. Army Medical Research and Materiel Command have been reported to be immunogenic, well-tolerated and safe. While DPIV and TVDV are in phase I clinical trial, TDEN is in phase II clinical trial<sup>36–38</sup>. Challenges associated with these vaccines include general body pain, neutralizing antibody titers decrease over time, TVDV presents only envelop (E) and membrane glycoprotein precursor (M) proteins and requires both adjuvant and high amount of vaccine antigen<sup>39</sup>. Due to the tetravalent nature of DENV, effective vaccine, long-term surveillance for potential adverse events, compatibility with current vaccine schedules, cost and stability are the current challenges in the implementation of the DENV vaccines. Therefore, the improvement in vaccine development against dengue is much needed.

Structurally, the genome of DENV is made of 10,696 nucleotides and 10,173 of these nucleotides are responsible for the single open reading frame (ORF) which encodes a polypeptide of 3,391 amino acid residues<sup>40</sup>. In addition to its single and positive sense RNA stranded (+ ssRNA), it encodes for both structural and non-structural proteins. Immunologically, DENVs are related but genetically and antigenically distinct with approximately 75% relatedness<sup>41,42</sup>. DENV serotypes are subdivided according to the envelope gene responsible for incomplete cross-defensive immunity towards other serotypes in *Homo sapiens*<sup>43</sup>. As perceived by the immune system, disease development, viral replication and antigenic determinants have been linked with DENV proteins<sup>44,45</sup>. Therefore, any antigenic DENV protein sequences that is fit to elicit immune responses its host could be employed in the identification of vaccine for dengue<sup>46</sup>.

| RefSeq      | Protein name                        | Length (aa) | Accession ID |
|-------------|-------------------------------------|-------------|--------------|
| NP_073286.1 | Polyprotein                         | 3387        | NP_073286    |
| NP_740313.1 | Capsid protein (C)                  | 99          | NP_740313    |
| NP_740315.1 | Membrane glycoprotein precursor (M) | 166         | NP_740315    |
| NP_740317.1 | Envelope protein (E)                | 495         | NP_740317    |
| NP_740318.1 | Nonstructural protein NS1           | 352         | NP_740318    |
| NP_740319.1 | Nonstructural protein NS2A          | 218         | NP_740319    |
| NP_740320.1 | Nonstructural protein NS2B          | 130         | NP_740320    |
| NP_740321.1 | Nonstructural protein NS3           | 618         | NP_740321    |
| NP_740322.1 | nonstructural protein NS4A          | 127         | NP_740322    |
| NP_740324.1 | Nonstructural protein NS4B          | 245         | NP_740324    |
| NP_740325.1 | RNA-dependent RNA polymerase NS5    | 900         | NP_740325    |

**Table 1.** Structural and non-structural protein sequences of DENV 4 retrieved from NCBI.

| Parameters  | prM       | E         | NS1       | NS2A      | NS2B      | NS3       | NS4A      | NS4B      | NS5        |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|
| Total AAs   | 166       | 495       | 352       | 218       | 130       | 618       | 127       | 245       | 900        |
| Mol. w      | 18,709.67 | 53,988.10 | 39,635.01 | 24,039.33 | 13,939.24 | 69,397.18 | 13,910.58 | 26,580.16 | 102,969.54 |
| Theo. pI    | 6.35      | 7.90      | 6.55      | 9.72      | 4.27      | 8.35      | 5.26      | 9.01      | 8.80       |
| Ext. coeff  | 33,835    | 72,140    | 96,660    | 19,730    | 26,470    | 99,600    | 13,980    | 36,565    | 214,125    |
| Half-life   | 1.1       | 30        | 1.1       | 30        | 1.9       | 1.9       | 1.9       | 1.4       | 30         |
| LI          | 47.04     | 24.64     | 44.12     | 48.82     | 29.02     | 33.53     | 34.42     | 28.62     | 38.21      |
| Alip. index | 75.78     | 80.67     | 67.84     | 131.88    | 116.23    | 78.11     | 132.20    | 105.59    | 71.13      |
| GRAVY       | -0.122    | -0.72     | -0.460    | 0.852     | 0.422     | -0.479    | 0.667     | 0.273     | -0.607     |

**Table 2.** Physicochemical properties of the viral protein. <https://web.expasy.org/cgi-bin/protparam/protaram>.

Asides palliative therapy, the treatment of both dengue and severe dengue are not specific and the mode of action by which the immune system react to this infection is unclear. Moreover, successful mosquito control measure to limit the spread of DF is still in its infant stage. As such, epitope-based vaccine development remains a clear solution for this infection.

In this study, several computational approaches were applied to design multiple epitope-based vaccines against DENV that comprises multiple epitope, which elicit the activation of cytotoxic T lymphocytes (CTLs) and helper T lymphocytes (HTLs). These vaccines require simple peptides that are easily chemically synthesized rather than pathogens. These approaches present an advantage over conventional methods of vaccine development by reducing the chance of allergenic reactions. In addition, these methods are rapid, easy and cost-effective. Although, similar approaches have been widely employed for the construction of vaccines against viruses such as SARS-CoV-2<sup>47</sup>, Ebola<sup>48</sup>, Zika<sup>49</sup>, and Chikungunya<sup>50</sup>, to date, no commercial and licensed epitope-based vaccine are available on the market using these bioinformatics approaches<sup>51,52</sup>.

## Results

**Sequence selection.** In order to develop an epitope/ multi-epitope based vaccine for DEN-4, the protein sequences of DENV (structural and non-structural sequences) were obtained from NCBI and UniProt. The process of component sequence selection was modified from Ali et al.<sup>53</sup>. These components are presented in Table 1. Their physical and chemical properties are also presented in Table 2. Structural proteins are responsible for host invasion and particle assembly. On the other hand, enzyme synthesis which aids viral replication and structural protein synthesis are attributed to non-structural proteins. After infection, they prompt different immune responses. Prior to downstream analysis, the retrieved sequences were subjected to antigenicity and transmembrane helices prediction, and physicochemical analysis. All the three structural proteins were confirmed antigenic at a threshold of 0.4. While the non-structural components NS2A, NS2B and NS4A were non-antigenic due to their low prediction score, NS1, NS3 and NS5 were predicted to be antigenic. With a prediction score of 0.544, non- structural NS5 was the most antigenic viral component of DEN-4 (Table 3).

The TMHMM Server DTU Bioinformatics prediction tool revealed the position and amino acid range of the sequences of the selected viral components (Table 4). The capsid protein and selected sequence length of the envelop protein (466–471) were predicted to be localized inside the cell. All selected antigenic non-structural protein components and membrane glycoprotein precursor sequences were localized outside the cell.

| Viral component            | VaxiJen score | Probability   | Localization      |
|----------------------------|---------------|---------------|-------------------|
| Capsid protein             | 0.4938        | Antigenic     | Mitochondria      |
|                            | 0.4725        | Antigenic     | Plasma membrane   |
| Envelope protein           | 0.4090        | Antigenic     | Cytoplasmic       |
| Nonstructural protein NS1  | 0.4340        | Antigenic     | Extracellular     |
| Nonstructural protein NS2A | 0.0365        | Non-antigenic | Plasma membrane   |
| Nonstructural protein NS2B | 0.1671        | Non-antigenic | Plasma membrane   |
| Nonstructural protein NS3  | 0.4756        | Antigenic     | Mito, Cytoplasmic |
| Nonstructural protein NS4A | 0.0957        | Non-antigenic | Plasma membrane   |
| Nonstructural protein NS4B | 0.1549        | Non-antigenic | Plasma membrane   |
| Nonstructural protein NS5  | 0.54368       | Antigenic     | Cytoplasmic       |

**Table 3.** Antigenicity prediction of dengue virus proteins. (Threshold=0.4 with tumor model).

| Viral component                 | Position | Amino acids |
|---------------------------------|----------|-------------|
| Capsid protein                  | Inside   | 1–99        |
| Membrane glycoprotein precursor | Outside  | 1–166       |
| Envelope protein                | Outside  | 1–442       |
|                                 | TMhelix  | 443–465     |
|                                 | Inside   | 466–471     |
|                                 | TMhelix  | 472–494     |
|                                 | Outside  | 495–495     |
| Nonstructural protein NS1       | Outside  | 1–352       |
| Nonstructural protein NS3       | Outside  | 1–618       |
| Nonstructural protein NS5       | Outside  | 1–900       |

**Table 4.** Prediction of transmembrane helices in proteins using the TMHMM Server v2.0.

**Prediction of B-cell epitope.** B-cell epitopes were predicted using BCPREDS for all antigenic structural and non-structural protein components of DENV-4. Predicted epitopes were shortlisted on the basis of prediction score, toxicity, allergenicity, antigenicity, hydrophobicity and surface accessibility. All epitopes with a score  $\geq 0.70$  were selected for vaccine design. A total of 47 epitopes were predicted for all protein components while only 9 epitopes satisfied the set criteria and were selected for further analysis. One epitope each was predicted for the capsid protein, membrane glycoprotein precursor and NS1. Although no B-cell epitope was predicted for the confirmed antigenic sequences of NS3, 4 and 2 B-cell epitopes were shortlisted for NS5 and the envelop components respectively. These selected epitopes were further validated using BepiPred for their structural and relative surface accessibility through NetsurfP's default threshold. Regions of the B-cell epitopes (italicized) were predicted by SVMTrip with a score of 1.00 (Table 5).

**T-cell epitopes predictions and prioritization.** Table 6 shows the final list of the selected epitopes for each of the antigenic viral components. Briefly, the selected viral components of DV were subjected to various MHC I and II epitopes prediction web-based servers (immune epitope database and analysis resources as well as the CTLPred). T cell epitopes (HTL and CTL) were ranked based on stringent in-house criteria before being shortlisted for downstream analysis. These rules of selection include: Good IEDB score, high conservancy, good binding affinity, B-cell epitope overlap,  $\geq 9$ mer for MHC I, 15mer for MHC II, significantly antigenic/immunogenic and topographically accessible to membrane-bound or free antibody. For high immunogenicity, lower percentile ranks and  $IC_{50}$  value were considered. According to these specifications, 21 epitopes each (C-2, M-3, E-4, NS1-1, NS3-3, and NS5-8) and (C-5, E-5, NS1-1, NS3-5, and NS5-5) were shortlisted for MHC-I and MHC II binders for multi-epitope based vaccine designed.

**Immune receptors.** The receptors considered for docking include two immune alleles (HLA-A\*11-01 and HLA-DRB1\*04-01) and two immune molecules toll-like receptors (TLR2 and TLR4). Their respective crystallographic structures with PDB IDs (HLA-A\*11-01: 5WJL and HLA-DRB1\*04-01: 5JLZ; TLR2: 3FXI and TLR4: 2Z7X) were retrieved from the protein data bank. Since proteins from the PDB database often lack coordinates for side-chains on the surface of the protein given the complexity to determine side chain coordinates from x-ray crystallographic data, all proteins were refined through the protein preparation wizard in Schrödinger to improve the overall quality of the output. These refined outputs (Fig. 1) were then used for docking analysis.

| VC  | Pos | Epitopes   | B-Scores | AT-Score | Allergenicity | A-score | Conservancy |
|-----|-----|--|----------|----------|---------------|---------|-------------|
| C   | 57  | <u>SIPPT</u> <sub>A</sub> <u>G</u> <sub>H</sub> <u>KR</u> <sub>W</sub> <u>GQ</u> <sub>I</sub> <u>KKNKA</u> | 0.975    | 0.7003   | Non-a         | 0.83    | 50% (2 / 4) |
| M   | 51  | <u>YKCPL</u> <u>LVNTEPEDIDC</u> <u>WCNL</u>  | 0.856    | 1.0357   | Non-a         | 0.75    | 50% (2 / 4) |
| E   | 68  | <u>ITTATRCPTOGE</u> <u>PYLKEEOD</u>  | 0.999    | 0.7964   | Non-a         | 0.71    | 25% (1 / 4) |
|     | 388 | <u>TLHWFRK</u> <u>GSSIGKMFESTYR</u>  | 0.794    | 0.4258   | Non-a         | 0.63    | 100% (4/4)  |
| NS1 | 104 | <u>LTPPVSD</u> <u>LKYSWKTW</u> <u>GKAKI</u>  | 1.000    | 0.7604   | Non-a         | 0.70    | 75% (3 / 4) |
| NS5 | 96  | <u>NVTE</u> <sub>V</sub> <u>K</u> <sub>G</sub> <u>KG</u> <sub>G</sub> <u>PGHEEPIP</u>                      | 1.000    | 0.5602   | Non-a         | 0.85    | 100% (4/4)  |
|     | 275 | <u>TIIGRR</u> <sub>R</sub> <u>QR</u> <sub>Q</sub> <u>QEEHKETWHY</u>  | 0.987    | 0.5902   | Non-a         | 0.77    | 25% (1 / 4) |
|     | 57  | <u>RGSSK</u> <sub>R</sub> <u>WTV</u> <u>ERGM</u> <sub>V</sub> <u>KPKGK</u>                                 | 0.976    | 0.4875   | Non-a         | 0.77    | 25% (1 / 4) |
|     | 533 | <u>DDTAG</u> <sub>D</sub> <u>DTR</u> <sub>T</sub> <u>TEDD</u> <sub>T</sub> <u>ONEEL</u>                    | 0.970    | 0.6909   | Non-a         | 0.87    | 75% (3 / 4) |

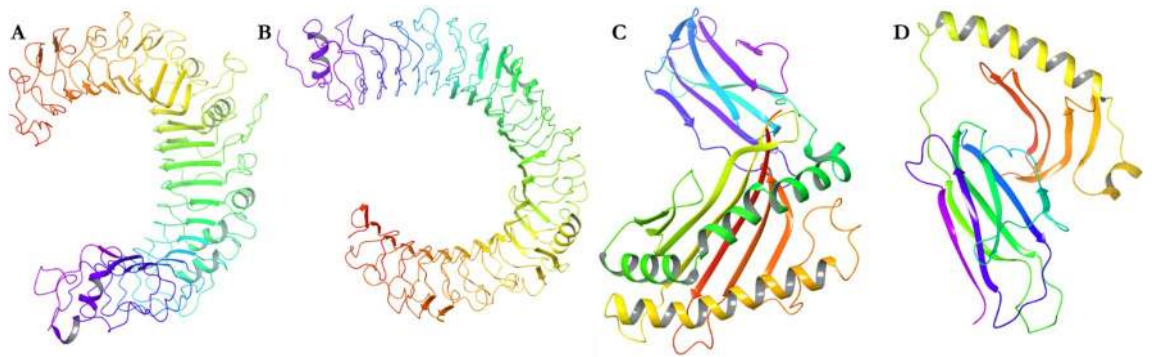
**Table 5.** Selected B-cell epitopes. The B-cell epitopes were predicted using BCPREDS web-based tool and were validated by BepiPred 2.0 software. Epitopes are represented as BCPREDS predicted. Green colored epitopes were predicted by BepiPred. Additionally, underlined amino acids are structurally coiled whereas the amino acids (not underlined) are structurally helix. For surface accessibility, subscripted amino acids are buried as against the exposed amino acids. VC viral component; Pos. position; B-score BCPred score, A-score ABCpred score, AT-score antigenicity score. For antigenicity prediction, the threshold was set to 0.4, which implies that any predicted epitope with antigenicity score  $\geq 0.4$  is said to be antigenic. In addition, regions of the B-cell epitopes predicted by SVMTrip with a score of 1.00 were italicized.

|     | MHC I  | MHC II  | B-Cell  |
|-----|--|---|---|
| C   | VVRPPFNMLK<br>SIPPTAGILK   | FLRVLSSIPPTAGILK<br>KEIGRMLNILNGRKR<br>LRVLSSIPPTAGILKR<br>PFNMLKRERNRVSTP<br>FNMLKRERNRVSTPQ | <u>SIPPT</u> <sub>A</sub> <u>G</u> <sub>H</sub> <u>KR</u> <sub>W</sub> <u>GQ</u> <sub>I</sub> <u>KKNKA</u>  |
| M   | LLFKTTEGINK<br>GTCTQSGER<br>VAKHERGRPLLFK  |   | <u>YKCPL</u> <u>LVNTEPEDIDC</u> <u>WCNL</u>   |
| E   | VTFKVPFAK<br>MSYTMCSGK<br>VTFKVPFAKR<br>TTAKEVALLR   | NHGVMTAMITPRSPV<br>HGVMTAMITPRSPVE<br>HKQWFLDPLPWTAG<br>YKERMVTFKVPFAKR<br>KQWFLDPLPWTAGA     | <u>ITTATRCPTOGE</u> <u>PYLKEEOD</u><br><u>TLHWFRK</u> <u>GSSIGKMFESTYR</u>  |
| NS1 | LMSAAIKDQK   | EQYKFQPEPARLAS  | <u>LTPPVSD</u> <u>LKYSWKTW</u> <u>GKAKI</u>   |
| NS3 | KVASAGISY<br>HVQTKPGLFK<br>VQTKPGLFK   | AAIFMTATPPGATDP<br>AIFMTATPPGATDPF<br>KSGDYVSAITQAERI<br>SGDYVSAITQAERIG<br>TKSGDYVSAITQAER   |   |
| NS5 | ATLKNVTEVK<br>YVDYMPVMK<br>RGAVMDIISRK<br>LTYQNKVVK<br>RSVSTETEK<br>KVDTRTPQPK<br>CTREEFISK<br>WSYYMATLK | LGKAYAQMWSLMYFH<br>TREEFISKVRSNAAI<br>CLGKAYAQMWSLMYF<br>PWDVIMPVTQLAMTD<br>ACLGKAYAQMWSLMY   | <u>NVTE</u> <sub>V</sub> <u>K</u> <sub>G</sub> <u>KG</u> <sub>G</sub> <u>PGHEEPIP</u><br><u>TIIGRR</u> <sub>R</sub> <u>QR</u> <sub>Q</sub> <u>QEEHKETWHY</u><br><u>RGSSK</u> <sub>R</sub> <u>WTV</u> <u>ERGM</u> <sub>V</sub> <u>KPKGK</u><br><u>DDTAG</u> <sub>D</sub> <u>DTR</u> <sub>T</sub> <u>TEDD</u> <sub>T</sub> <u>ONEEL</u> |

**Table 6.** Composite table of the prioritized epitopes for vaccine development. Underlined residues are structurally coiled whereas the amino acids (not underlined) are structurally helix. For surface accessibility, subscripted amino acids are buried as against the exposed amino acids. In addition, regions of the B-cell epitopes predicted by SVMTrip with a score of 1.00 were italicized.

**Vaccine construction.** From the prioritized epitopes, four vaccines were constructed, with the aim to elicit an immune response against DENV-4. Three different adjuvants i.e. beta defensin, L7/L12 ribosomal protein and HABA protein were used to construct four different vaccines. PADRE sequence is a pan HLA-DR epitope peptide that was used for vaccine construction to enhance their potency with minimal toxicity. The PADRE sequence improves the CTL response as it has the capability to bind with different MHC class-II molecules with high binding affinity (Wu et al., 2010). Also, appropriate linkers were used for the conjugation of the prioritized epitopes during the design process. The final vaccine lengths for DV1-4 were 352, 336, 336 and 400 respectively (Table 7).

**Designed vaccine properties analysis.** The various vaccine components including their specific properties are presented in Table 3. In addition to the different adjuvants used for the design, the antigenicity, allergenicity, toxicity and surface accessibility of these vaccines were also evaluated (Table 8). All designed vaccines



**Figure 1.** The prepared 3D structures of the Toll-like and Major histocompatibility complex receptors used in this study: (A) TLR-2; PDB ID: 2ZX; (B) TLR-4; PDB ID: 3FX1; (C) MHC class I; PDB ID: 5WJL; and (D) MHC class II; PDB ID: 5JLZ.

| Vaccine name                  | Length | Vaccine constructs   |
|-------------------------------|--------|--|
| Dengue virus vaccine-1 (DV-1) | 352    | GIINTLQKYYCRVRCVAVLSCLPKKEEQIGKCCSTRGRKCCRRKKEAAAK <b>KAFVAAWTLKAA</b> AGGGSVVRPPFNM-LKAAAYLLFKTTTEGINKAAYVTFKVPHAKAAYLMSAAIKDQKAAYKVASAGISYAAATLKNVTEVKGGSFL-RVLSIPPTAGILKGGPGPNHGVTAMITPRSPSVGPGPEQYKQPESPARLASGPGGAAIFMTATPPGATDPG-PGPGLGKAYAQMWSLMYFHGGSSIPPTAGILKRWGQLKKNKAKKYC <b>PLLVNTEPEDIDCWCNLKITTATRCPTQGEPYLKEEQDKKLT</b> PPVSDLKYSWKTWGWKAKIKKNVTEVKGYSKGGPGHEEPI                              |
| Dengue virus vaccine-2 (DV-2) | 336    | MAKLSTDELLKEMTLELSDVFKFEETFEVTAAPVAVAAAGAAPAGAAVEAAEEQSEFDVILEAAGDKKIGVIVVREIVS-GLGLKEAKDLVDGAPKPLLEKVAKEAADEAKLEAAGATVTVKEAAAK <b>KAFVAAWTLKAA</b> AGGGSIPPTAGILKAAAYGTC-TQSGERAAYMSYTMCSGKAAYHVQTKPGLFKAAYYVDYMPVMKGGGSKIEGRMLNINLGRKRGPGGPHGVTAMITPRSPSVEG-PGPAIFMTATPPGATDPFGPGPGTREFEFISKVRSNAAIGGGSTLHWFRKSSIGKM <b>FESTYRKKTIIGRRLQRLQEEHKETWHY</b>   |
| Dengue virus vaccine-3 (DV-3) | 336    | MAENPNIDDLPLAALGAADLALATVNDLIANLRERAETRAETRTTRVEERRARLTKFQEDLPEQFIELRDKFT-TEELRKAEEGYLEAATNRYNELVERGEAALQRLRSQTALEDASARAEGYVDQAVELTQEALGTVASQTRAVGERA-AKLVGIELEAAAK <b>KAFVAAWTLKAA</b> AGGGSVAKHERGRPLLFKAAYVTFKVPHAKRAAYVQTKPGLFKAAYRGA-VMDIISRKGGGSLRVLSIPPTAGILKRGPGPHKQWFLDPLPWTAGGPGPGKSGDYVSAITQAERIGPGPGCLGKAYAQM-WSLMYFGGSRGSSKIRWIVERGMVVKPKGK   |
| Dengue virus vaccine-4 (DV-4) | 400    | MSDINKLAENLKIVEVNDLAKILKEKYGLDPSANLAIPSLPKAEILDKSKEKTSFDLILKAGSAKLTVVKRIK-DLIGLGLKESKDLVDNVPKHLKGLSKEAEALSQQLEEVGAELKEAAAK <b>KAFVAAWTLKAA</b> AGGGSSTAK-EVALLRAAYLTYQNKVVKAAYRSVSTETEKAAKYVDTRTPQPKAACTREEFISKAAYWSYYMATLKGGSFPNM-LKRERNRVSTPGPGPFNMLKRERNRVSTPQGGPGYKERMVTFKVPHAKRGPQWFLDPLPWTAGAG-PGPGSGDYVSAITQAERIGPGPGTKSGDYVSAITQAERIGPGPGPVDVIMVMTQLAMTDGPGGACLGKAYAQMWSLMYGGGS-DDTAGWDRITREDDLQNEEL |

**Table 7.** Predicted multi-epitope based vaccines. The distinct bold sequences are the pan HLA DR-binding epitope (PADRE) while the B cell epitopes are italicized.

| Vaccines | Adjuvant used | Number of HTL | Number of CTL | Number of B cell | Length | Antigenicity | Allergenicity | Toxicity | Surface accessibility |
|----------|---------------|---------------|---------------|------------------|--------|--------------|---------------|----------|-----------------------|
| DV-1     | BD            | 6             | 5             | 5                | 352    | 0.6136       | NA            | NT       | Outside               |
| DV-2     | RP            | 5             | 4             | 2                | 336    | 0.6184       | NA            | NT       | Outside               |
| DV-3     | HBHA          | 4             | 4             | 1                | 336    | 0.6945       | NA            | NT       | Outside               |
| DV-4     | RP            | 6             | 8             | 1                | 400    | 0.6764       | NA            | NT       | Outside               |

**Table 8.** Composition and assessment of the primary sequence of the multi-epitope based vaccine construct. *BD* beta defensing, *RP* ribosomal protein, *HBHA* Heparin-binding hemagglutinin; *NA* non-allergen, *NT* non-toxic.

are antigenic in nature, non-allergenic and non-toxic. Furthermore, their surface accessibilities were predicted to be localized outside of the cell.

The physicochemical properties of these vaccines were computed using various web-based servers. Parameters evaluated include; solubility, amino acid composition (AAS), molecular weight Mol. W), theoretical pI (Theo. pI), extinction coefficient (Ext. coeff), half-life, instability index (I.I), aliphatic index (A.I), and the grand average of hydrophobicity (GRAVY) (Table 9). DV-1 had the highest predicted solubility and Theo. pI of 0.636 and 9.87 respectively. However, DV-4 had the highest AAs sequence and predicted ext. coeff of 400 and 59,360 M<sup>-1</sup> cm<sup>-1</sup> respectively. All of the designed vaccines were found to have similar predicted half-life of 30 h. However, DV-3 had the highest predicted aliphatic index (82.08) while DV-2 had the highest GRAVY of -0.167 among the designed vaccines. However, further laboratory studies are required to investigate the accuracy of these results.

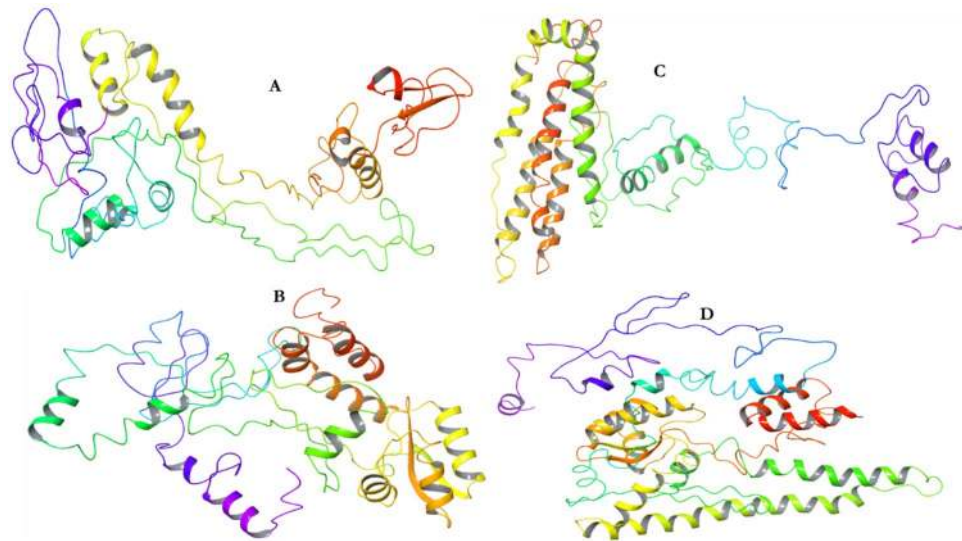
| Parameters         | DV-1      | DV-2      | DV-3      | DV-4      |
|--------------------|-----------|-----------|-----------|-----------|
| Solubility         | 0.636     | 0.555     | 0.556     | 0.597     |
| No. of amino acids | 352       | 336       | 336       | 400       |
| Molecular weight   | 37,586.86 | 35,295.60 | 36,255.51 | 42,958.22 |
| Theoretical pI     | 9.87      | 9.14      | 9.68      | 9.72      |
| Ext. coeff         | 55,975    | 29,910    | 40,910    | 59,360    |
| Half-life          | 30 h      | 30 h      | 30 h      | 30 h      |
| Instability Index  | 36.22     | 38.17     | 33.45     | 29.34     |
| Aliphatic index    | 68.58     | 75.39     | 82.08     | 76.17     |
| GRAVY              | -0.376    | -0.167    | -0.331    | -0.449    |

**Table 9.** Physicochemical properties of the predicted multi-epitope based vaccine construct.

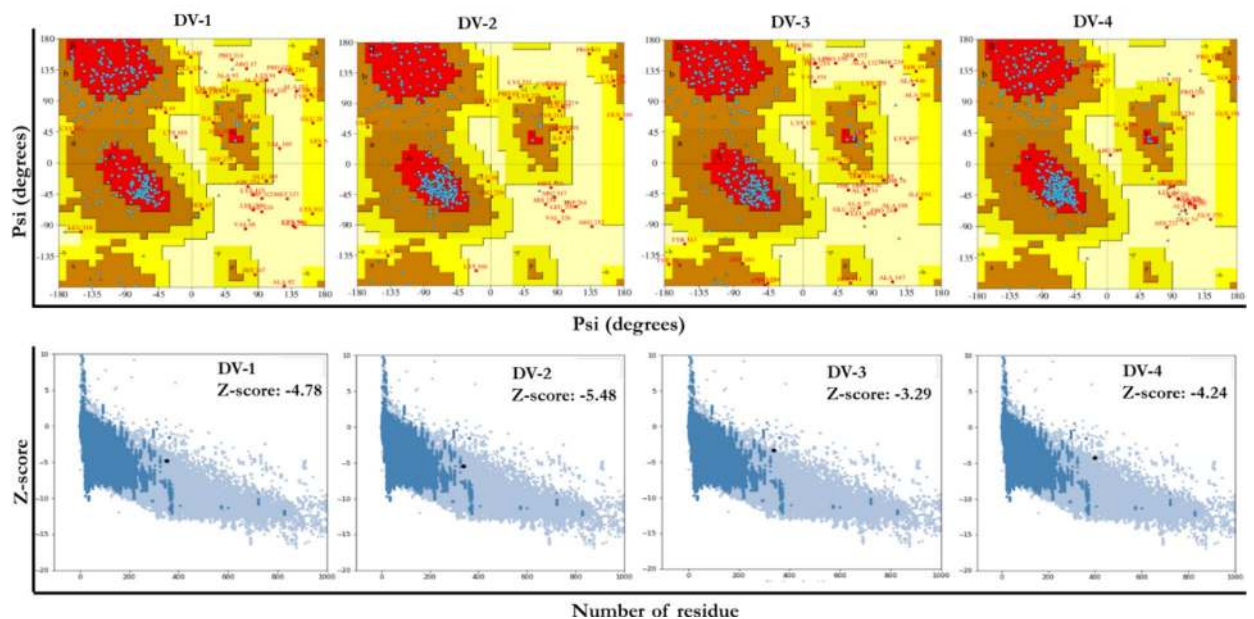


**Figure 2.** The secondary structures of the predicted dengue virus vaccines (PSIPRED).

**Secondary structural prediction and validation of the designed vaccines.** The secondary structure of the four designed vaccines were predicted by PSIPRED (Fig. 2). The structures are composed of coils, helices and strands with no structural deformities. Epitopes often are found in coils and needs to be exposed. The sequences were then modelled by several prediction tools, including Schrödinger, to ensure accuracy (Fig. 3). The knowledge-based method of homology modeling of prime module was employed under Schrödinger soft-



**Figure 3.** The predicted multi-epitope vaccines for DENV. A) DV-1; B) DV-2; C) DV-3; and (D) DV-4 as modelled by Schrödinger.



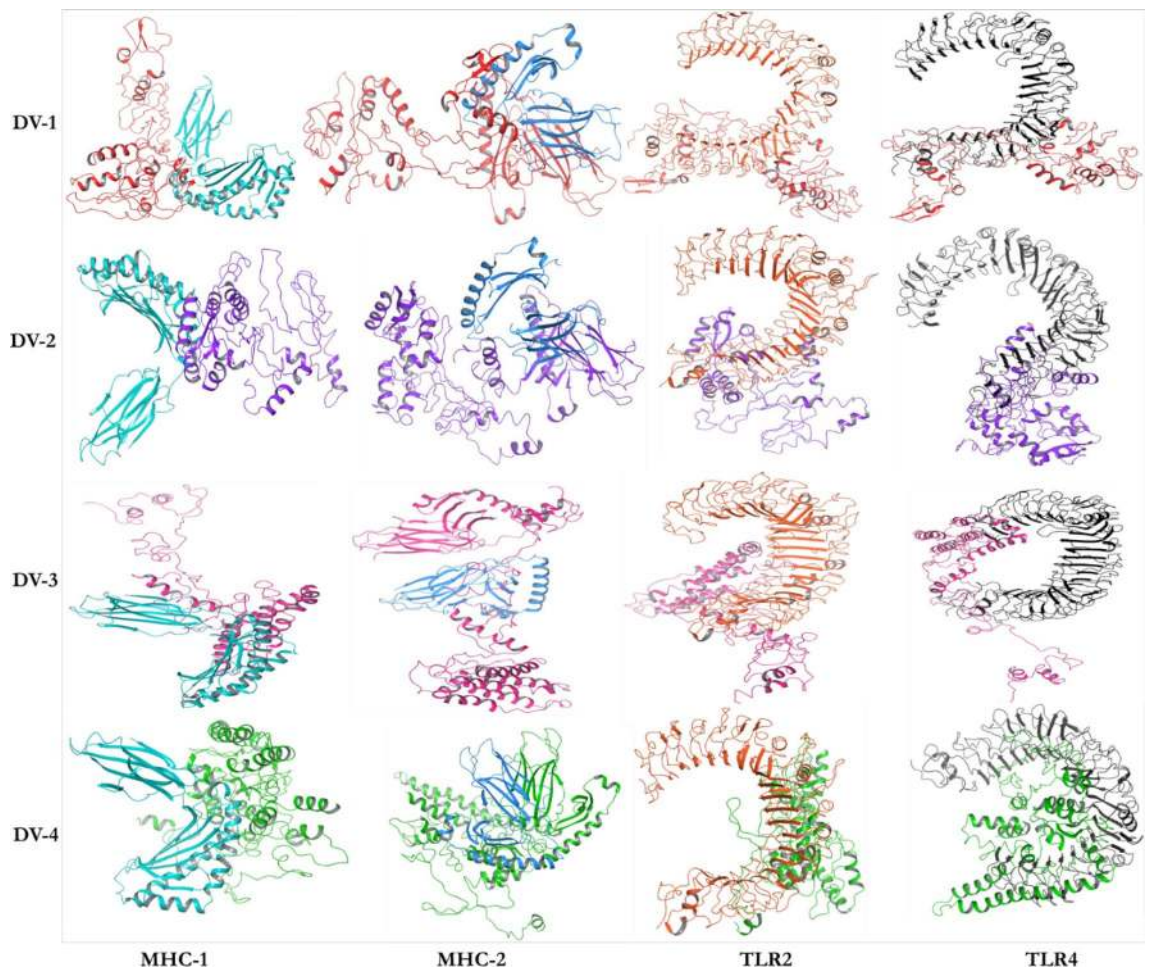
**Figure 4.** Validation of the modeled-predicted DENV vaccines with ProSA-Web.

ware. For DV-1, the template ID: 2LX0 with percentage identity of 58% and expected value (e-value) of  $1 \times 10^{-8}$  was taken into consideration. For DV-2 to 4, the templates PDB IDs include 1DD3, 3J00, and 3J7Z with % identity of 84%, 58% and 43%, respectively. The e-values of DV-2 to 4 were  $2.4 \times 10^{-16}$ ,  $1.2 \times 10^{-2}$ , and  $6.0 \times 10^{-21}$ , respectively. The coverage of these modeled templates are 37%, 39%, 35% and 30%, respectively for DV-1 to 4.

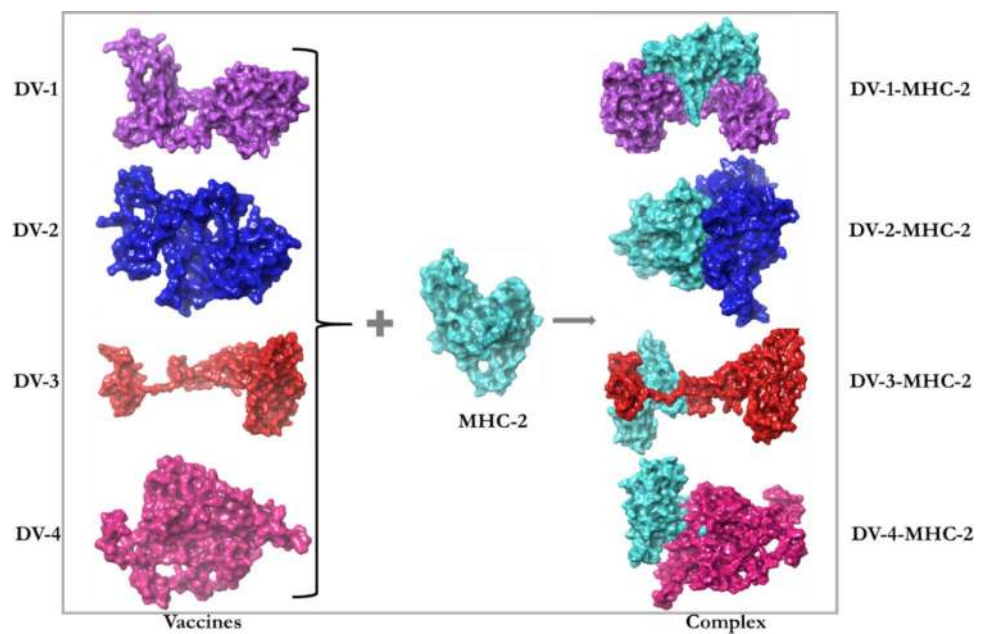
The modelled outputs were further validated by PROCHECK and ProSA-web servers prior to docking analysis (Fig. 4).

**Docking study of the designed vaccines against the immune molecules.** The interactions between an antigenic molecules and an immune receptors molecules are crucial to effectively transport the antigenic molecule and to activate immune response<sup>54,55</sup>. Therefore, docking analysis was carried out between the immune receptor molecules (MCHs and TLRs) and the designed vaccines (DV-1–4) in order to investigate possible interactions, binding energy and poses. Docking analysis was carried out using Schrödinger as well as PATCHDOCK and FIREDOCK to enhance the accuracy of the prediction (Fig. 5). Schrödinger was specifically used to evaluate the generated poses including their solid molecular surface display (Figs. 6 and 7)<sup>56</sup>. The PATCH-

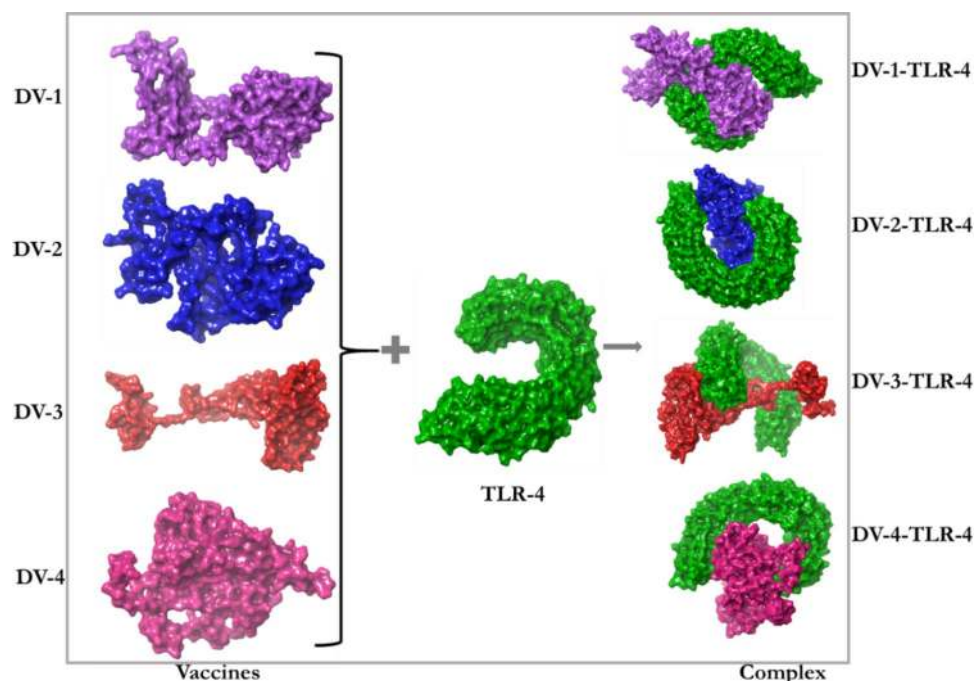




**Figure 5.** Docking poses of the predicted DENV vaccines against selected immunological receptors (Schrodinger suite v2020-3).



**Figure 6.** The solid molecular surface display of the predicted vaccines docked to the MHC class-2 receptor.

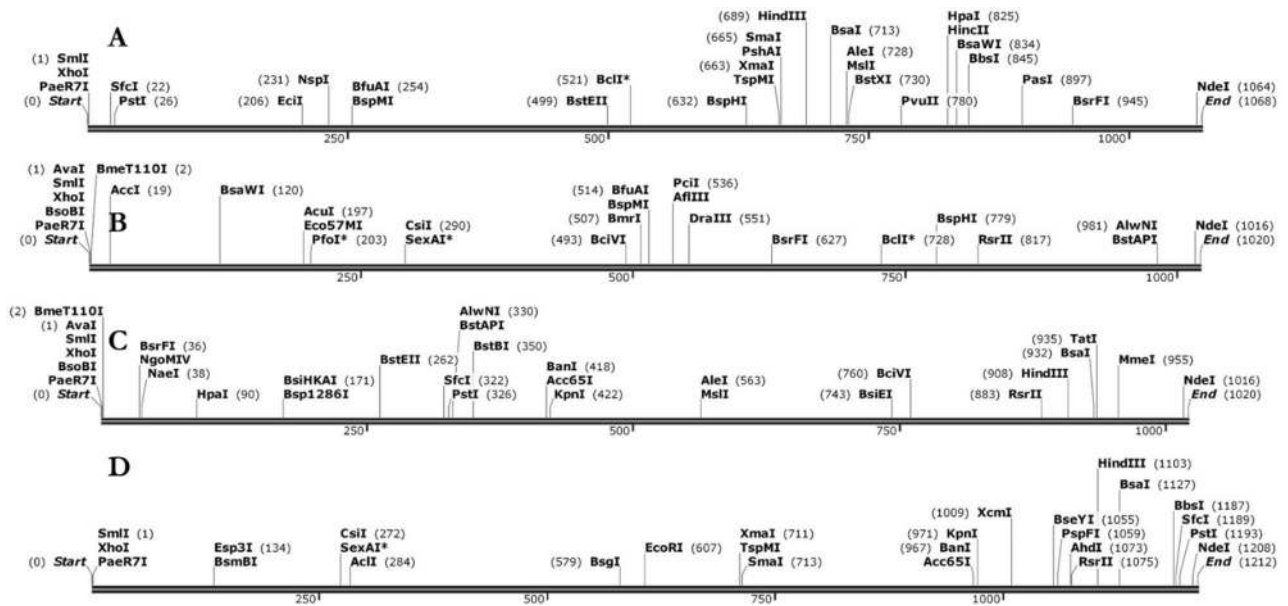


**Figure 7.** The solid molecular surface display of the predicted vaccines docked to the TLR-4 receptor.

| Vaccine construct | Receptor | Name of target | Score  | Area     | H-bond E | Global E | ACE      |
|-------------------|----------|----------------|--------|----------|----------|----------|----------|
| DV-1              | 5WJL     | HLA-A*11-01    | 10,204 | 14,12.80 | - 4.81   | - 60.07  | - 114.00 |
|                   | 5JLZ     | HLA-DRB1*04-01 | 10,056 | 1240.30  | - 2.19   | - 51.02  | - 140.36 |
|                   | 3FXI     | TLR2           | 15,766 | 2591.40  | - 3.20   | - 63.40  | - 33.79  |
|                   | 2Z7X     | TLR4           | 16,202 | 1880.60  | - 5.73   | - 69.89  | 103.14   |
| DV-2              | 5WJL     | HLA-A*11-01    | 9098   | 1523.60  | - 1.90   | - 48.11  | 145.04   |
|                   | 5JLZ     | HLA DRB1*04-01 | 9072   | 1387.70  | - 1.24   | - 59.30  | - 89.11  |
|                   | 3FXI     | TLR2           | 15,784 | 2264.40  | - 1.14   | - 47.31  | 162.29   |
|                   | 2Z7X     | TLR4           | 17,282 | 2301.00  | - 1.69   | - 50.15  | 13.11    |
| DV-3              | 5WJL     | HLA-A*11-01    | 10,162 | 1642.20  | - 2.51   | - 41.64  | 307.70   |
|                   | 5JLZ     | HLA DRB1*04-01 | 10,740 | 1373.90  | - 3.83   | - 63.84  | - 111.34 |
|                   | 3FXI     | TLR2           | 15,328 | 1816.70  | - 4.09   | - 45.59  | 182.27   |
|                   | 2Z7X     | TLR4           | 15,214 | 1893.10  | - 2.17   | - 60.23  | 121.10   |
| DV-4              | 5WJL     | HLA-A*11-01    | 10,444 | 2290.20  | - 3.85   | - 57.16  | 256.86   |
|                   | 5JLZ     | HLA DRB1*04-01 | 10,350 | 1326.40  | - 3.48   | - 67.22  | - 196.17 |
|                   | 3FXI     | TLR2           | 15,726 | 2291.70  | - 5.90   | - 52.53  | - 99.35  |
|                   | 2Z7X     | TLR4           | 15,420 | 3052.30  | - 6.31   | - 50.39  | 464.55   |

**Table 10.** The binding properties of the docking analysis using PATHDOCK and refined by FIREDOCK. *H-Bond E* Hydrogen bond energy; *Global E* Global energy, *ACE* Atomic contact energy \*Allelic variants.

DOCK server was used to generate the binding scores, area and ACE. The top 100 poses were refined using the FIREDOCK server where the global and hydrogen bond energies were generated for the docking calculation (Table 10). TLRs can efficiently induce the immune response after virus recognition. For TLRs, DV-1 showed the best binding among the four designed vaccines for TLR2 and TLR4 with binding energies of - 63.40 kcal/mol and - 69.89 kcal/mol. Respectively, followed by DV-4 (- 52.53 kcal/mol) and DV-3 (- 60.23 kcal/mol). For the MHC alleles, DV-1 and DV-4 showed the best binding interactions with HLA-A\*11-01 and HLA-DRB1\*04-01 with global energies of - 60.07 kcal/mol and - 67.22 kcal/mol. respectively. For these alleles, DV-4 and DV-3 also showed lower energies compared to DV-2. The docking analysis showed strong interactions as well as effective binding between the designed vaccines and MCH1, MCH2, TLR2 and TLR4. The docking statistics are shown in Table 10.

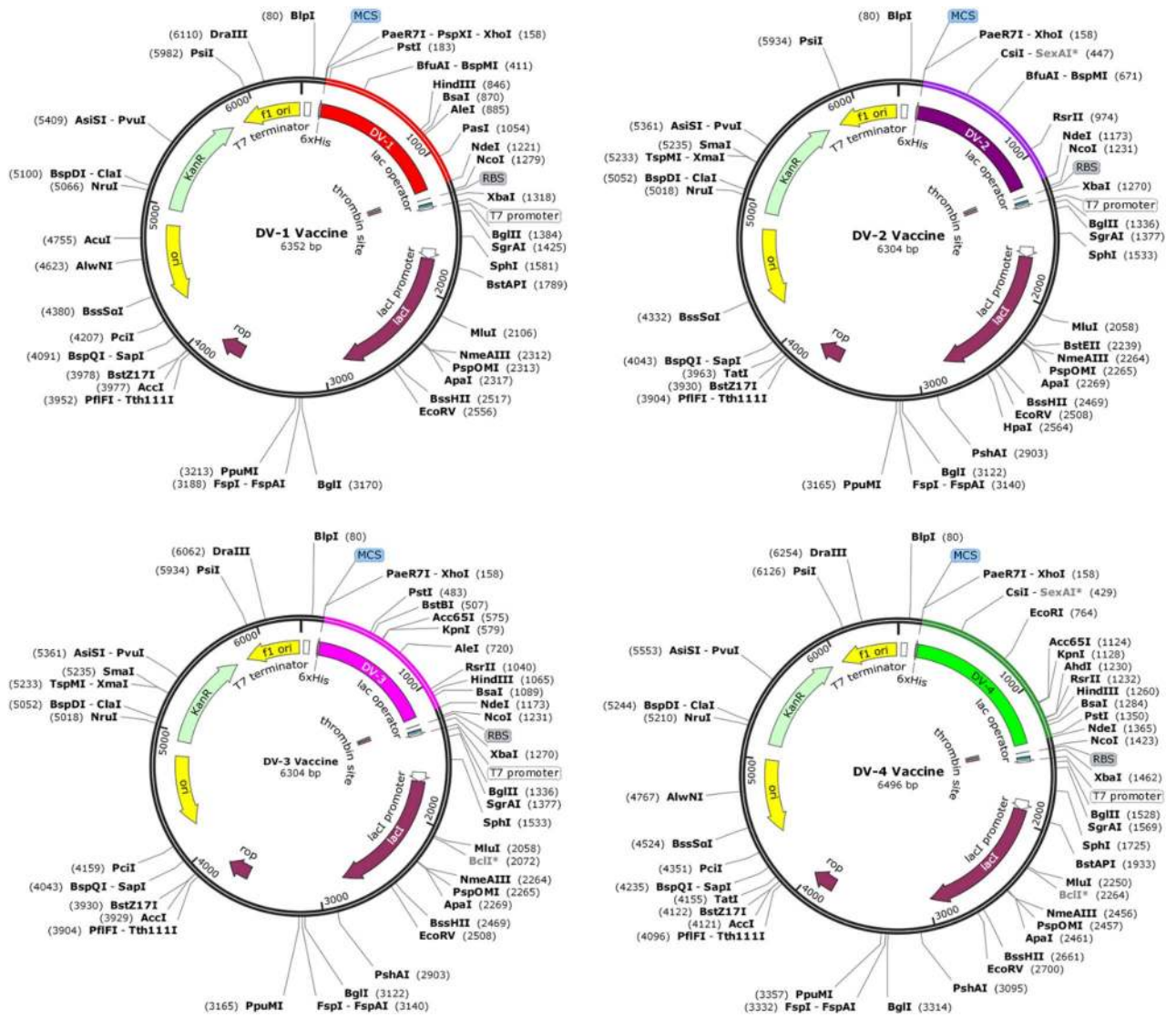


**Figure 8.** The linear vaccine constructs with *XhoI* and *NdeI* sites. (A) DV-1, (B) DV-2, (C) DV-3, and (D) DV-4 (SnapGene software).

**Optimization and in silico cloning of the designed vaccines.** Efficient expression of vaccines into *E. coli* expression system is a crucial step in in silico cloning. Four vaccines (DV-1–4) were designed with the prioritized B-cells and T-cell predicted epitopes with appropriate adjuvants and linkers. The sequences of these vaccines were used as input (individually) in the Java codon adaptation tool (JCAT) in order to adapt the codon usage to most sequenced prokaryotic organisms. The result of this analysis shows that the DNA sequences of DV-1–4 were 1056, 1008, 1008 and 1200 nucleotides respectively. The observed CAIs (DV-1 = 1.0; DV-2 = 1.0; DV-3 = 0.99 and DV-4 = 0.98) indicated that the adapted sequences were made up of codons capable of cellular machinery of the target organism. Furthermore, the GC content of the improved sequences were in the range of 50.99% to 54.27%. This sequence information implies the efficient expression and reliability of the designed vaccines in *E. coli*. Restriction recognition sites of the restriction enzymes *XhoI* and *NdeI* were conjugated at the N and C-terminal of reverse translated nucleotides of DV-1–4 respectively (Fig. 8) and the final optimized sequences were inserted into the pET28a (+) vector in order to clone the designed vaccines using SnapGene software. The final lengths of the cloned plasmids were 6.352, 6.304, 6.304 and 6.396 kbp for DV-1–4, respectively (Fig. 9).

**Immune simulation analysis.** The result of the immune simulation after injection of DV-1 is presented in Fig. 10. In response to several exposures of DV-1 to the host immune system, a significantly high level of the secondary and tertiary antibodies was observed (higher antibody detection) when compared to the primary antibodies. Consequently, rapid clearance of the antigen was detected (decrease in antigen concentration) (Fig. 10A). The different immune cell population/state count is presented in Fig. 10B–H. These immune cells were significantly increased with significant memory cell development; however, T-cells were gradually decreased after their initial increase. Similarly, an increase in the levels of cytokines were observed for IFN- $\gamma$  (>400,000 ng/ml) (Fig. 10). These observations suggest that immune memory development and the vaccine constructs may therefore confer immunity against the dengue virus.

**Conformational stability analysis of the designed vaccines.** Stability of the complex interactions as well as the dynamic behaviour observed between the designed vaccines and the receptors were elucidated through 100 ns MDs at 300 °K. The statistical parameters such as the root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration (rGyr) were used to probe the system for stability and structural adjustment of the vaccine in the binding domains of the selected receptor (Fig. 11A–D). As a measure of the distance between the Ca atoms of superimposed proteins, RMSD was used to evaluate the stability of DV-1 in the binding domain of TLR4 (Fig. 11B). Residual flexibility of DV-1-TLR4 was investigated by RMSF (Fig. 11C) in order to determine the extent and effect of contact deviation on the binding of DV-1 to TLR4. The average RMSF detected for the system was 1.2 Å. In addition; rGyr was used to characterize the secondary elements regular packing in the 3D structure of the DV-1-TLR4 complex and to evaluate complex compactness. The rGyr provides effective information on the tendency of protein structures to expand during MDs<sup>37</sup>. At the time of simulation, the rGyr starts with a sudden increase of 4 ns and remained constant with an average value of 5.98 Å (Fig. 11D). The higher the rGyr the less compact the protein, with less folded stability. Interestingly, the rGyr profile of the system is consistent with the RMSD and RMSF profiles for complex fluctuation<sup>38</sup>.

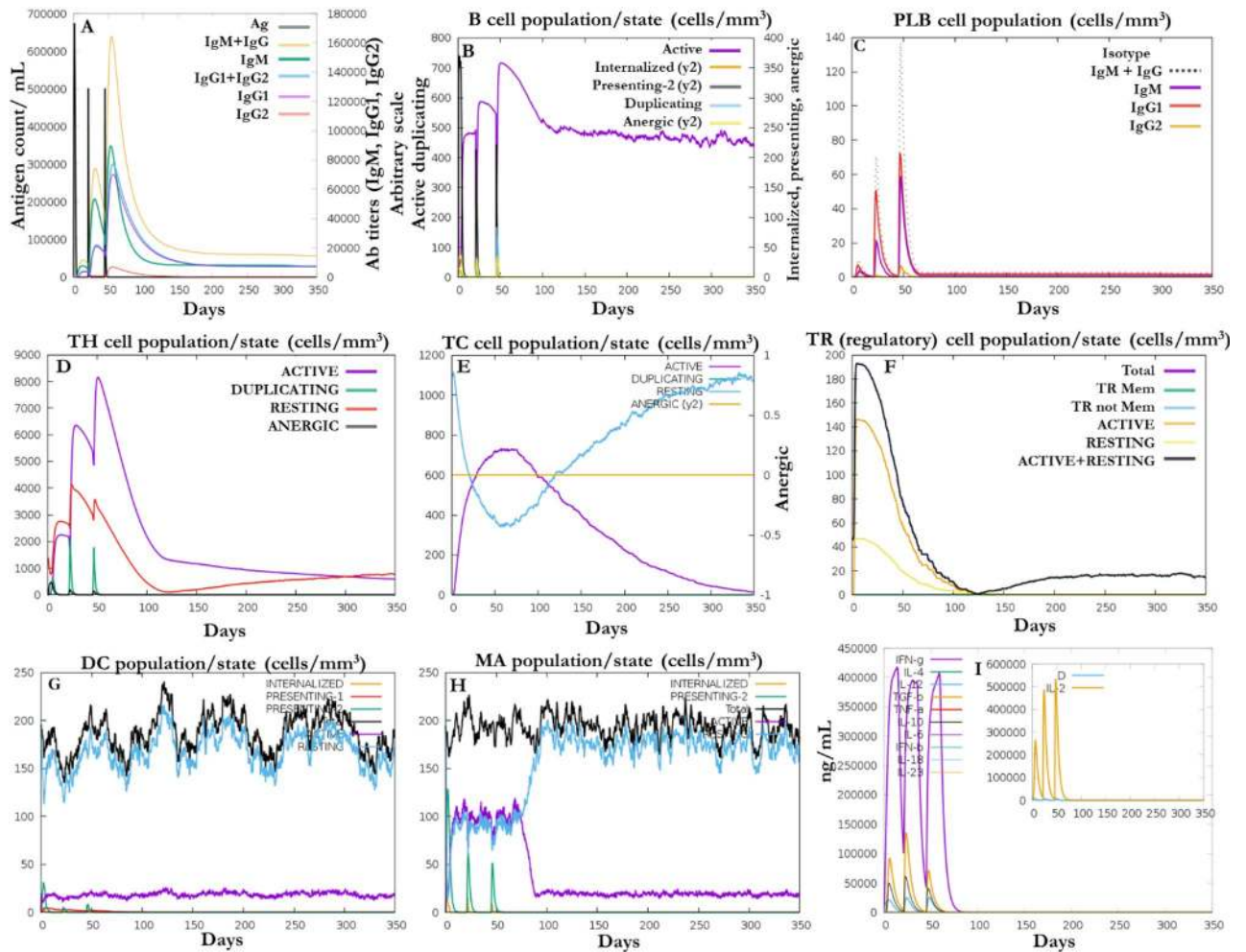


**Figure 9.** In silico restriction cloning of the codon optimized final designed vaccines in the pET-28a (+) vector between the *XhoI* (158) and *NdeI* (238) restriction enzyme sites (SnapGene software). The final constructs can further be expressed in *E coli* (strain K12) for efficient vaccine production.

### Discussion

The use of immunoinformatics approaches in the construction of potent vaccines against various microorganisms especially viruses, are becoming increasingly acceptable as the first line of vaccine development. In recent years, immunoinformatics-guided approaches have been employed in the design of epitope-based subunits for SARS-CoV-2<sup>59–64</sup>. Kar and Srivastava<sup>65</sup> also identified potential vaccine candidates against *Theileria* parasites and the T-cell epitope vaccine candidate for parasitic helminth infection<sup>66</sup>. Dengue is a mosquito-borne viral infection caused by different serotypes of DENV. These serotypes (DENV-1–4) cross-react immunologically with each other<sup>67</sup>. On a global scale, about three billion people are susceptible to DENV infection, of which approximately 96 million cases occur annually<sup>68</sup>. Individuals with primary infection are highly prone to develop dengue hemorrhagic fever and dengue shock syndrome during secondary infection due to antibody-dependent enhancement<sup>69,70</sup>. To date, there is no specific treatment for dengue fever and the main preventive measure is by vaccination in order to reduce the burden of the disease. Hence, this study aimed at designing novel multi-epitope based DENV vaccine candidates capable of evoking immune responses in DENV infected individuals using computational approaches. Epitope-based vaccines have potential advantages over conventional vaccines and they have received considerable attention in recent years.

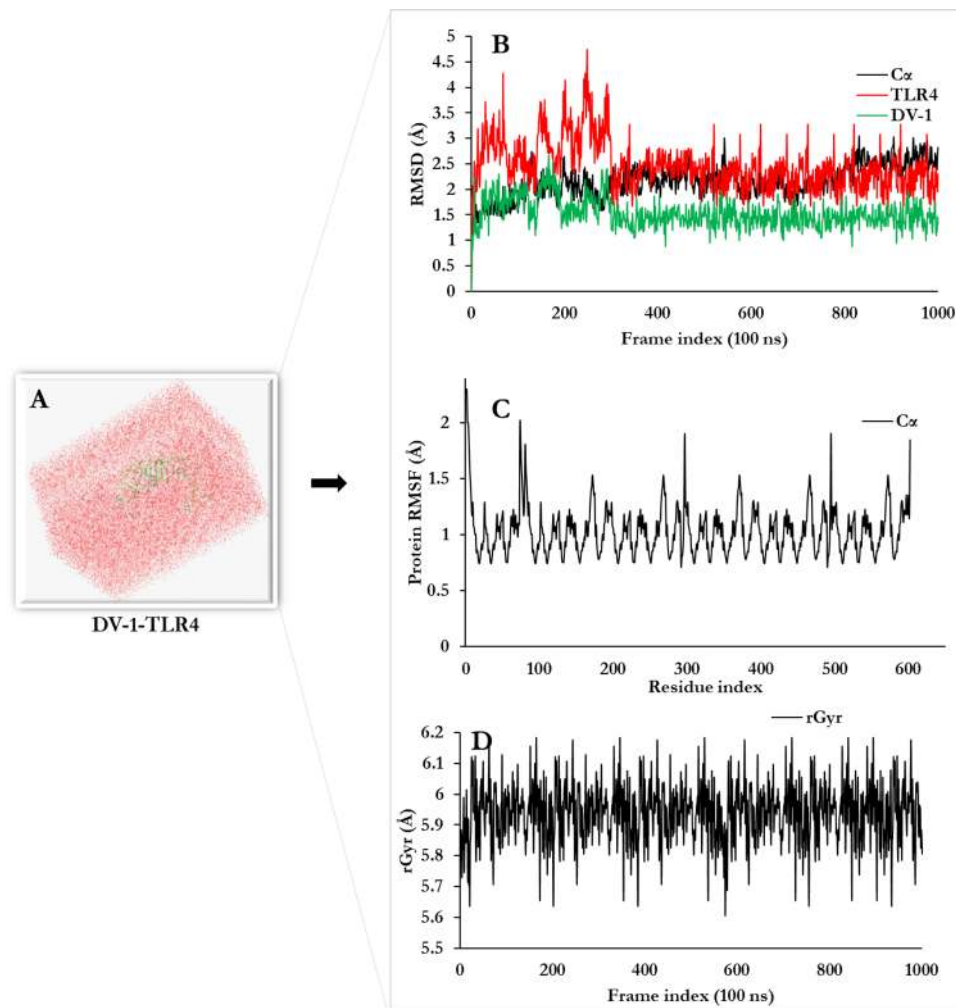
DENV protein components were retrieved and used for the multi-epitope based design of dengue vaccine in this study. The selected components were subjected to physical and chemical properties prediction prior to antigenic determination of the sequences. Using the threshold of 0.4, the viral components were classified into antigens and non-antigens using the tumor model of VaxiJen (VaxiJen v2.0). Scores below this threshold were classified as non-antigen and discarded. For accurate component selection, the localization and transmembrane helices of the protein sequences were further determined. This process ascertains the suitability of candidate



**Figure 10.** Immunogenic potential of the designed vaccine (DV-1). In response to the exposure of DV-1 after three injections, (A) production of immunoglobulins (B) active B-cell populations/state; (C) plasma B-lymphocytes and their isotypes per state; (D) helper T-cell population/state; (E) cytotoxic T-cell population/state; (F) reduction in the level of T regulatory cells; (G) dendritic cell population per state; (H) activity of macrophage population/state; (I) cytokine level and interleukins (smaller plot) in different states with the Simpson index (dotted line). All units are in cells/mm<sup>3</sup> in three subsequent immune responses.

vaccine selection for experimental validation in vaccine development process<sup>71,72</sup>. The final result of all the structural proteins and NS1, NS3, and NS5 from the non-structural proteins were considered for epitope-based vaccine design in this study. Furthermore, the B and T-cell epitopes were predicted from these antigenic viral components. A good vaccine should be able to induce immunity through the antigen with a durable adaptive immunity. A total of 21 antigenic and non-allergenic epitopes were predicted for both MHC classes, while nine antigenic, non-toxic and non-allergenic epitopes were predicted for the B-cells. The CTL epitopes are responsible for developing durable immunity capable of eliminating circulating virus and infected cells<sup>73</sup>. On the other hand, HTL epitopes are associated with the production of both humoral and cellular immune responses. These epitopes elicit a CD4<sup>+</sup> helper response for the generation of protective CD8<sup>+</sup> T-cell memory and activation of B-cells<sup>74</sup>.

The homology modeling of all the predicted multi-epitope vaccines were predicted by Schrödinger and structurally confirmed by three web-based protein structural prediction tools (Phyre<sup>2</sup>, RaptorX, and SCRATCH predictions). The predicted 3D structures of the multi-epitope vaccines were validated using PROCHECK at PDBSum Generate server and ProSA-web. The result of the predicted structural analysis by PROCHECK server is a Ramachandran plot. This plot gives information about the torsional angles (phi ( $\phi$ ) and psi ( $\psi$ )) of the residues contained in the input PDB file. The statistics confirmed that approximately 90% of the predicted vaccine residues are in both the most favoured and additional allowed regions while only few residues were detected in the generously allowed and disallowed regions. The model structures were further queried for potential errors by ProSA-Web. This software tool provides a graphical output of the overall and local model quality based on the input and Z-score of the predicted model. The modeled protein structures of Yadav et al.<sup>75</sup> were reported to be  $-5.98$  kcal/mol and  $6.66$  kcal/mol for template and target, respectively. In addition, an investigation of immunogenic properties of Hemolin using an immunoinformatic approach, showed that the Z-score of the modeled protein was  $-7.08$  kcal/mol<sup>76</sup>. With Z-scores of  $-5.26$  kcal/mol,  $-9.5$  kcal/mol and  $-2.11$  kcal/mol,



**Figure 11.** Molecular dynamics simulation of (DV-1-TLR-4 complex at 100 ns (ns). (A) Molecular dynamics trajectory (system set-up for the simulation), (B) Root Mean Square Deviation of docked complex shows very minimal deviation which reflects the stable microscopic interaction between DV1 and TLR4 molecule (C) RMSF-Root Mean Square Fluctuation plot of docked protein complex side chain fluctuation in plot generates peak which reflects the flexibility of side chain of docked protein complex. (D) Time evolution of the radius of gyration (rRyr) during 100 ns of MD simulation.

Droppa-Almeida et al.<sup>77</sup>, Rekik et al.<sup>78</sup>, and Hashemzadeh et al.<sup>79</sup> concluded that the predicted 3D structures were reliable and of good quality. The overall model quality suggests the qualities of X-ray crystallography with Z-score of  $-4.78$ ,  $-5.48$ ,  $-3.29$ , and  $-4.24$  kcal/mol. for the predicted DV-1-DV-4 vaccines, respectively. Overall the results of the validation tools provide evidence that the predicted vaccine structures are acceptable and of good quality. Therefore, the 3D-structures of the predicted vaccines are reliable and were considered for downstream analysis.

Molecular docking is a crucial tool employed in the field of Bioinformatics to explore specific interactions and binding affinities of ligands towards selected receptors based on the lock and key mechanism. In the field of immunoinformatics, protein-protein docking is usually employed to investigate the best, stable and effective vaccine through their poses, interacting atoms and their binding energies. This procedure was used to examine the best possible vaccine construct(s) against the MHC alleles (HLA-A\*11-01 and HLA-DRB1\*04-01) as well as the immune molecules (TLR2 and TLR4). Several immunoinformatics studies have shown the importance of these receptors in the production of an effective immune response<sup>80-83</sup>. TLRs have a significant function in innate immunity such as immune activation and adaptive immune response. Also, the involvement of TLR2 and TLR4 were previously investigated in the recognition of viral structural proteins leading to inflammatory cytokine production<sup>84</sup>. The docking calculation between the designed constructs (DV-1-4) and the receptors (MHCs and TLRs) showed that the designed vaccines perfectly fit into the binding pockets of the receptors, and therefore, the vaccines may generate stable immune responses. Due to the global energies of the DV-1 and its excellent performance, it was considered for MDs study.

In silico cloning was carried out in a pET28a (+) vector after codon optimization by the JCAT web server in order to avoid codon bias<sup>85</sup>. The CAIs and the GC content of the optimized nucleotides were within the accepted

ranges of 0.8–1.0 and 30–70% respectively. These data confirmed that the designed constructs were reliable with efficient expression in the *E. coli* strain K12.

C-ImmSim is a data-driven immune prediction algorithm and an agent-based simulator of the immune response previously used to study the immunogenic potentials of predicted vaccines<sup>86</sup>. This algorithm simulates the major functional mammal system components (bone marrow, thymus and lymph node). Since a potent vaccine must imitate the natural immunity induced by antigen with the production of long-lasting adaptive immunity, the response of the immune cells (HTL, CTL, B-cells, dendritic cells, immunoglobulins and cytokines) were reported against the designed vaccines. Immune simulation with DV-1 was predicted to trigger IgG, IgM, B-cell, T-cell and cytokines. Similar to the results of previous immunoinformatics vaccine designed studies<sup>87–89</sup>, the designed vaccines may offer protection against DENV.

The stability of the complex (DV-1-TLR4) was validated by performing a 100 ns at 300° Kelvin MDs using the Schrödinger software package. OPLS\_2005 force field and a predefined solvent model of TIP-3P was chosen for the calculation. Furthermore, NaCl was added as (Na<sup>+</sup> and Cl<sup>-</sup>) and the final salt concentration was 0.15 M in order to simulate the background salt at physiological conditions. From the result obtained, a slight fluctuation of TLR4-vaccine complex was observed from the start of the simulation period to about 300 ns after which there was no major fluctuation in the RMSD profile of the complex. The RMSF profile showed the presence of a highly fluctuating N-terminal region at the beginning of the simulation. This result was consistent with previous studies where immunoinformatics of vaccine–receptor complex stabilization was achieved within a similar time-scale using MDs<sup>59,73,90,91</sup>. The rGyr measures the ‘extendedness’ of a ligand, and is equivalent to its principal moment of inertia. The rGyr of the DV-1-TLR4 complexes ranged between 5.5 Å to 6.1 Å. This change in rGyr reveals the compactness of the vaccine construct-TLR4 complexes. This result implies that the designed vaccines can strongly interact with immune receptors.

**Limitations.** This study highlighted an alternative vaccine approach based on multi-epitope construction of the protein components of dengue genome to handle antigenic complexity. Although the predicted vaccines were recommended based on immunoinformatics strategies and believed to be immunogenic, the extent of protection from dengue infection is unknown. Immunoinformatics approaches are extremely useful to perform *in silico* studies and can guide laboratory experiments contributing to save time and money. Nonetheless, the next phase is to carry out *in vitro* immunological assays to validate the predicted vaccines, determine their immunogenicity, and further devise challenge-protection preclinical studies to eventually certify these approaches.

There are some limitations to this current study. First, challenges such as standard benchmark, insufficient prediction methods, and unavailability of precise datasets for different computational analysis are associated with immunoinformatics approach. Secondly, little is known about the pathogenic entry and specific immune-bypass mechanism of dengue within the human host. With this knowledge, the vaccine could be developed more effectively and efficiently to control and/or prevent the pathogenesis of dengue. Thirdly, the proposed vaccine constructs await experimental investigation through *in-vitro* and *in-vivo* bioassays to demonstrate its safety, efficacy, and immunogenicity against dengue by producing an active immunological memory.

On a whole, the application of these results is pending validation in the wet lab experimental models.

## Methodology

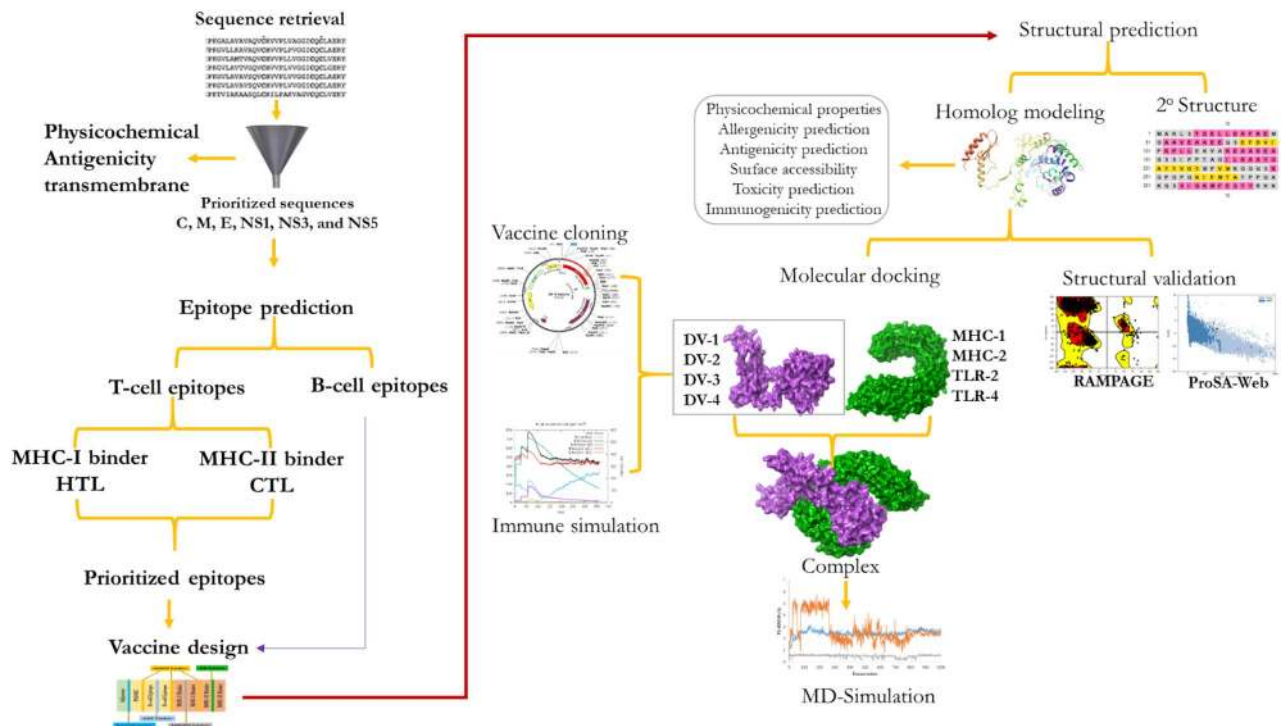
Figure 12 illustrates the stepwise flow of the methodology followed to design a vaccine against DENV.

**Sequence retrieval.** All the protein components of DENV 1–4 were retrieved from the virus pathogen resource (VIPR) database at <https://www.viprbrc.org/brc/home.spg?decorator=vipr> and confirmed with NCBI. These protein components include: the structural proteins (Capsid protein (C), membrane glycoprotein precursor (M), and envelop protein (E)) and the 7 non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). A single file containing all the sequences was generated in FASTA format.

**Protein selection for vaccine preparation.** The complete amino acid sequences of the proteins were retrieved from the NCBI database in the FASTA format. The proteins were subjected to signal peptide analyses using SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>) to discriminate classical secretory and non-secretory proteins. SignalP integrates a prediction of cleavage sites and signal/non-signal peptide predictions based on a combination of several artificial neural networks<sup>92</sup> signal was reported to outperform other predictors in benchmark tests<sup>93</sup>. Localization predictions were conducted using DeepLoc (<http://www.cbs.dtu.dk/services/DeepLoc/>), a template-free algorithm that uses deep neural networks to predict protein subcellular localization using only sequence information with good accuracy<sup>94</sup>.

**Multiple sequence alignment and antigen selection.** The NCBI BLAST program at <https://blast.ncbi.nlm.nih.gov/Blast.cgi> was used to identify highly conserved sequences of the DV. In addition, multiple sequence alignment was achieved using the Multalin server at <https://www.multalin.toulouse.inra.fr/multalin>. The antigenicity of the viral components (structural and non-structural) were evaluated using VaxiJen 2.0 server (<http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html>)<sup>95</sup>. Finally, the most highly conserved sequences and antigenic peptides were selected for downstream analysis.

**Allergenicity prediction.** Allergenicity prediction is a critical step in therapeutics and bio-pharmaceuticals due to their involvement in foods and/or food products. Sequence identity of known allergens is important for predicting the cross-reactive potential of novel proteins and their resistance to pepsin digestion and glyco-



**Figure 12.** Overview of the study design.

sylation status is used to evaluate denovo allergenicity potential<sup>96</sup>. AllerTOP v2.0 is a bioinformatics tool and web-based server used for protein allergenicity prediction in this study. As outlined by Dimitrov et al.<sup>97</sup>, this tool (<https://www.ddg-pharmfac.net/AllerTOP/index.html>) uses an auto cross covariance transformation of protein sequences into uniform equal-length vectors. The method was previously applied to quantitative structure–activity relationship studies of peptides with different lengths using five descriptors such as hydrophobicity, molecular size, helix-forming propensity, relative abundance of amino acids and  $\beta$ -strand forming propensity<sup>98</sup>.

**Antigenicity.** VaxiJen is a web-based server for prediction of protective antigens, tumour antigens and subunit vaccines<sup>95</sup>. This server accessed at <http://www.ddg-pharmfac.net/vaxijen/> was used to determine the best antigenic protein components and to identify epitopes from DENV-4. The viral components' sequences, predicted epitopes, and constructed vaccines were used as input with 'tumor' as selected organism target at a threshold of 4.0.

**B-cell epitope prediction and selection.** The B-cell epitopes were predicted using the BCPreds server at <http://ailab.ist.psu.edu/bcpred/>. This server uses a novel kernel method for predicting linear B-cell epitopes. The epitope length was kept at 20 amino acids with 75% sensitivity. The use of multiple tools in epitope prediction has been reported to improve the rate of true positives in epitopes prediction<sup>99</sup>. Therefore, other prediction servers such as IEDB analysis Resource (<http://tools.iedb.org/population>)<sup>100</sup> and the BepiPred-2.0 web server (<http://www.cbs.dtu.dk/services/BepiPred/>) were used for validation. The method employed by BepiPred-2.0 is more superior to other available tools for sequence-based epitope prediction based on the epitope data derived from solved 3D structures and a large collection of linear epitopes downloaded from the IEDB database. In addition, the proteins were subjected to linear epitope prediction using ABCpred at <https://webs.iitd.edu.in/raghava/abcpred/index.html> to compute ABCpred epitope scores at a threshold of 0.5. At this threshold, the tool shows 65.93% accuracy with equal sensitivity and specificity using a window length of 16. Finally, the server, SVMTrip at <http://sysbio.unl.edu/SVMTriP/prediction.php> was used to improve the prediction performance for the predicted B-cell epitopes.

**T-Cell epitope prediction analysis.** Following the combined methods of Shey et al.<sup>101</sup>, Ali et al.<sup>53</sup> and Ullah et al.<sup>48</sup>, the IEDB epitope analysis resource at [https://www.iedb.org/home\\_v3.php](https://www.iedb.org/home_v3.php) was used to predict the MHC class I and MHC class II binding epitopes. Briefly, the MHC class I restricted CD8<sup>+</sup> cytotoxic T lymphocyte (CTL) epitopes from the antigenic viral components of DENV-4 were predicted using the NetMHCpanEL 4.0 prediction method for HLA-A\*11-01 allele. The result of this method was further subjected to the SMM method in order to prioritize the selected CTL epitopes using the IC<sub>50</sub> threshold of 500 nM and percentile ranks of  $\leq 0.2$ . In addition, the 15-mer MHC class II-restricted CD4<sup>+</sup> helper T lymphocyte (HTL) epitopes were predicted using the NetMHCIIpanEL 4.0 prediction method for HLA-DRB1\*04-01 alleles. The predicted classes of the MHCs were further screened for antigenicity, non-allergenicity, non-toxicity, highly conservancy and transmembrane topology.



**Toxicity and transmembrane topology analysis.** ToxinPred is an *in silico* software tool used for the prediction as well as the design of both toxic and non-toxic peptides. This tool predicts the toxicity of the epitopes according to their physicochemical properties of the input sequences. With the default settings, the predicted sequences were used as input in order to determine their toxicity. Epitopes that were not predicted as allergens by AllerTOP were subjected to toxicity analysis by the ToxinPred webserver at <http://crdd.osdd.net/raghava/toxinpred/>. All the peptides predicted as toxic were discarded. In addition, the transmembrane topology experiment of the viral component and the predicted epitopes were evaluated using the TMHMM v2.0 server (<http://www.cbs.dtu.dk/services/TMHMM/>). This tool is user-friendly for determining the transmembrane topology of epitopes using their sequences as input<sup>102</sup>.

**Analysis of the physicochemical properties and construct solubility.** The physicochemical properties of the viral components and the constructed vaccines were predicted by ExPASy ProtParam at <https://web.expasy.org/protparam/><sup>103</sup>. Characteristics such as number of amino acids, molecular weight, theoretical pI, half-lives, instability indexes, extinction coefficient, aliphatic index and grand average of hydrophobicity were evaluated<sup>104</sup>. Protein-Sol (<https://protein-sol.manchester.ac.uk/>) was used to evaluate the solubility of the multi-epitope based vaccines<sup>105</sup>.

**Epitope conservancy analysis.** The predicted epitopes were subjected to conservancy analysis in order to predict the degree of similarity within the serotypes of dengue virus at sequence identity threshold of  $\geq 80\%$ . Using the IEDB analysis resource at <http://tools.iedb.org/conservancy/>, each of the epitopes predicted were used as input against all the serotypes of dengue (DENV 1–4).

**Vaccine construction.** Vaccine construction requires three requisite elements in order to create an efficient molecule capable of eliciting an immune response. These elements include the identified epitopes (the B- and T-cell epitopes), adjuvants and the spacer sequences. The predicted epitopes from the DENV-4 were selected for the final vaccine construction after the in-house screening processes. In order to amplify the immune system, molecules with immunomodulatory properties were added to the vaccine constructs as safe vaccine adjuvants. The adjuvants were connected to the epitopes using the rigid linker EAAAK. The glycine-proline rich GPGPG linkers were used to space the B-cell and HTL epitopes. In addition, the CTL epitopes were linked using an effective and flexible linker AAY. To prevent the polymorphism of HLA-DR molecules, the Pan DR T Helper Epitope (PADRE epitope) was introduced to the final vaccine sequences. The sequence order of the final vaccines are as follows: Adjuvant-Rigid linker (EAAAK)-PADRE-GGGS-BCL-flexible linker (KK)-BCL-flexible linker (GPGPG)-HTL-flexible linker (GGGS)-HTL-flexible linker (AAY)-CTL. Four vaccines were constructed with three adjuvants namely heparin-binding hemagglutinin (HBHA), beta defensin, and ribosomal protein L7/L12 for the vaccines DV-1 to 4 respectively.

**Homology modelling of the predicted epitopes.** The 3D structures of the epitopes and the designed vaccines were modeled using the structure prediction wizard of prime module in the Schrodinger suites v2020-3 using the Maestro v12.3 software interface. Since the process of constructing a 3D atomic-resolution model is based on the principle that evolutionary related proteins share similar sequence and protein structures, homology modeling therefore, relies on high quality sequence alignment and template structure (knowledge-based method of modeling). The predicted epitopes and designed sequences were used as input individually in FASTA format followed by the BLAST homology search. Homology modeling steps such as template recognition and initial alignment, alignment correction, backbone generation, loop and side chain modeling, model optimization and validation were all performed on the 3D predicted models. In addition, the online protein structure predictions; Protein Homology/analogy Recognition Engine V 2.0 (Phyre<sup>2</sup>) at <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>, RaptorX at <http://raptorx.uchicago.edu/ContactMap/>, and SCRATCH protein predictor at <http://scratch.proteomics.ics.uci.edu/> were used to remodel and ascertain the modeled 3D structures of the designed vaccines.

**Secondary structure prediction.** PSIPRED V3.3 at <http://bioinf.cs.ucl.ac.uk/psipredtest> was utilized to predict the secondary structures of the designed vaccines using their primary sequences as input. PSIPRED is a simple and accurate 2D structure prediction method, incorporating two feed-forward neural networks that performs an analysis on output obtained from PSI-BLAST (Position Specific Iterated-BLAST)<sup>106</sup>.

**Structural validation of the modeled vaccines.** Following the modified method of Saha and Prasad<sup>107</sup>, the PDBSum module in the EMBL-EBI database (<https://www.ebi.ac.uk/>) and ProSA-Web at <https://prosa.servi.ces.sbg.ac.at/prosa.php> were used to analyze the designed vaccine constructs before and after minimization and refinement in order to validate their purities<sup>108</sup>. Specifically, the PDBSum Generate at <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html> use the PDB files of the prepared constructs after homology modeling and preparation as input to evaluate purity using the PROCHECK and produces a graphical output in the form of a Ramachandran plot<sup>109</sup>. The higher the percentage of atoms in the most favored region the greater the resolution and purity and the models can be used for molecular docking calculations.

**Receptor selection.** It is crucial for designed vaccines to interact with immune cells in order to confer the stability of the immune response. The molecular docking calculation study was performed to investigate these interactions using the Major histocompatibility complex (MHCs) receptors as well as the Toll-like receptors

(TLRs). Following previously methods of receptor selections<sup>64,101,110</sup>, the crystal structures of MHC class I allele (HLA-A\*11-01) and MHC class II allele (HLA DRB1\*04-01) with PDB IDs, 5WJL and 5JLZ were retrieved from the Protein Data Bank (PDB) at <https://www.rcsb.org/>. In addition, the 3D structure of TLR4 dimer and TLR2 (PDB ID: 3FXI and PDB ID: 2Z7X) were retrieved for additional docking purposes. These proteins were refined, minimized and optimized using the protein preparation wizard in the Schrodinger suite and were further considered for molecular docking analysis with the multi-epitope constructed vaccines<sup>111</sup>.

**In silico docking calculation.** To inspect the binding affinity of the modeled vaccines to MHCs and TLRs molecules, the designed vaccines were docked against the selected receptors using the Prime protein–protein docking module in Schrodinger<sup>112</sup>. The binding score and the atomic contact energy were analyzed using the PATCHDOCK server at <https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php> while the refinement of the generated poses including the global energy were carried out using the FIREDOCK server at <http://bioinfo3d.cs.tau.ac.il/FireDock/php.php>.

**Optimization of designed vaccine candidate.** Reverse translation and codon optimization were performed using the Java Codon Adaptation Tool (JCat) server (<http://www.jcat.de/>) in order to express the multi-epitope vaccine construct in a selected expression vector<sup>113</sup>. Codon optimization was performed in order to express the final vaccine construct in the *E. coli* (strain K12) host, as codon usage of *E. coli* differs from that of ‘homo sapiens’, where the sequence of the final vaccine construct is derived. In the additional options section, the ‘avoid the rho-independent transcription termination’, ‘avoid prokaryote ribosome binding site’, and ‘avoid restriction enzymes cleavage sites’ were all checked. The JCat output includes the codon adaptation index (CAI) and GC content (%), which can be used to assess protein expression levels. CAI provides information on codon usage biases; the ideal CAI score is 1.0 but >0.8 is considered a good score<sup>114</sup>. The GC content of a sequence should range between 30–70%. GC content values outside this range suggest unfavorable effects on translational and transcriptional efficiencies<sup>53</sup>. To clone the optimized gene sequence of the final vaccine constructs in *E. coli* pET-28a (+) vector, *XhoI* and *NdeI* restriction sites were added to the N and C-terminals of the DNA sequences of the predicted vaccines, respectively. Finally, the optimized DNA sequences (with restriction sites) were inserted into the pET-28a (+) vector using SnapGene tool to ensure vaccine expression.

**Immune simulation of the designed vaccines.** In order to understand the immune system relative to the predicted vaccines, C-ImmSim at <http://150.146.2.1/C-IMMSIM/index.php?page=1> were employed to assess the molecular binding of immune complexes. The principle of operation of this server is based on the position-specific scoring matrix and machine learning approaches to predict subunit vaccine and binding interactions. Following the clinical vaccine recommended procedures<sup>115,116</sup>, and immune simulation methods previously reported<sup>89,101,117,118</sup>, three injections were administered for a total of 1000 steps of simulation using the designed vaccine sequences as input.

**Molecular Dynamic simulation (MDs).** MDs of the designed vaccines of the immunogenicity were characterized by the system build and molecular dynamic Desmond package in Schrödinger. Prior to simulation, the docked complexes were prepared using the system builder module in Maestro v12.4. The optimized potentials for the liquid simulations OPLS-2005 force field were used to determine the receptor interactions which were solvated with the simple point charged (TIP-3P) water model<sup>119,120</sup>. The orthorhombic water box was used to create a 10 Å buffer region between the atoms on the receptors and box sides. The volume of the box was minimized and the overall charge of the system was neutralized by adding Na<sup>+</sup>. The temperature and pressure were kept constant at 300° Kelvin and 1.01325 bar using the Nose–Hoover thermostat<sup>121</sup> and Martyna–Tobias–Klein barostat methods. The simulations were performed using NPT ensemble by considering the number of atoms, pressure and timescale and the simulation time at 100 ns. The MD results were analyzed by simulation interactions diagram module and MS-MD trajectory analysis.

## Conclusion

In this study, immunoinformatics approaches were employed in the development of potential vaccine candidates against DENV considering their ease of use in experimental investigations. Since the only viable defense for DENV infection is a vaccine; four potential vaccines were designed and predicted to elicit specific immune responses in individuals with dengue infection. The results of this study are based on computational approaches. Therefore, it is required that both *in-vivo* and *in-vitro* laboratory experiments be carried out to ascertain the results of this study.

## Data availability

All data generated or analysed with respect to this study are included in the submitted manuscript. The sequences of the protein analysed can be retrieved from NCBI (<https://www.ncbi.nlm.nih.gov/>), UniProt (<https://www.uniprot.org/>), and Protein Data Bank (<https://www.rcsb.org/>) using their accession numbers.

Received: 26 March 2021; Accepted: 22 September 2021

Published online: 05 October 2021

## References

1. Hotez, P. J., Aksoy, S., Brindley, P. J. & Kamhawi, S. *What Constitutes a Neglected Tropical Disease?* (Public Library of Science, 2020).

2. Fitzpatrick, C., Nwankwo, U., Lenk, E., de Vlas, S. J. & Bundy, D. A. *An Investment Case for Ending Neglected Tropical Diseases* (The World Bank, 2017).
3. Rees, C. A., Hotez, P. J., Monuteaux, M. C., Niescierenko, M. & Bourgeois, F. T. Neglected tropical diseases in children: An assessment of gaps in research prioritization. *PLoS Negl. Trop. Dis.* **13**, e0007111 (2019).
4. Adekiya, T. A., Aruleba, R. T., Klein, A. & Fadaka, A. O. In silico inhibition of SGT4 as a therapeutic target for the treatment of schistosomiasis. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1850363> (2020).
5. Hotez, P. J. Ten global “hotspots” for the neglected tropical diseases. *PLoS Negl. Trop. Dis.* **8**, e2496 (2014).
6. Relman, D. A. & Choffnes, E. R. *The Causes and Impacts of Neglected Tropical and Zoonotic Diseases: Opportunities for Integrated Intervention Strategies* (National Academies Press, 2011).
7. Waggoner, J. J. *et al.* Viremia and clinical presentation in Nicaraguan patients infected with Zika virus, chikungunya virus, and dengue virus. *Clin. Infect. Dis.* **63**, 1584–1590 (2016).
8. Organization WH, UNICEF. *Global vector control response 2017–2030* (UNICEF, 2017).
9. Organization, W. H. A *Global Brief on Vector-Borne Diseases* (World Health Organization, 2014).
10. Gurugama, P., Garg, P., Perera, J., Wijewickrama, A. & Seneviratne, S. L. Dengue viral infections. *Indian J. Dermatol.* **55**, 68 (2010).
11. Malavige, G., Fernando, S., Fernando, D. & Seneviratne, S. Dengue viral infections. *Postgrad. Med. J.* **80**, 588–601 (2004).
12. da Silveira, L. T. C., Tura, B. & Santos, M. Systematic review of dengue vaccine efficacy. *BMC Infect. Dis.* **19**, 750 (2019).
13. Crill, W. D. & Roehrig, J. T. Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. *J. Virol.* **75**, 7769–7773 (2001).
14. Lai, C. Y. *et al.* Antibodies to envelope glycoprotein of dengue virus during the natural course of infection are predominantly cross-reactive and recognize epitopes containing highly conserved residues at the fusion loop of domain II. *J. Virol.* **82**, 6631–6643 (2008).
15. Beltramello, M. *et al.* The human immune response to Dengue virus is dominated by highly cross-reactive antibodies endowed with neutralizing and enhancing activity. *Cell Host Microbe.* **8**, 271–283 (2010).
16. Oliphant, T. *et al.* Antibody recognition and neutralization determinants on domains I and II of West Nile Virus envelope protein. *J. Virol.* **80**, 12149–12159 (2006).
17. Guy, B. *et al.* A recombinant live attenuated tetravalent vaccine for the prevention of dengue. *Expert Rev. Vaccines* **16**, 671–684 (2017).
18. Dayan, G. H. *et al.* Immunogenicity and safety of a recombinant tetravalent dengue vaccine in children and adolescents ages 9–16 years in Brazil. *Am. J. Trop. Med. Hyg.* **89**, 1058 (2013).
19. Sridhar, S. *et al.* Effect of dengue serostatus on dengue vaccine safety and efficacy. *N. Engl. J. Med.* **379**, 327–340 (2018).
20. Lanata, C. F. *et al.* Immunogenicity and safety of tetravalent dengue vaccine in 2–11 year-olds previously vaccinated against yellow fever: Randomized, controlled, phase II study in Piura, Peru. *Vaccines* **30**, 5935–5941 (2012).
21. Amar-Singh, H. *et al.* Safety and immunogenicity of a tetravalent dengue vaccine in healthy children aged 2–11 years in Malaysia: A randomized, placebo-controlled, Phase III study. *Vaccine* **31**, 5814–5821 (2013).
22. Halstead, S. B. Safety issues from a Phase 3 clinical trial of a live-attenuated chimeric yellow fever tetravalent dengue vaccine. *Hum. Vaccine Immunother.* **14**, 2158–2162 (2018).
23. Halstead, S. B. Neutralization and antibody-dependent enhancement of dengue viruses. *Adv. Virus Res.* **60**, 421–467 (2003).
24. Pinheiro-Michelsen, J. R. *et al.* Anti-dengue Vaccines: From development to clinical trials. *Front. Immunol.* **11**, 1252 (2020).
25. Men, R., Bray, M., Clark, D., Chanock, R. M. & Lai, C. J. Dengue type 4 virus mutants containing deletions in the 3′ noncoding region of the RNA genome: Analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. *J. Virol.* **70**, 3930–3937 (1996).
26. Durbin, A. P. *et al.* Attenuation and immunogenicity in humans of a live dengue virus type-4 vaccine candidate with a 30 nucleotide deletion in its 3′-untranslated region. *Am. J. Trop. Med. Hyg.* **65**, 405–413 (2001).
27. Durbin, A. P. *et al.* rDEN4delta30, a live attenuated dengue virus type 4 vaccine candidate, is safe, immunogenic, and highly infectious in healthy adult volunteers. *J. Infect. Dis.* **191**, 710–718 (2005).
28. Kirkpatrick, B. D. *et al.* The live attenuated dengue vaccine TV003 elicits complete protection against dengue in a human challenge model. *Sci. Transl. Med.* **8**, 33036 (2016).
29. Osorio, J. E. *et al.* Safety and immunogenicity of a recombinant live attenuated tetravalent dengue vaccine (DENVax) in flavivirus-naïve healthy adults in Colombia: A randomised, placebo-controlled, phase 1 study. *Lancet. Infect. Dis.* **14**, 830–838 (2014).
30. Sáez-Llorens, X. *et al.* Immunogenicity and safety of one versus two doses of tetravalent dengue vaccine in healthy children aged 2–17 years in Asia and Latin America: 18-month interim data from a phase 2, randomised, placebo-controlled study. *Lancet. Infect. Dis.* **18**, 162–170 (2018).
31. Sáez-Llorens, X. *et al.* Safety and immunogenicity of one versus two doses of Takeda’s tetravalent dengue vaccine in children in Asia and Latin America: Interim results from a phase 2, randomised, placebo-controlled study. *Lancet Infect. Dis.* **17**, 615–625 (2017).
32. Jackson, L. A. *et al.* A phase 1 study of safety and immunogenicity following intradermal administration of a tetravalent dengue vaccine candidate. *Vaccine* **36**, 3976–3983 (2018).
33. Manoff, S. B. *et al.* Preclinical and clinical development of a dengue recombinant subunit vaccine. *Vaccine* **33**, 7126–7134 (2015).
34. Clements, D. E. *et al.* Development of a recombinant tetravalent dengue virus vaccine: Immunogenicity and efficacy studies in mice and monkeys. *Vaccine* **28**, 2705–2715 (2010).
35. Manoff, S. B. *et al.* Immunogenicity and safety of an investigational tetravalent recombinant subunit vaccine for dengue: Results of a Phase I randomized clinical trial in flavivirus-naïve adults. *Hum. Vaccines Immunother.* **15**, 2195–2204 (2019).
36. Danko, J. R. *et al.* Safety and immunogenicity of a tetravalent dengue DNA vaccine administered with a cationic lipid-based adjuvant in a phase 1 clinical trial. *Am. J. Trop. Med. Hyg.* **98**, 849 (2018).
37. Schmidt, A. C. *et al.* Phase 1 randomized study of a tetravalent dengue purified inactivated vaccine in healthy adults in the United States. *Am. J. Trop. Med. Hyg.* **96**, 1325 (2017).
38. Thomas, S. J. *et al.* A phase II, randomized, safety and immunogenicity study of a re-derived, live-attenuated dengue virus vaccine in healthy adults. *Am. J. Trop. Med. Hyg.* **88**, 73 (2013).
39. Porter, K. R. *et al.* Immunogenicity and protective efficacy of a vaxfectin-adjuvanted tetravalent dengue DNA vaccine. *Vaccine* **30**, 336–341 (2012).
40. Osatomi, K. & Sumiyoshi, H. Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology* **176**, 643–647 (1990).
41. Wang, B. *et al.* Phylogenetic analysis of dengue virus reveals the high relatedness between imported and local strains during the 2013 dengue outbreak in Yunnan, China: A retrospective analysis. *BMC Infect. Dis.* **15**, 1–7 (2015).
42. Bäck, A. & Lundkvist, A. Dengue viruses: An overview. *Infect. Ecol. Epidemiol.* **3**, 19839 (2013).
43. Cao, J. *et al.* Epidemiological and clinical characteristics of Dengue virus outbreaks in two regions of China, 2014–2015. *PLoS ONE* **14**, e0213353 (2019).
44. Fajardo-Sánchez, E., Galiano, V. & Villalain, J. Molecular dynamics study of the membrane interaction of a membranotropic dengue virus C protein-derived peptide. *J. Biomol. Struct. Dyn.* **35**, 1283–1294 (2017).
45. Tuiskunen Bäck, A. & Lundkvist, A. Dengue viruses: An overview. *Infect. Ecol. Epidemiol.* **3**, 19839 (2013).

46. Naz, A. *et al.* Identification of putative vaccine candidates against *Helicobacter pylori* exploiting exoproteome and secretome: A reverse vaccinology based approach. *Infect. Genet. Evol.* **32**, 280–291 (2015).
47. Ka, T. *et al.* A Candidate multi-epitope vaccine against SARS-CoV-2. *Sci. Rep.* **10**, 10895 (2020).
48. Ullah, A., Sarkar, B. & Islam, S. S. Exploiting the reverse vaccinology approach to design novel subunit vaccine against ebola virus. *Immunobiology* **8**, 151949 (2020).
49. Alam, A., Ali, S., Ahamad, S., Malik, M. Z. & Ishrat, R. From ZikV genome to vaccine: In silico approach for the epitope-based peptide vaccine against Zika virus envelope glycoprotein. *Immunology* **149**, 386–399 (2016).
50. Anwar, S., Mouroosi, J. T., Khan, M. F. & Hosen, M. J. Prediction of epitope-based peptide vaccine against the chikungunya virus by immuno-informatics approach. *Curr. Pharm. Biotechnol.* **21**, 325–340 (2020).
51. Chong, L. C. & Khan, A. M. *Vaccine Target Discovery. Encyclopedia of Bioinformatics and Computational Biology* (Elsevier, 2019).
52. Maria, R., Arturo, C., Alicia, J. A., Paulina, M. & Gerardo, A. O. *The Impact of Bioinformatics on Vaccine Design and Development* (InTech, 2017).
53. Ali, M. *et al.* Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci. Rep.* **7**, 1–13 (2017).
54. Chaplin, D. D. Overview of the immune response. *J. Allergy Clin. Immunol.* **125**, S3–S23 (2010).
55. Fadaka, A. O. *et al.* Inhibitory potential of repurposed drugs against the SARS-CoV-2 main protease: A computational-aided approach. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1847197> (2020).
56. Fadaka, A. O., Pretorius, A. & Klein, A. MicroRNA assisted gene regulation in colorectal cancer. *Int. J. Mol. Sci.* **20**, 4899 (2019).
57. Abedi Karjiban, R. *et al.* Molecular dynamics study of the structure, flexibility and dynamics of thermostable L1 lipase at high temperatures. *Protein. J.* **28**, 14–23 (2009).
58. Fadaka, A. O., Sibuyi, N. R. S., Madihe, A. M. & Meyer, M. Computational insight of dexamethasone against potential targets of SARS-CoV-2. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1819880> (2020).
59. Jyotisha, S. S. & Qureshi, I. A. Multi-epitope vaccine against SARS-CoV-2 applying immunoinformatics and molecular dynamics simulation approaches. *J. Biomol. Struct. Dyn.* **1**, 17. <https://doi.org/10.1080/07391102.2020.1844060> (2020).
60. Sarkar, B., Ullah, M. A., Johora, F. T., Taniya, M. A. & Araf, Y. Immunoinformatics-guided designing of epitope-based subunit vaccines against the SARS Coronavirus-2 (SARS-CoV-2). *Immunobiology* **225**, 151955 (2020).
61. Ghaebi, M., Osali, A., Valizadeh, H., Roshangar, L. & Ahmadi, M. Vaccine development and therapeutic design for 2019-nCoV/SARS-CoV-2: Challenges and chances. *J. Cell. Physiol.* **235**, 9098–9109 (2020).
62. Dong, R., Chu, Z., Yu, F. & Zha, Y. Contriving Multi-Epitope Subunit of Vaccine for COVID-19: Immunoinformatics Approaches. *Front. Immunol.* **11**, 1784 (2020).
63. Lim, H. X., Lim, J., Jazayeri, S. D., Poppema, S. & Poh, C. L. Development of multi-epitope peptide-based vaccines against SARS-CoV-2. *Biomed. J.* **44**, 18–30 (2020).
64. Rahman, N. *et al.* Vaccine design from the ensemble of surface glycoprotein epitopes of SARS-CoV-2: An immunoinformatics approach. *Vaccines*. **8**, 423 (2020).
65. Kar, P. P. & Srivastava, A. Immuno-informatics analysis to identify novel vaccine candidates and design of a multi-epitope based vaccine candidate against theileria parasites. *Front. Immunol.* **9**, 2213 (2018).
66. Zawawi, A. *et al.* In silico design of a T-cell epitope vaccine candidate for parasitic helminth infection. *PLoS Pathog.* **16**, e1008243 (2020).
67. Tripathi, N. K. & Shrivastava, A. Recent developments in recombinant protein-based dengue vaccines. *Front. Immunol.* **2018**, 1919 (2018).
68. Thomas, S. J. & Rothman, A. L. Trials and tribulations on the path to developing a dengue vaccine. *Vaccine* **33**, D24–D31 (2015).
69. Halstead, S. B. Antibody, macrophages, dengue virus infection, shock, and hemorrhage: A pathogenetic cascade. *Rev. Infect. Dis.* **11**, S830–S839 (1989).
70. Thomas, S. J. Preventing dengue: Is the possibility now a reality?. *N. Engl. J. Med.* **372**, 172–173 (2015).
71. Naz, K. *et al.* PanRV: Pangenome-reverse vaccinology approach for identifications of potential vaccine candidates in microbial pangenome. *BMC Bioinform.* **20**, 123 (2019).
72. Kumar Jaiswal, A. *et al.* An in silico identification of common putative vaccine candidates against *Treponema pallidum*: A reverse vaccinology and subtractive genomics based approach. *Int. J. Mol. Sci.* **18**, 402 (2017).
73. Kar, T. *et al.* A candidate multi-epitope vaccine against SARS-CoV-2. *Sci. Rep.* **10**, 10895 (2020).
74. Sauer, K. & Harris, T. An effective COVID-19 vaccine needs to engage T cells. *Front. Immunol.* **11**, 581807 (2020).
75. Yadav, S. *et al.* In silico and in vitro studies on the protein-protein interactions between *Brugia malayi* immunomodulatory protein calreticulin and human C1q. *PLoS ONE* **9**, e106413 (2014).
76. Aathmanathan, V. S., Jothi, N., Prajapati, V. K. & Krishnan, M. Investigation of immunogenic properties of Hemolin from silkworm, *Bombyx mori* as carrier protein: an immunoinformatic approach. *Sci. Rep.* **8**, 1–10 (2018).
77. Droppa-Almeida, D., Franceschi, E. & Padilha, F. F. Immune-informatic analysis and design of peptide vaccine from multi-epitopes against *Corynebacterium pseudotuberculosis*. *Bioinform. Biol. Insights* **12**, 1177932218755337 (2018).
78. Rezik, I. *et al.* In silico characterization and Molecular modeling of double-strand break repair protein MRE11 from *Phoenix dactylifera* v deglet nour. *Theor. Biol. Med. Model.* **12**, 1–14 (2015).
79. Hashemzadeh, P., Ghorbanzadeh, V., Lashgarian, H. E., Kheirandish, F. & Dariushnejad, H. Harnessing bioinformatic approaches to design novel multi-epitope subunit vaccine against *Leishmania infantum*. *Int. J. Pept. Res. Ther.* **26**, 1417–1428 (2020).
80. Vijay, K. Toll-like receptors in immunity and inflammatory diseases: Past, present, and future. *Int. Immunopharmacol.* **59**, 391–412 (2018).
81. Ahmad, S. *et al.* Design of a novel multi epitope-based vaccine for pandemic coronavirus disease (COVID-19) by vaccinomics and probable prevention strategy against avenging zoonotics. *Eur. J. Pharm. Sci.* **151**, 105387 (2020).
82. Pfarr, K. M., Fischer, K. & Hoerauf, A. Involvement of Toll-like receptor 4 in the embryogenesis of the rodent filaria *Litomosoides sigmodontis*. *Med. Microbiol. Immunol.* **192**, 53–56 (2003).
83. Kerepesi, L. A., Leon, O., Lustigman, S. & Abraham, D. Protective immunity to the larval stages of *Onchocerca volvulus* is dependent on Toll-like receptor 4. *Infect. Immun.* **73**, 8291–8297 (2005).
84. Compton, T. *et al.* Human cytomegalovirus activates inflammatory cytokine responses via CD14 and Toll-like receptor 2. *J. Virol.* **77**, 4588–4596 (2003).
85. Marin, A. *et al.* Relationship between G+ C content, ORF-length and mRNA concentration in *Saccharomyces cerevisiae*. *Yeast* **20**, 703–711 (2003).
86. Rapin, N., Lund, O., Bernaschi, M. & Castiglione, F. Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system. *PLoS ONE* **5**, e9862 (2010).
87. Devi, A. & Chaitanya, N. S. In silico designing of multi-epitope vaccine construct against human coronavirus infections. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1804460> (2020).
88. Ismail, S., Ahmad, S. & Azam, S. S. Vaccinomics to design a novel single chimeric subunit vaccine for broad-spectrum immunological applications targeting nosocomial Enterobacteriaceae pathogens. *Eur. J. Pharm. Sci.* **146**, 105258 (2020).
89. Tahir ul Qamar, M. *et al.* Multi-epitope-based subunit vaccine design and evaluation against respiratory syncytial virus using reverse vaccinology approach. *Vaccines* **8**, 288 (2020).

90. Khan, A. *et al.* Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PLoS ONE* **13**, e0196484 (2018).
91. Banisharif-Dehkordi, F., Mobini-Dehkordi, M., Shakhshi-Niaei, M. & Mahnam, K. Design and molecular dynamic simulation of a new double-epitope tolerogenic protein as a potential vaccine for multiple sclerosis disease. *Res. Pharm. Sci.* **14**, 20–26 (2019).
92. Nielsen, H. *Predicting Secretory Proteins with SignalP* 59–73 (Springer, 2017).
93. Petersen, T. N., Brunak, S., Von Heijne, G. & Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
94. Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H. & Winther, O. DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics* **33**, 3387–3395 (2017).
95. Doytchinova, I. A. & Flower, D. R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **8**, 4 (2007).
96. Westerhout, J. *et al.* Allergenicity prediction of novel and modified proteins: Not a mission impossible! Development of a random forest allergenicity prediction model. *Regul. Toxicol. Pharmacol.* **107**, 104422 (2019).
97. Dimitrov, I., Flower, D. R. & Doytchinova, I. AllerTOP: A server for in silico prediction of allergens. *BMC Bioinform.* **14**, 1–9 (2013).
98. Venkatarajan, M. S. & Braun, W. New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical–chemical properties. *Mol. Model. Annu.* **7**, 445–453 (2001).
99. Faria, A. R. *et al.* High-throughput analysis of synthetic peptides for the immunodiagnosis of canine visceral leishmaniasis. *PLoS Negl. Trop. Dis.* **5**, e1310 (2011).
100. Beaver, J. E., Bourne, P. E. & Ponomarenko, J. V. EpitopeViewer: a Java application for the visualization and analysis of immune epitopes in the Immune Epitope Database and Analysis Resource (IEDB). *Immunome Res.* **3**, 3 (2007).
101. Shey, R. A. *et al.* In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci. Rep.* **9**, 1–18 (2019).
102. Möller, S., Croning, M. D. & Apweiler, R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**, 646–653 (2001).
103. Walker, J. M. *The Proteomics Protocols Handbook* (Springer, 2005).
104. Gasteiger, E. *et al.* *Protein Identification and Analysis Tools on the ExpASY Server* 571–607 (Springer, 2005).
105. Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R. & Warwicker, J. Protein-Sol: A web tool for predicting protein solubility from sequence. *Bioinformatics* **33**, 3098–3100 (2017).
106. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349–W357 (2013).
107. Saha, R. & Prasad, B. V. In silico approach for designing of a multi-epitope based vaccine against novel Coronavirus (SARS-COV-2). *bioRxiv* <https://doi.org/10.1101/2020.03.31.017459> (2020).
108. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410 (2007).
109. Lovell, S. C. *et al.* Structure validation by C $\alpha$  geometry:  $\phi$ ,  $\psi$  and C $\beta$  deviation. *Proteins* **50**, 437–450 (2003).
110. Ullah, M. A., Sarkar, B. & Islam, S. S. Exploiting the reverse vaccinology approach to design novel subunit vaccines against Ebola virus. *Immunobiology* **225**, 151949 (2020).
111. Fadaka, A. O., Sibuyi, N. R. S., Madihe, A. M. & Meyer, M. MicroRNA-based regulation of Aurora a kinase in breast cancer. *Oncotarget* **11**, 4306 (2020).
112. Ojo, O. A. *et al.* Deciphering the interaction of puerarin with cancer macromolecules: An in silico investigation. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1819425> (2020).
113. Grote, A. *et al.* JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* **33**, W526–W531 (2005).
114. Morla, S., Makhija, A. & Kumar, S. Synonymous codon usage pattern in glycoprotein gene of rabies virus. *Gene* **584**, 1–6 (2016).
115. Castiglione, F., Mantile, F., De Berardinis, P. & Prisco, A. How the interval between prime and boost injection affects the immune response in a computational model of the immune system. *Comput. Math. Methods Med.* **2012**, 1–9 (2012).
116. Kroger, A. *General Recommendations on Immunization; US Department of Health and Human Services* (Public Health Service, Centers for Disease Control, 2013).
117. Nain, Z., Karim, M. M., Sen, M. K. & Adhikari, U. K. Structural basis and designing of peptide vaccine using PE-PGRS family protein of *Mycobacterium ulcerans*: An integrated vaccinomics approach. *Mol. Immunol.* **120**, 146–163 (2020).
118. Chauhan, V. & Singh, M. P. Immuno-informatics approach to design a multi-epitope vaccine to combat cytomegalovirus infection. *Eur. J. Pharm. Sci.* **147**, 105279 (2020).
119. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
120. Blessy, J. J. & Sharmila, D. J. S. Molecular simulation of N-acetylneuraminic acid analogs and molecular dynamics studies of cholera toxin-Neu5Gc complex. *J. Biomol. Struct. Dyn.* **33**, 1126–1139 (2015).
121. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695 (1985).

## Acknowledgements

We would like to acknowledge the National Integrated Cyberinfrastructure system, Center for High Performance Computing (CHPC), Department of Science and Technology, Republic of South Africa for the license to the Lengau cluster and other modules under the Schrodinger suit.

## Author contributions

All authors have made significant contributions to the submission of this manuscript. A.O.F., N.R.S.S., A.K., A.M., and M.M conceived the concept of the work. A.O.F., A.K., A.M., M.G., and M.M. designed the study. A.O.F., O.O.B., A.K., A.M., D.R.M., and M.M. carried out the research, data analyses and supervision. Manuscript draft by A.O.F., D.R.M., M.G., N.R.S.S., O.O.B., M.M., A.K., and N.R.S.S. critically appraised the manuscript. All authors agreed with the submission of the final version of the article.

## Funding

Department of Science and Innovation/Mintek Nanotechnology Innovation Centre, Biolabels Node.

## Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to A.O.F. or M.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021