



## RESEARCH ARTICLE

# Discovery of potential immune epitopes and peptide vaccine design - a prophylactic strategy against Rift Valley fever virus [version 1; peer review: 2 approved with reservations]

Maruf Ahmed Bhuiyan <sup>1</sup>, Syeda Tasnim Quayum <sup>2</sup>, Foysal Ahammad <sup>3,4</sup>, Rahat Alam<sup>4,5</sup>, Abdus Samad <sup>4,5</sup>, Zulkar Nain <sup>4,6,7</sup>

<sup>1</sup>Department of Virology, Bangabandhu Sheikh Mujib Medical University, Dhaka, 1000, Bangladesh

<sup>2</sup>Department Biochemistry and Microbiology, North South University, Bashundhara, Dhaka, 1229, Bangladesh

<sup>3</sup>Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

<sup>4</sup>Biological Solution Centre (BioSol Centre), Farmgate, Dhaka, 1215, Bangladesh

<sup>5</sup>Department of Genetic Engineering and Biotechnology, Faculty of Biological Science and Technology, Jashore University of Science and Technology, Jashore, 7408, Bangladesh

<sup>6</sup>Department of Genetic Engineering and Biotechnology, Faculty of Sciences and Engineering, East West University, Dhaka, 1212, Bangladesh

<sup>7</sup>Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Islamic University, Kushtia, 7003, Bangladesh

**V1** First published: 18 Aug 2020, 9:999  
<https://doi.org/10.12688/f1000research.24975.1>  
Latest published: 18 Aug 2020, 9:999  
<https://doi.org/10.12688/f1000research.24975.1>

## Abstract

**Background:** Rift Valley fever virus (RVFV) is an emerging arbovirus infecting both animals and humans. Any form of direct contact with body fluids, blood or tissue of infected animals is the mode of transmission of this pathogen. Despite being an emerging virus, no proper vaccinations are yet available for the public. Our objective is to compose a multiepitope vaccine utilizing immuno-bioinformatics as a strategy against RVFV.

**Methods:** To identify immunodominant epitopes and design a potent vaccine candidate, we applied a series of immunoinformatic approaches with molecular dynamics and immune response simulation frameworks.

**Results:** A glycoprotein with the highest antigenicity was selected and employed for determining promising epitopes. We selected T cell epitopes based on their immunological potencies and cytokine inducing properties, while B cell epitopes were selected based on their antigenic features. Finally, we selected four cytotoxic T-lymphocyte, two helper T-lymphocyte, and three linear B-lymphocyte epitopes that were arranged into a vaccine construct with appropriate adjuvants and linkers. The chimera protein was modeled, refined, and validated prior to docking against toll-like receptor 4. Docking studies suggest strong binding interactions while dynamics simulation revealed the stable nature of the docked complex. Furthermore, the immune simulation showed robust and prolonged immune responses with

## Open Peer Review

### Reviewer Status

#### Invited Reviewers

1 ..... 2 .....

version 1  
18 Aug 2020

report

report

1. **Firzan Nainu** , Hasanuddin University, Makassar, Indonesia
2. **Trina E. Tallei** , Universitas Sam Ratulangi, Manado, Indonesia
3. **Talha Bin Emran** , BGC Trust University, Bangladesh, Chittagong, Bangladesh
4. **Sandra Moreno**, National Institute for Agriculture and Food Research and Technology (INIA), Madrid, Spain
5. **Alejandro Brun** , Animal Health Research Centre (CISA), National Institute for

rapid antigen clearance. Finally, codon optimization and cloning conducted with *Escherichia coli* K12 suggests high translation efficiency within the host system.

**Conclusion:** We believe that our designed multiepitope vaccine is a promising prophylactic candidate against RVFV pathogenesis.

#### Keywords

Vaccine Design, Immunoinformatics, Rift Valley fever virus, Immune Simulation

Agriculture and Food Research and Technology (INIA), Madrid, Spain

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding authors:** Abdus Samad ([kazisamad50@gmail.com](mailto:kazisamad50@gmail.com)), Zulkar Nain ([znain.bd@gmail.com](mailto:znain.bd@gmail.com))

**Author roles:** **Bhuiyan MA:** Data Curation, Formal Analysis, Investigation, Visualization, Writing – Original Draft Preparation; **Quayum ST:** Formal Analysis, Investigation, Visualization, Writing – Original Draft Preparation; **Ahammad F:** Resources, Writing – Review & Editing; **Alam R:** Formal Analysis, Investigation, Visualization; **Samad A:** Methodology, Resources, Supervision, Writing – Review & Editing; **Nain Z:** Conceptualization, Methodology, Supervision, Validation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** The author(s) declared that no grants were involved in supporting this work.

**Copyright:** © 2020 Bhuiyan MA *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Bhuiyan MA, Quayum ST, Ahammad F *et al.* **Discovery of potential immune epitopes and peptide vaccine design - a prophylactic strategy against Rift Valley fever virus [version 1; peer review: 2 approved with reservations]** F1000Research 2020, 9:999 <https://doi.org/10.12688/f1000research.24975.1>

**First published:** 18 Aug 2020, 9:999 <https://doi.org/10.12688/f1000research.24975.1>

## Abbreviations

2D, two-dimensional; 3D, three-dimensional; ACC, auto cross co-variance; CAI, codon adaptation index; CTL, cytotoxic T-lymphocyte; *E. coli*, *Escherichia coli*; GC, guanine and cytosine; GRAVY, grand average of hydropathicity; HTL, helper T-lymphocyte; IEDB, immune epitope database; IFN, interferon; IFN- $\gamma$ , interferon-gamma; IL, interleukin; IL-10, interleukin 10; IL-4, interleukin 4; kNN, k-nearest neighbors; LBL, linear B-lymphocyte; MHC, major histocompatibility complex; NCBI, National Center for Biotechnology Information; NMA, normal mode analysis; PDB, Protein Data Bank; PI, isoelectric point; RMSD, root mean square deviation; RVF, Rift Valley fever; RVFV, Rift Valley fever virus; SVM, support vector machine; TAP, transporter associated with antigen processing; TLR4, toll-like receptor 4.

## Introduction

Rift Valley fever virus (RVFV) is a risk to worldwide public health and farming, especially in parts of Africa, Madagascar, and the Middle East<sup>1</sup>. RVFV epidemics have killed hundreds of thousands of animals, more than a thousand humans, and caused significant economic losses<sup>2</sup>. RVFV is a negative sense, single-stranded RNA (ssRNA) virus<sup>3</sup> and belongs to the family of *Bunyaviridae*. RVFV was first identified in Kenya among sheep, in the vicinity of Lake Naivasha<sup>4</sup>. Using mosquitoes as a vector it can cause large scale transmission, causing mild symptoms like fever, back pain and nausea to fatal illnesses including critical eye diseases, encephalitis in humans, and lethal hemorrhage in animals<sup>5</sup>. Mortality of up to 90% has been reported in newborn animals and up to 30% in adult animals<sup>6</sup>, although the mortality rate for humans has been reported to be approximately 2%<sup>7</sup>. Epidemic alarms have persuaded several national and international health organizations to issue cautions about the rising risk of infection in Rift Valley fever (RVF) uninfected countries, like Europe and USA, due to the existence of vectors of transmission which are highly permissive, further compounded by global animal trade<sup>7-9</sup>. These reports unanimously concluded that coordinated efforts are needed in order to prepare for preventive measures against the recurrent emergence of RVFV.

RVF is an arthropod-borne zoonotic infectious viral disease caused by RVFV<sup>9</sup>. Direct contact via tissue, body fluids or blood of RVFV infected animals acts as the prime mode of transmission for humans<sup>10</sup>. Mosquitoes are a major vector for RVFV spread<sup>11</sup>, infecting humans over long distances and even causing vertical transmission between livestock<sup>8</sup>. The incubation period for RVFV is about 2–6 days in humans. The virus consists of a negative sense, triple segmented (large, medium, and small) ssRNA molecule and has a viral genome encoding four proteins: glycoprotein, RNA-dependent RNA polymerase, non-structural protein, and nucleocapsid protein<sup>12-14</sup>. Although the non-structural proteins facilitate RVF to survive inside its host by inhibiting first-line immunogenic responses, its glycoproteins are essential and highly crucial for invasion, entry and viral replication inside the host cell<sup>13,14</sup>. Thus, the viral glycoproteins were targeted for our multiepitope vaccine design, which would be constructed using glycoprotein epitope

sequences evoking an immune response inside the human system<sup>13,14</sup>. At present, epitope-based candidate vaccine design against viruses and bacteria as well as parasites has become very popular and has been done previously<sup>15-20</sup>. Multiepitope vaccines consist of short peptide fragments of immunogenic stimulants, which trigger a strong immune response and allow for a significantly lower chance of allergenic reactions inside the host system<sup>21</sup>. The identification of immunogenic epitopes derived from viral glycoprotein or nucleocapsid sequences has significantly enhanced the *in silico* development of peptide vaccines<sup>22</sup>.

In our study, we screened the RVFV proteome to find the highest antigenic glycoprotein to predict T and B cell epitopes using a computational approach. Subsequently, cytotoxic T-lymphocyte (CTL), helper T-lymphocyte (HTL) and linear B-lymphocyte (LBL) epitopes predicted to be immunogenic and antigenic were shortlisted, which were further subjected to toxicity and allergenicity analysis. A vaccine design was assembled by combining all the assessed CTL, HTL and LBL epitopes using linkers and an appropriate adjuvant. Physico-chemical analysis and solubility prediction were performed in *Escherichia coli* (*E. coli*) to assess the vaccine. Next, secondary and tertiary vaccine models were predicted using structure analysis tools. The predicted tertiary structure was refined and validated. Moreover, a disulfide bond was introduced by disulfide engineering, improving vaccine stability. Interactions within the vaccine-TLR4 complex were studied using molecular docking and evaluated using molecular dynamics simulation. Additionally, codon optimization was carried out to increase the translation efficiency of the designed vaccine within a *E. coli* K12 host. Finally, immune simulation was carried out to predict real-life immunogenic potency of the vaccine. The employed steps for the development of vaccine are summarized in Figure 1.

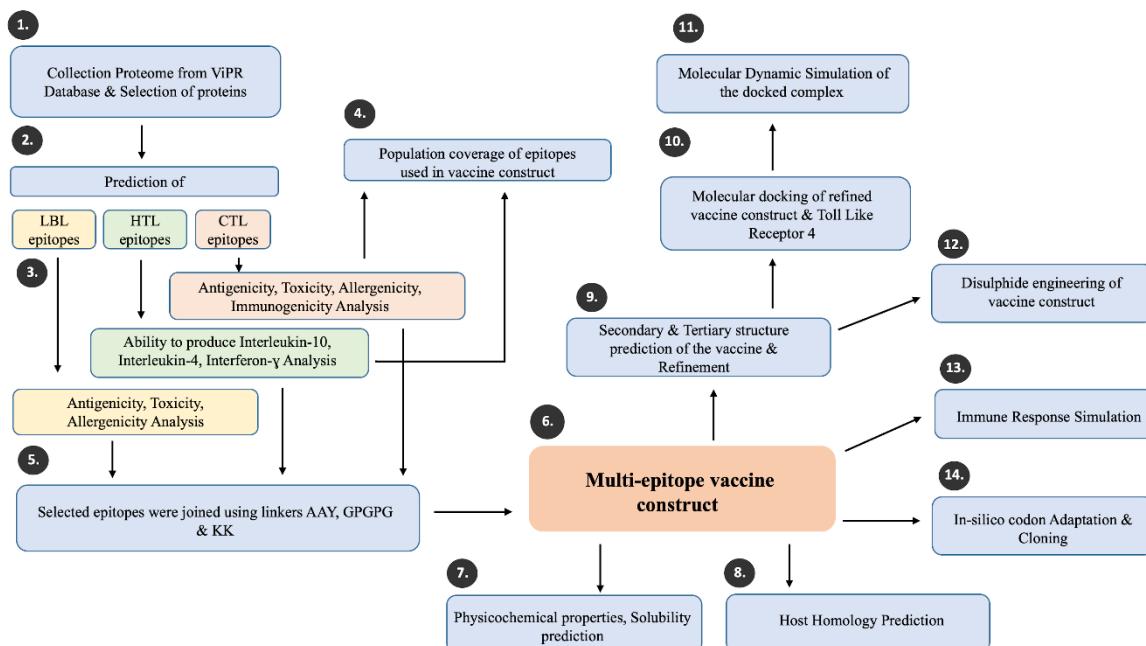
## Methods

### Proteome collection and antigenicity evaluation

The 232 complete proteomes of RVFV were obtained from **ViPR** (Virus Pathogen Resource) database (Supplementary Table 1, *Extended data*)<sup>23</sup>. ViPR is a reliable and open database for pathogenic virus families<sup>24</sup>. The viral proteome sequences retrieved were evaluated using **VaxiJen v2.0 server** for scores indicating antigenic influence. The threshold for antigenicity was fixed at 0.5<sup>25</sup>. VaxiJen v2.0 sever possesses high prediction powers and utilizes auto cross-covariance (ACC) transformation methods. The glycoprotein sequence with the highest antigenicity score was taken from this proteome for further analysis.

### CTL epitopes and their MHC class I alleles predictions

the **NetCTL v1.2 tool** was utilized to isolate CTL peptides from the RVF glycoprotein sequence. This server generates different nonamer epitopes against 12 supertypes (A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, B62). NetCTL v1.2 identifies CTL epitopes depending on C-terminal cleavage, TAP competence and MHC-1 complex binding<sup>26</sup>. The threshold was determined at 0.5 with a corresponding sensitivity of 0.89 and



**Figure 1.** Flowchart representing the overall vaccine design process from designing to validation and simulation *in silico*. LBL, linear B-lymphocyte; HTL, helper T-lymphocyte; CTL, cytotoxic T-lymphocyte.

specificity of 0.94. The MHC-I affinity for the CTL peptides was identified using the [MHC-I IEDB web server](#)<sup>27</sup> and a consensus percentile of  $\leq 5$  was set to narrow down CTL epitopes, for obtaining high-affinity epitopes with MHC-1 alleles.

#### Assessment of CTL epitopes

The isolated CTL epitopes were assessed for antigenicity by submitting them to the VaxiJen v2.0 server. To ensure proper induction of immune response in the human body, the immunogenicity of these epitopes was evaluated using the [IEDB Immunogenicity tool](#)<sup>28</sup>. Furthermore, [AllerTOP v2.0 server](#) was chosen for predicting allergenicity and ensure that the vaccine construct does not induce an allergic reaction in humans. AllerTOP v2.0 uses k-nearest neighbors (kNN) methods to distinguish allergens from non-allergens<sup>29</sup>. Lastly, the selected nonamers were screened for probable toxicity using [ToxinPred server](#), which predicts toxic epitopes using quantitative matrix methods and machine learning technology<sup>30</sup>.

#### Identification of HTL epitopes and binding MHC class II alleles

HTL acts as the orchestrator to stimulate B cells, macrophages, and CD8+ cells to fight against pathogens. Consequently, HTL epitopes are important for making an effective vaccine<sup>31</sup>. HTL epitopes, each 15-mer in length, were identified by utilizing the [IEDB MHC-II tool](#). The anticipation of binding of the epitopes to class II alleles, HLA-DP, HLA-DR and HLA-DQ, were determined utilizing the consensus 2.22 prediction method on the same server<sup>32</sup>. Due to obtaining a large number of epitopes, a percentile rank of  $\leq 0.3$  was set as a threshold.

#### Selection of cytokine-inducing HTL epitopes

HTLs release cytokines such as interferon-gamma (IFN- $\gamma$ ), interleukin 4 (IL-4), and interleukin 10 (IL-10) that can activate immune cells in the body. In addition, cytokines released by HTLs can survive past inflammatory responses and avert tissue damage. Therefore, HTL epitopes that can induce the release of cytokines are important in vaccine development. Therefore, in order to incorporate the epitopes that induce IFN-  $\gamma$ , we used the [IFNepitope server](#), through the hybrid method (motif and SVM)<sup>33</sup>. Furthermore, the peptides were evaluated to check if they induce IL-4 and IL-10 using [IL4pred](#) and [IL10pred](#) servers, respectively<sup>34,35</sup>.

#### Evaluation of LBL epitopes

The activation of B cells plays a vital part in the activation of the humoral immune response and generation of plasma cells against a specific antigen, and distinguishing LBL epitopes is another vital step in the development of epitope-based vaccine constructs. LBL epitopes were predicted using a combinatorial algorithm of gradient boosting and a randomized tree method using the [iBCE-EL server](#)<sup>36</sup>. The predicted LBL epitopes were then reevaluated for their antigenicity score with VaxiJen v2.0 server, and non-allergenic and non-toxic epitopes were predicted using the AllerTOP v2.0 server and ToxinPred tool, respectively.

#### Estimation of population coverage

For an epitope to be antigenic and evoke a strong immune reaction it needs to be acknowledged by the MHC complex molecule. MHC alleles are the most polymorphic and occur

in thousands of HLA combinations in humans. Therefore, an HLA allele with a high frequency of occurrences in the majority population of the world would have a high chance of exerting an effect immunogenic<sup>37</sup>. In our study, we wanted to find out the distribution, presence and frequency of the T cell epitopes, which were selected for the purpose of designing the vaccine structure. Allele Population Coverage of [IEDB population coverage tool](#) was utilized for the calculation of population coverage<sup>38</sup>. A prediction analysis was run on regions of Africa where the RVFV outbreak initially started and the neighboring countries that were mostly affected and also across the entire world.

### Assembling of multiepitope vaccine

All the predicted and assessed epitopes i.e. CTL, LBL, and HTL peptides were joined together using linkers and an adjuvant sequence was added upstream of all of them to form a vaccine structure. 50S ribosomal protein L7/L12 (GenPept Accession: [P9WHE3](#)) is a toll-like receptor 4 (TLR4) agonist and was chosen as an adjuvant for the construct to boost the immune response against it<sup>21</sup>. TLR4 can recognize viral glycoproteins and bacterial ligands<sup>39,40</sup> stimulates the production of cytokines against them<sup>41</sup>. The CTL peptides and the 50S L7/L12 protein adjuvant were linked together with the EAAAK linker to ensure adequate separation between each component for their effective interaction with their respective targets. This linker was chosen as it separates bifunctional fused protein domains<sup>42</sup>. Each CTL epitope was separated using AAY linkers, and HTL epitopes by GPGPG linkers. Independent immunogenic potential of each LBL epitope was preserved by separating them using KK linkers<sup>43</sup>.

### Host homology crosschecking

To avoid the risk of elicitation of autoimmunity by molecular mimicry, we screened the vaccine construct against the *Homo sapiens* protein sequences (NCBI: txid9606) through the use of NCBI BLASTp<sup>44</sup>.

### Physicochemical evaluation of the vaccine construct

Vaccines are developed to provide protection against diseases by evoking an immune response after injection against specific antigens, without causing any disease. Thus, vaccines need to have antigen inducing capability without being allergenic to the receiver. Therefore, the multiepitope vaccine construct was tested for antigenicity and allergenicity in each step of the initial design, using VaxiJen v2.0<sup>25</sup> and AllerTOP v2.0 servers, respectively<sup>29</sup>. Predictive assessment of different physiochemical attributes of the subunit vaccine was carried out using the [ProtParam server](#)<sup>45</sup>. Lastly, the [SOLpro server](#) in the SCRATCH suite was utilized to assess the solubility of the vaccine structure in *E. coli*, with a view to determining the bioavailability of the vaccine<sup>46</sup>.

### Secondary and tertiary modelling of the construct

Predicted two dimensional (2D) configuration of our vaccine design was generated using the [PSIPRED v4.0 web tool](#). The PSIPRED tool utilizes the query amino acid residues to predict a 2D model using two-feed-forward neural networking along

with PSI-BLAST<sup>47</sup>. The three dimensional (3D) model of our designed vaccine was attained through the [I-TASSER web tool](#), which exhibits protein modeling via a hierarchical procedure to estimate and bring about suitable structure and function<sup>48</sup>. The I-TASSER site helps to generate 3D structure of a protein and determine its functions using a state-of-the-art algorithm with high precision. This web tool enables determination of C-score, TM-score value and root mean square deviation (RMSD), along with the top five predicted structure models of the given protein sequence<sup>48</sup>. The produced 3D structure was chosen based on its C-score value and downloaded in PDB format. The server provides a C-score ranging between -5 to 2, where a higher value indicates a better model<sup>48</sup>.

### Refinement and verification of 3D vaccine

Our 3D vaccine structure was refined using the [GalaxyRefine server](#). The web-based program reconstructs sidechains, then repacks them using dynamic simulations for proper structural relaxation<sup>49</sup>. The [ProSA tool](#) was used to recognize possible errors in the predicted tertiary model and for structural validation<sup>16</sup>. If the calculated score of ProSA deviates from the given range, this indicates that the primary sequence and predicted structure possibly contain flaws<sup>50</sup>. Additionally, the [Verify3D server](#) was employed to examine the congruity of a tertiary structure with its primary sequence. It is done by specifying a basic grade depending upon its position and surroundings and correlating the findings to known verified structures<sup>51</sup>. Ramachandran plot generation was carried out using the [RAMPAGE tool](#)<sup>52</sup>. The Ramachandran analysis plot is a visual representation of energetically permitted and rejected dihedral planar angles based on weak Van der Walls force of interactions between amino acids of the side chain. The RAMPAGE score enumerates the residues residing in favored, allowed and disallowed regions (in %) based on the PROCHECK principle<sup>53</sup>.

### Disulfide bridging of vaccine

Disulfide bonds are intra-protein bonds that can help stabilize tertiary or quaternary interactions in a protein<sup>54</sup>. Therefore, disulfide engineering was accomplished by introducing two novel bonds in the vaccine framework by employing the [Disulfide by Design v2.12 web tool](#). The refined vaccine complex was uploaded, checked for a residue-pair match within parameters of  $\chi_3$  angle  $-87^\circ$  or  $+97^\circ \pm 30$  and  $C\alpha-C\beta-S\gamma$  angle  $114.6^\circ \pm 10$ . Appropriate amino acid duos were chosen and mutated to cysteine residues<sup>55</sup>.

### Codon optimization and *in silico* cloning

Codon refinement is a vital step for the expression of the protein inside the host system. Therefore, codon adaptation was carried out by utilizing the [JCcat web tool](#) in order to improve protein translation. *E.coli* K12 was selected for cloning since unadjusted codon sequences will reduce protein expression<sup>56</sup>. Termination of transcription (rho-independent), prokaryotic ribosome region of attachment and restriction enzyme cut-off sites were disregarded by setting the parameters to “avoid”. The server further provided two additional measures such as the codon adaptation index (CAI) and GC content of the

refined sequences<sup>57</sup>. The codon nucleotide adapted sequence was cloned using SnapGene v5.0.8 software into a *E. coli* pET28a(+) vector obtained from SnapGene<sup>58</sup>. ApE and Genome Compiler are open access alternative to this software that could be used for this purpose.

### Molecular docking between TLR4 and vaccine

Molecular docking is a computerized process that evaluates and assesses the contact between a protein and its receptor. The binding affinity and interaction between the proteins provide a simulated score<sup>59</sup>. TLR4 can recognize viral glycoprotein<sup>60</sup> and therefore, the TLR4 structure was obtained from PDB (PDB ID: 4G8A) and was chosen as a receptor to dock with the refined vaccine<sup>61</sup>. ClusPro v2.0 web docking server was appointed to determine the binding position and inclination for binding between the designed vaccine and TLR4 receptor<sup>62</sup>. The vaccine-receptor composite was selected based on the binding position (on active site), docking efficiency and the lowest energy scoring of the docked complexes.

### Molecular dynamics simulation

Determining the stability and firmness of the protein-receptor docked structure by molecular dynamics is essential for carrying out an *in silico* study. Protein stability was resolved by contrasting protein dynamics and their counterpart normal modes<sup>63,64</sup>. The iMODY tool was used to evaluate and plot amino acid residues and their motion within their inner order via normal mode analysis (NMA)<sup>65</sup>. The NMA tool surfaced the extent of motion of the vaccine-receptor complex in terms of covariance and ability to deform; along with eigenvalues and B-factor results. The deformability of the complex depends on whether it can rotate each of its residues while staying in its position. The eigenvalue demonstrates the rigidity of motion, being directly proportional to the energy required for the deformation of the

structure. Lower eigenvalues represent easier structural deformation<sup>65,66</sup>.

### Estimation of immune response

Simulations of immune response of the vaccine design were executed with the C-IMMSIM v10.1 server. This server uses position-specific scoring matrix (PSSM) and AI-based technology to predict the intensity of immune reaction caused by the multiepitope vaccine<sup>67</sup>. It evaluates the immunologic response of the subunit vaccine according to an *in vivo* system and can also simulate immune response through its agent-based dynamic system. The minimum recommended interval is at least four weeks between the first and second dose<sup>68</sup>. Time-steps for each vaccine dose were fixed: 1, 84 and finally 168 respectively, with an interval of eight hours in-between. Analysis of three doses was conducted in respect to immune cell rise and activity with at least four weeks in between each of the dose sessions.

## Results

### RVFV proteome and selected antigen

Amino acid sequences of a total of 232 RVFV proteins were collected from the ViPR database. The VaxiJen v2.0 server revealed the two proteins with the highest antigenicity (Table 1). The glycoprotein sequence was selected for further analysis.

### Potential CTL, HTL and LBL epitopes

NetCTL v1.2 server predicted 135 unique 9-mer CTL epitopes in total. Out of them, 27 epitopes were predicted as having the ability to act as both antigen and immunogen and were also nontoxic and nonallergenic (Supplementary Table 2, Extended data)<sup>23</sup>. Only six CTL peptides were shortlisted for the ultimate vaccine design based on a combined score (Table 2).

**Table 1. Highest antigenic proteins of Rift Valley fever virus with antigenicity score, and amino acid length.**

Class	Segment	GenBank accession	Length (AA)	Antigenic score
Non-structural	Segment: S	AGB13810	265	0.6010
Glycoprotein	Segment: M	AVK51661	1197	0.5277

**Table 2. Shortlisted cytotoxic T-lymphocyte epitopes for epitope-based peptide vaccine design.**

CTL epitope	C-score*	Antigenicity	Immunogenicity	Allergenicity	Toxicity
VFALAPVVF	1.7143	Yes	Yes	No	Non-Toxic
<b>STAHEVVPF</b>	<b>1.7042</b>	<b>Yes</b>	<b>Yes</b>	<b>No</b>	<b>Non-Toxic</b>
<b>KLTLEITDF</b>	<b>1.627</b>	<b>Yes</b>	<b>Yes</b>	<b>No</b>	<b>Non-Toxic</b>
<b>RDNETS AEF</b>	<b>1.553</b>	<b>Yes</b>	<b>Yes</b>	<b>No</b>	<b>Non-Toxic</b>
<b>FSSVAIICL</b>	<b>1.3561</b>	<b>Yes</b>	<b>Yes</b>	<b>No</b>	<b>Non-Toxic</b>
ALSIGLFFL	1.3051	Yes	Yes	No	Non-Toxic

\*C-score is the combined score provided by the NetCTL v1.2 server.

Similarly, 182 15-mer HTL epitopes and their MHC-II binding alleles were predicted and evaluated by utilizing the IEDB MHC-II tool. Next, cytokine production capability of the selected HTL epitopes was predicted for IL-10, IL-4 and IFN- $\gamma$ . Only 14 epitopes were found to be capable of inducing any of the three cytokines' production and to have antigenicity (Supplementary Table 3, *Extended data*)<sup>23</sup>. Of those that were capable of producing IFN- $\gamma$ , two epitopes were chosen with the ability to produce IL-4 but not IL-10 and one epitope was chosen with the ability to produce IL-10 but not IL-4. Therefore, three epitopes were finally shortlisted prioritizing for IFN- $\gamma$  and IL-4 production (Table 3).

B cells act as antigen-presenting cells that recognize epitopes present on a protein and can bring about a humoral immune response<sup>69</sup>. The iBCE-EL sever program was used to predict unique LBL epitopes from the RVFV glycoprotein sequence. A total of 310 LBL epitopes were found. After further evaluation, 35 unique epitopes were found to be non-allergenic and non-toxic (Supplementary Table 4, *Extended data*)<sup>23</sup>. Amidst those, only five LBL epitopes were chosen for vaccine construction based on iBCE-EL server's probability scores. LBL epitopes with higher probability scores were chosen. (Table 4).

#### Estimated population coverage

The CTL and HTL epitopes across African regions and Southwest Asia have a mean population coverage of 77.39 %, with an average hit of 1.78. The total combined class 1 and 2 epitopes have significant coverage across Europe (99.92%), North America (100%), South America (99.39%) and South Asia (99%), which indicates the multiepitope vaccine would

be a good candidate in eliciting an immune response in these worldwide regions. In the African population, where the initial outbreak began, there is good coverage in the majority of regions, except for South Africa (18.36%), as shown in Figure 2.

#### Epitope based subunit vaccine construct

To design our multiepitope vaccine, we picked four CTL epitopes (Table 2) based on their high immunogenicity and antigenicity scores and which were non-allergic and nontoxic. For the HTL epitope selection, we looked for epitopes that could produce all three types of cytokine; however, no epitope was found that had the capacity to induce all three types of cytokines (IFN- $\gamma$ , IL-10, and IL-4). Hence, only two HTL epitopes were chosen from the glycoprotein sequence, which were capable of inducing IFN- $\gamma$  and IL-4 (Table 3). Three LBL epitopes that were nontoxic and non-allergic and had the best iBCE-EL predicted probability, high antigenicity scores were selected (Table 4). Two extra epitopes were shortlisted for each of the three cases for the purpose of randomization to find the optimum vaccine construct. The choice of adjuvant was 50S ribosomal protein L7/L12, retrieved from the NCBI Protein database (Accession no. P9WHE3) and was linked to the selected epitopes using an EAAAK linker. For the final vaccine construct, four out of six CTL, two out of four HTL and three out of five LBL shortlisted epitopes were chosen based on different combinations and randomization to generate seven potential vaccine candidates, provided in Supplementary Table 5 (see *Extended data*)<sup>23</sup>, and were merged with AAY, GPGPG and KK linkers, respectively (Figure 3). The prospective vaccine with the ideal physicochemical property was chosen. The final assembled vaccine is 262 amino acid residues long.

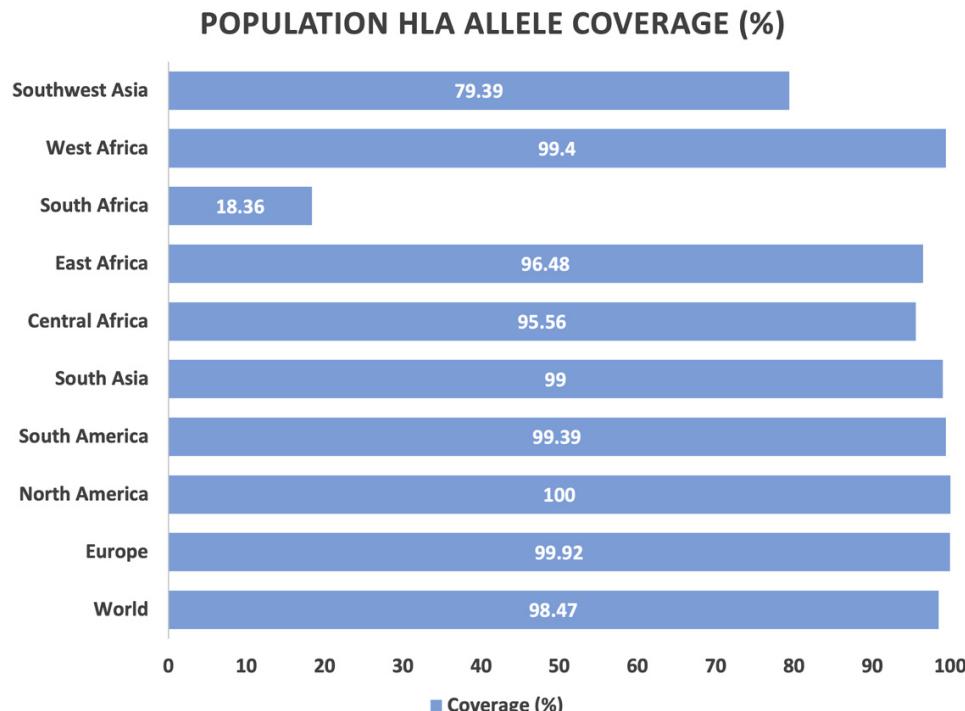
**Table 3. Shortlisted helper T-lymphocyte epitopes for epitope-based peptide vaccine design.**

HTL epitopes	Antigenicity	IFN- $\gamma$ inducer	IL-4 inducer	IL-10 inducer
NTKCRLSGTALIRAG	<b>1.0947</b>	Positive	IL4 inducer	IL10 non-inducer
CNAGARVCLSITSTG	<b>1.115</b>	Positive	IL4 inducer	IL10 non-inducer
TILLICLYVALSIGL	1.2038	Positive	Non-inducer	IL10 inducer
VVFVFSSVAIICLAI	1.108	Positive	Non-inducer	IL10 inducer

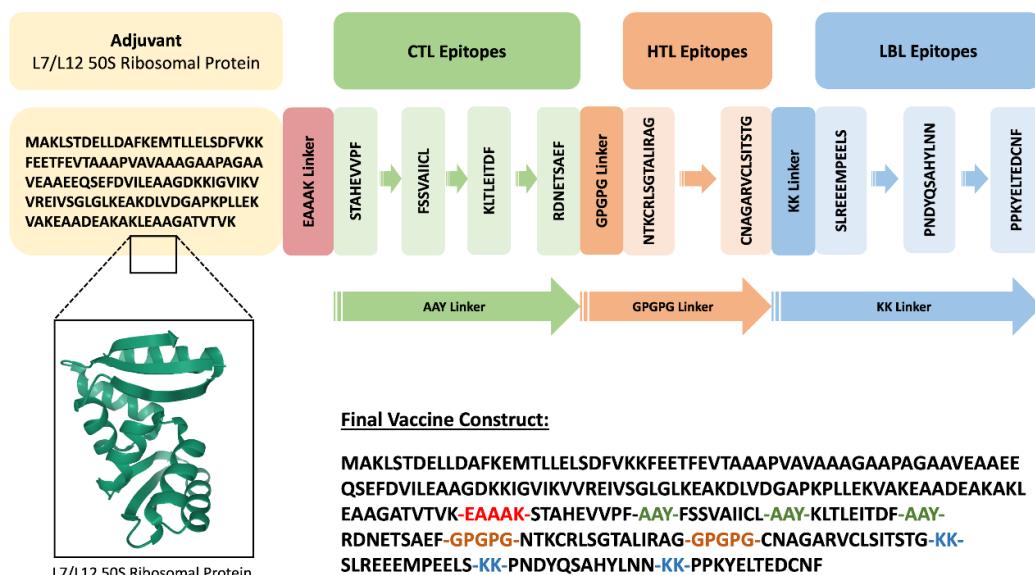
IFN- $\gamma$ , interferon gamma; IL-4, interleukin 4; IL-10, interleukin 10.

**Table 4. Shortlisted linear B-lymphocyte epitopes for epitope-based peptide vaccine design.**

LBL epitopes	Probability	Antigenicity	Allergenicity	Toxicity
KGTMGSGQTKRE	0.8213	Antigen	Non-allergen	Non-Toxic
PNDYQSAHYLNN	<b>0.8102</b>	Antigen	Non-allergen	Non-Toxic
PPKYELTEDCNF	<b>0.8043</b>	Antigen	Non-allergen	Non-Toxic
VTGSQSPSTEIT	0.8042	Antigen	Non-allergen	Non-Toxic
SLREEEMPEELS	<b>0.8039</b>	Antigen	Non-allergen	Non-Toxic



**Figure 2.** Population coverage with selected epitopes and their respective alleles.



**Figure 3.** Diagram of the assembled vaccine construct outlining 262 amino acids. The EAAAK linker was utilized to join the adjuvant at the 5' end to the rest of the vaccine. CTL epitopes linked with AAY linkers, HTL epitopes linked with GPGPG linkers and form a bridge with the last CTL epitope and first HTL epitope. KK linkers were used to connect LBL epitopes and also form a bridge between the last HTL epitope and the first LBL epitope. LBL, linear B-lymphocyte; HTL, helper T-lymphocyte; CTL, cytotoxic T-lymphocyte.

### Host-pathogen antigenic similarity

Evaluation of the multiepitope vaccine protein for similarity against the *Homo sapiens* proteome was done using BLASTp to ensure safety. No similar proteins were found, which negates the chance of autoimmunity as a result of the vaccine.

### Physicochemical properties and solubility

Physicochemical attributes of the vaccine structure were assessed using the ProtParam tool. The vaccine's molecular weight was estimated to be ~28 kD and showed good antigenic characteristics. The predicted Isoelectric point (pI) of the

designed vaccine is 4.96, indicating an acidic nature, and it has an instability index of 27.79, suggesting that the vaccine will maintain good stability inside the host system. The aliphatic index is 84.73, denoting the thermostable nature of the vaccine. The hydrophilic nature of the vaccine was estimated by the grand average of hydropathicity (GRAVY) score of -0.125. Half-life ( $t_{1/2}$ ) was calculated to be greater than 30 hours in a mammalian reticulocyte, greater than 20 hours in yeast, and more than 10 hours in *E. coli*. The overall vaccine construct is not allergenic and demonstrated good solubility as shown by the SOLpro server. The above-mentioned assessments suggest that our multiepitope vaccine design has the potential to be a good candidate for RVFV (Table 5).

### Secondary structural features

The 2D structure of the vaccine design was acquired through the PSIPRED v4.0 workbench. The secondary structure consisted of 49.62%  $\alpha$ -helices (130 amino acids), 11.83%  $\beta$ -strands (31 amino acids) and 38.55% random coils (101 amino acids) (Figure 4).

### Refined 3D structure of the vaccine

The multiepitope vaccine was speculated and constructed using I-TASSER server. The PDB structure (PDB ID: 1DD4) was regarded as the best template for modeling. This server carries out 3D modeling based on the consequence of threading template alignment and simulations are run by merging parameters of the structure assembly. It ranks the confidence of models quantitatively on C-score. The obtained C-score for the modeled 3D vaccine structure was -4.16. The refinement of the 3D structure was carried out by the GalaxyRefine web-tool. The refined and polished model was used to generate a Ramachandran plot, which showed an increased percentage

of residues in the favored region. The refined vaccine construct showed 90.0% residues in the favorable region in Ramachandran plot. The global distance test (GDT-HA) score was found to be 0.9055, RMSD was 0.534, MolProbity was 2.209, clash score was 14.2 and poor rotamers were non-existent (Figure 5).

### Validation properties of the 3D vaccine

The refined tertiary structure was validated using the RAMPAGE server, ProSA-Web tool and Verify3D server. Validation carried out by the RAMPAGE server using a Ramachandran plot of the unrefined 3D structure showed 69.2% of residues in the favorable region, 20.8% in acceptable regions, and 10.0% of the residues in disallowed regions. After structural refinement was carried out, RAMPAGE revealed 90.4% of the vaccine's residues resided in favorable regions, and 6.2% and 3.5% of residues were found in acceptable and disallowed regions, respectively (Figure 6A). The ProSA-web and Verify3D (Figure 7) servers were used to validate the unrefined tertiary structure. After refinement, ProSA-web revealed a Z-score of -6.2 for the best vaccine protein model, implying equivalence to the native protein conformation (Figure 6B).

### Disulfide bridging in the vaccine structure

Disulfide bridging was used to stabilize the refined vaccine model. The introduction of cysteine residues was carried out via Disulfide by Design v2.12. Although 23 potential pairs of residues were found (Supplementary Table 7, *Extended data*)<sup>23</sup> suitable for disulfide engineering, only two pairs of residues were chosen based on the energy of binding and  $\chi_3$  angle. Accordingly, two mutations were created on a pair of selected amino acids, based on the lowest kcal/mol value. For, Leu67-Ala100 residues, the energy value was 2.40 kcal/mol and the  $\chi_3$  angle was -103.13 degrees and for Phe28-Ala37 duos, the  $\chi_3$  angle was -111.70 degrees and energy value was 2.86 kcal/mol (Figure 8).

### Adapted codons and *in silico* cloning

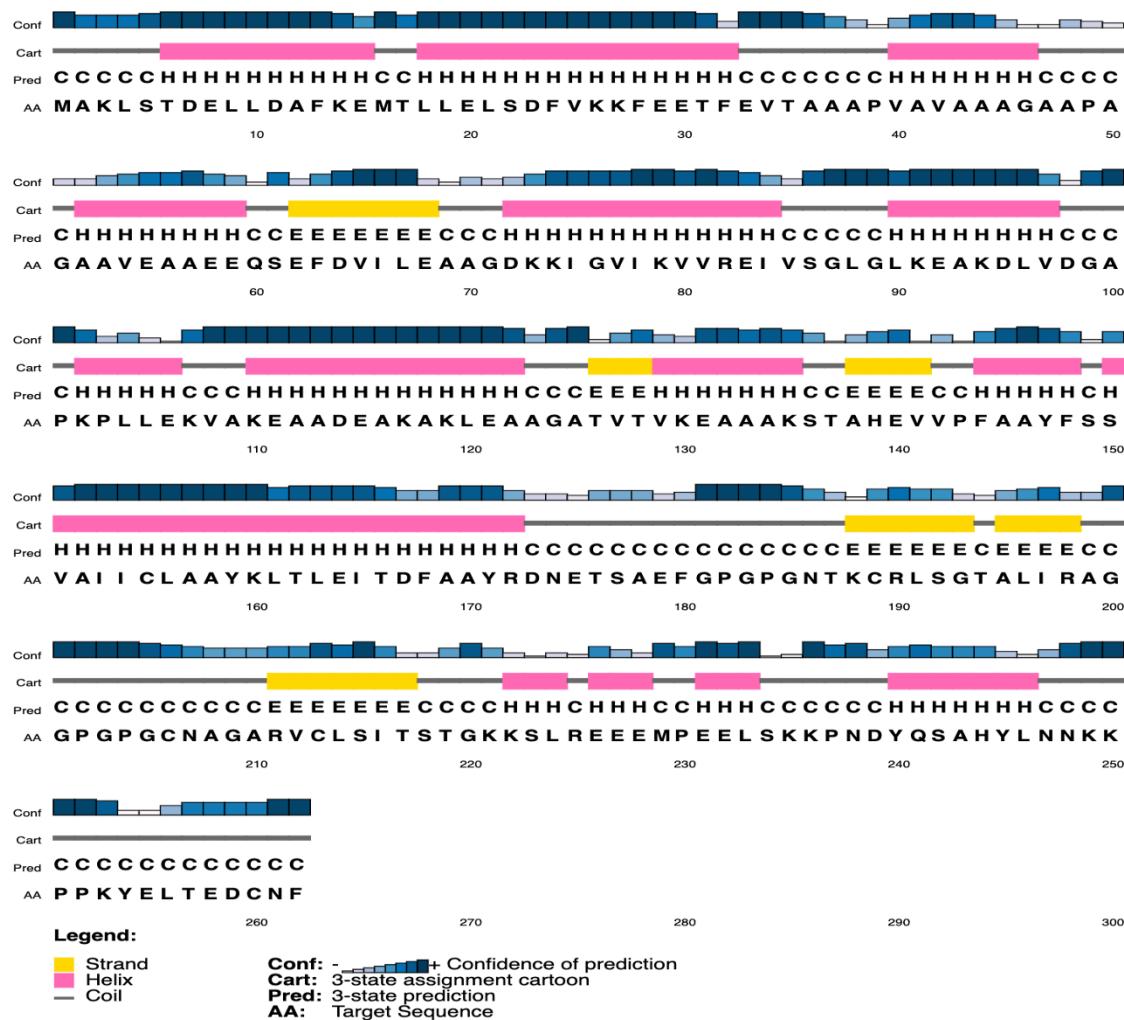
The primary objective is to express our constructed multiepitope vaccine sequence using *in silico* cloning into the *E. coli* expression system. As such, the subunit vaccine construct requires codon optimization according to the codon usage for the expression pattern system in *E. coli*. To optimize for maximal expression of our vaccine design in *E. coli* K12, the JCcat tool was used. The optimized codon sequence was 786 base pairs in length. The GC content was 49.75%, which lies within the ideal range of 30–70%. Codon Adaptation Index was 1.0, which is also within the ideal range (0.8–1.0). This ensures the efficient expression of the multiepitope vaccine inside *E. coli*. XbaI and BamHI restriction sites were inserted into *E. coli* plasmid pET28a(+) and the adapted vaccine's codon sequences were inserted into a recombinant *E. coli* pET28a(+) plasmid using SnapGene software v5.0.8 (Figure 9). The final length of the recombinant construct is 6121 bp.

### Molecular docking studies

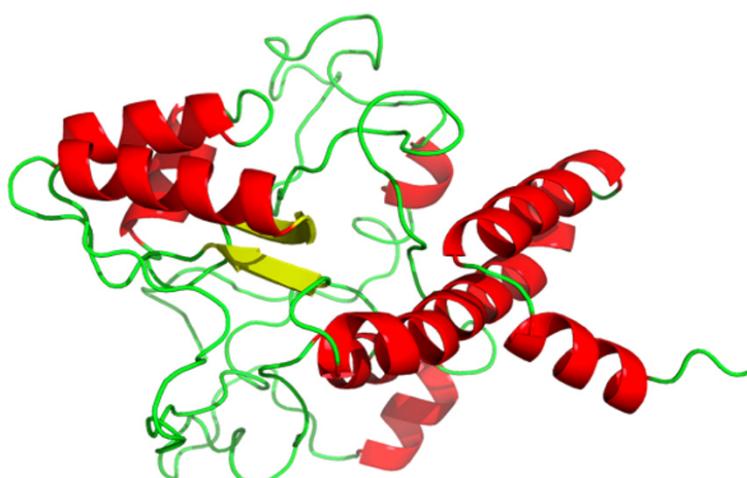
The ClusPro v2.0 server tool was used to conduct molecular docking for characterizing contact between the illustrated refined vaccine and TLR4 (PDB ID: 4G8A). This server

**Table 5. Physicochemical assessments of construct along with antigenic and allergenic properties.**

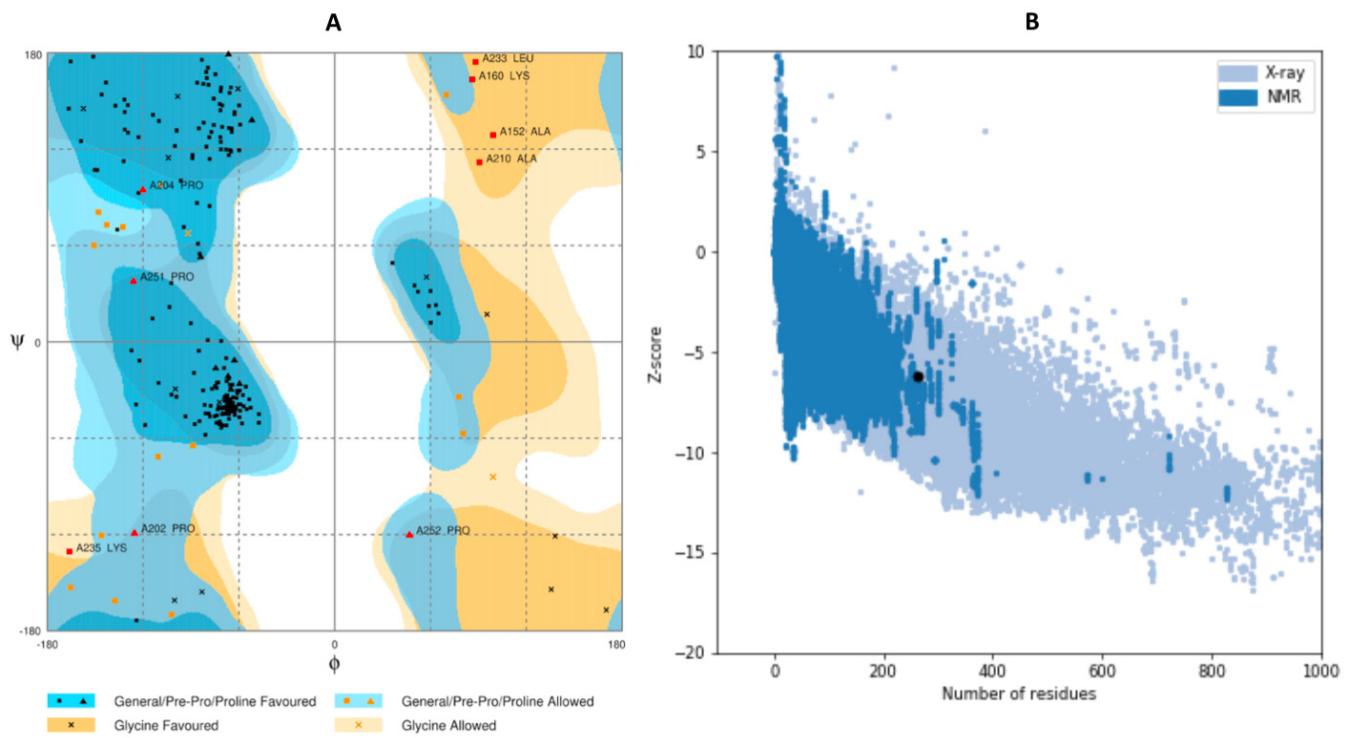
Attributes	Value
Length (amino acids)	262
Molecular weight (Dalton)	27603.43
Chemical formula	C <sub>1223</sub> H <sub>1964</sub> N <sub>318</sub> O <sub>389</sub> S <sub>8</sub>
Isoelectric point (pI)	4.96
Total atoms	3902
Extinction coefficient (at 280 nm in H <sub>2</sub> O)	9190 M <sup>-1</sup> cm <sup>-1</sup>
Instability index (II)	27.79
Aliphatic index (AI)	84.73
Grand average of hydropathicity (GRAVY)	-0.125
Solubility	0.826
Antigenicity (Vaxijen v2.0)	0.5590
Allergenicity (AllerTOP v2.0)	Non-allergen



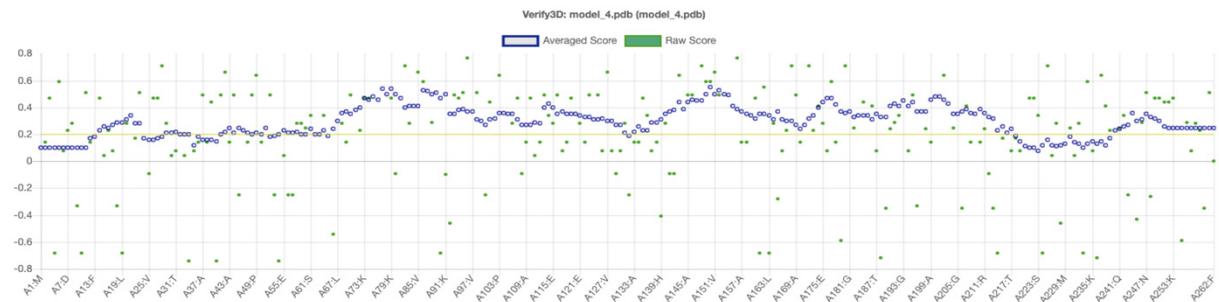
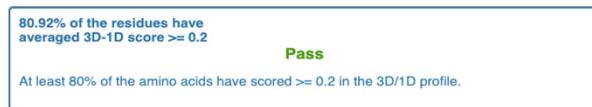
**Figure 4. Secondary structural features of the designed vaccine.** Herein,  $\alpha$ -helix,  $\beta$ -strands, and random coils are represented with pink, yellow and blue colors, respectively.



**Figure 5. The tertiary structure of the vaccine where  $\alpha$ -helix,  $\beta$ -strand, and random coils are shown in red, yellow, and green, respectively.**



**Figure 6. Qualification of our predicted 3D arrangement of the vaccine.** (A) Ramachandran plot analysis of the refined model showing favored, allowed and disallowed regions are 90.4%, 6.2% and 3.5%, respectively. (B) Validation using ProSA web tool, revealing a Z-score of -6.2.



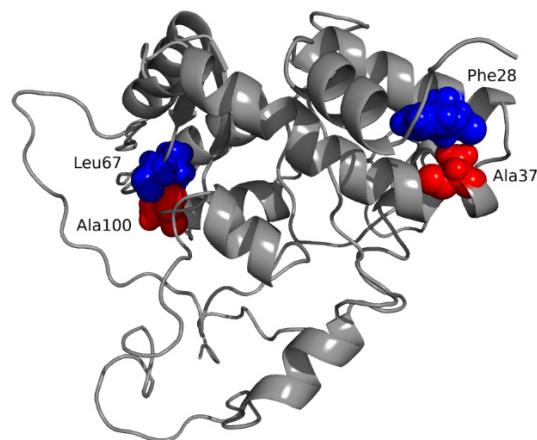
**Figure 7. Verify3D shows a score of 80.92% for our selected vaccine model.**

provided 30 different docked models, the best 10 of which were analyzed for the selection of the appropriate vaccine-receptor complex (Supplementary Table 6, *Extended data*)<sup>23</sup>. The model showing a strong interaction between vaccine residues and the active site of TLR4 along with the lowest energy value was selected. Model 7 fulfilled the above-mentioned criteria, and thus, was picked as the suitable vaccine-receptor complex (Figure 10). Model 7 had an energy score of -809.1 and indicated the

highest binding affinity as it had the lowest energy score (Supplementary Table 6, *Extended data*)<sup>23</sup>.

#### Molecular dynamics simulation

To understand the vaccine-receptor compound stability and its large-scale mobility, NMA was carried out using iMODS tool, which is simulated based on intramural configuration of the whole compound. Deformability of the vaccine-immune

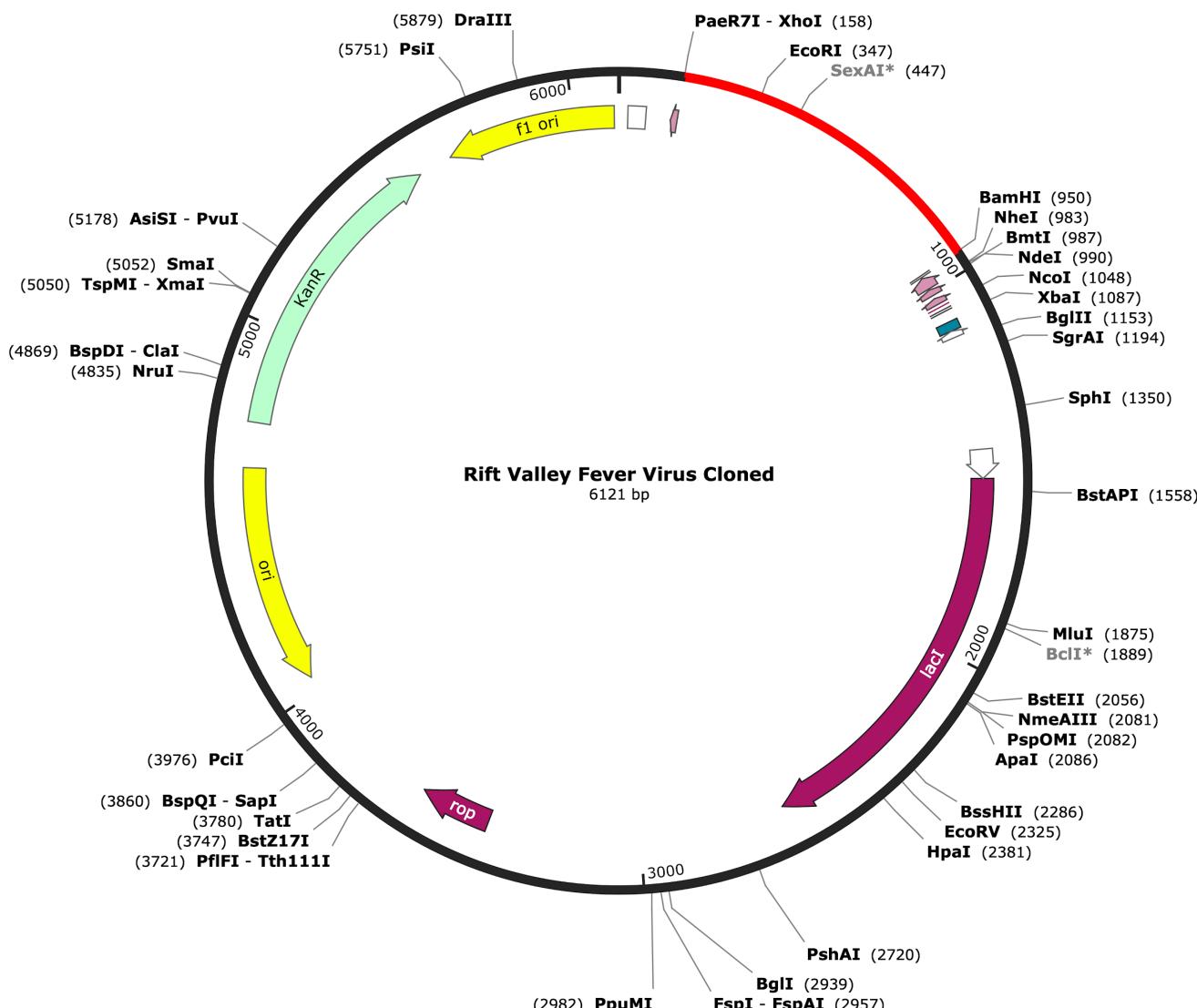


**Figure 8. Disulfide bridging of the vaccine model to increase stability.** The engineered pairs are indicated by blue and red colors.

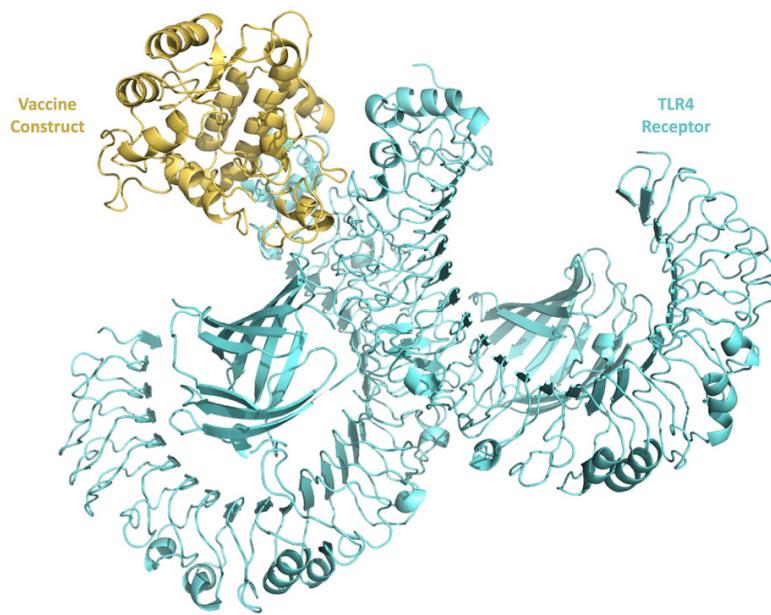
receptor complex depends on the isolated movement of every amino acid residue, illustrated via chain hinges (Figure 11). The eigenvalue was  $2.443455e-5^{57}$  (Figure 11A). The covariance matrix demonstrates the linkage between duplets of amino acids, the correlated residues marked in red, anti-related residues in white and non-correlated residues in blue, interspersed in dynamical zones (Figure 11B). The elastic mesh-work paradigm classifies which pairs of atoms are interlinked by springs and is visualized as a linking matrix. Dots are colored in accordance to their rigidity; the higher the rigidity, the darker the color (Figure 11C)<sup>70</sup>.

#### In silico immune responses

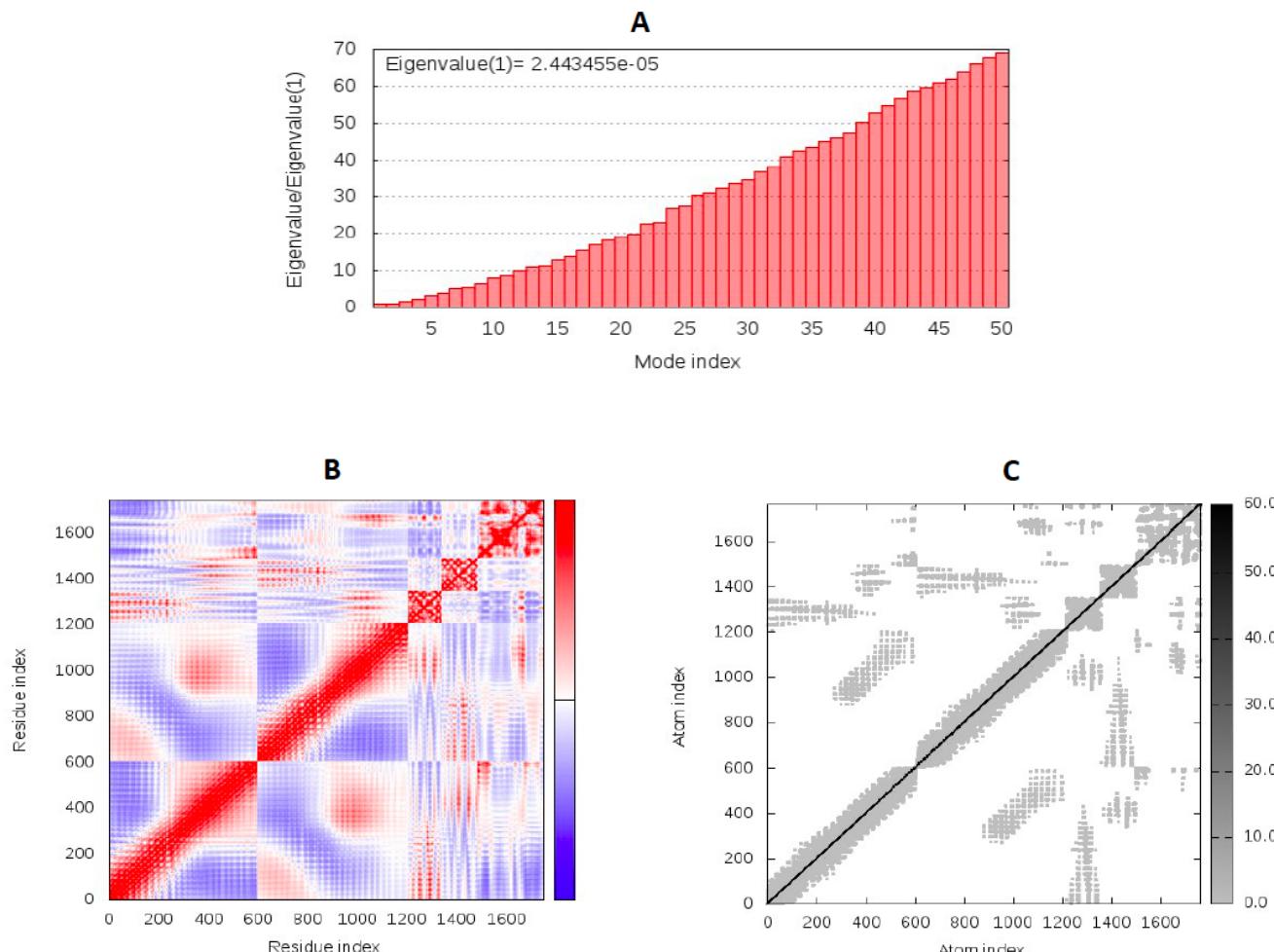
The simulated immune response showed resemblance to an actual immune response against a pathogen (Figure 12). Subsequent exposures produced a higher tier immune response compared to the primary immune response. Secondary and tertiary responses showed higher levels of antibodies (i.e.,



**Figure 9. Recombinant plasmid construct pET28a(+) containing multiepitope vaccine was inserted between XhoI (158 bp) and BamHI (950 bp) restriction sites shown in red.** Outer dark margin represents the vector skeleton.



**Figure 10. Interaction between the vaccine and TLR4 receptor.** Vaccine surface is represented in yellow, while the receptor is represented in cyan.

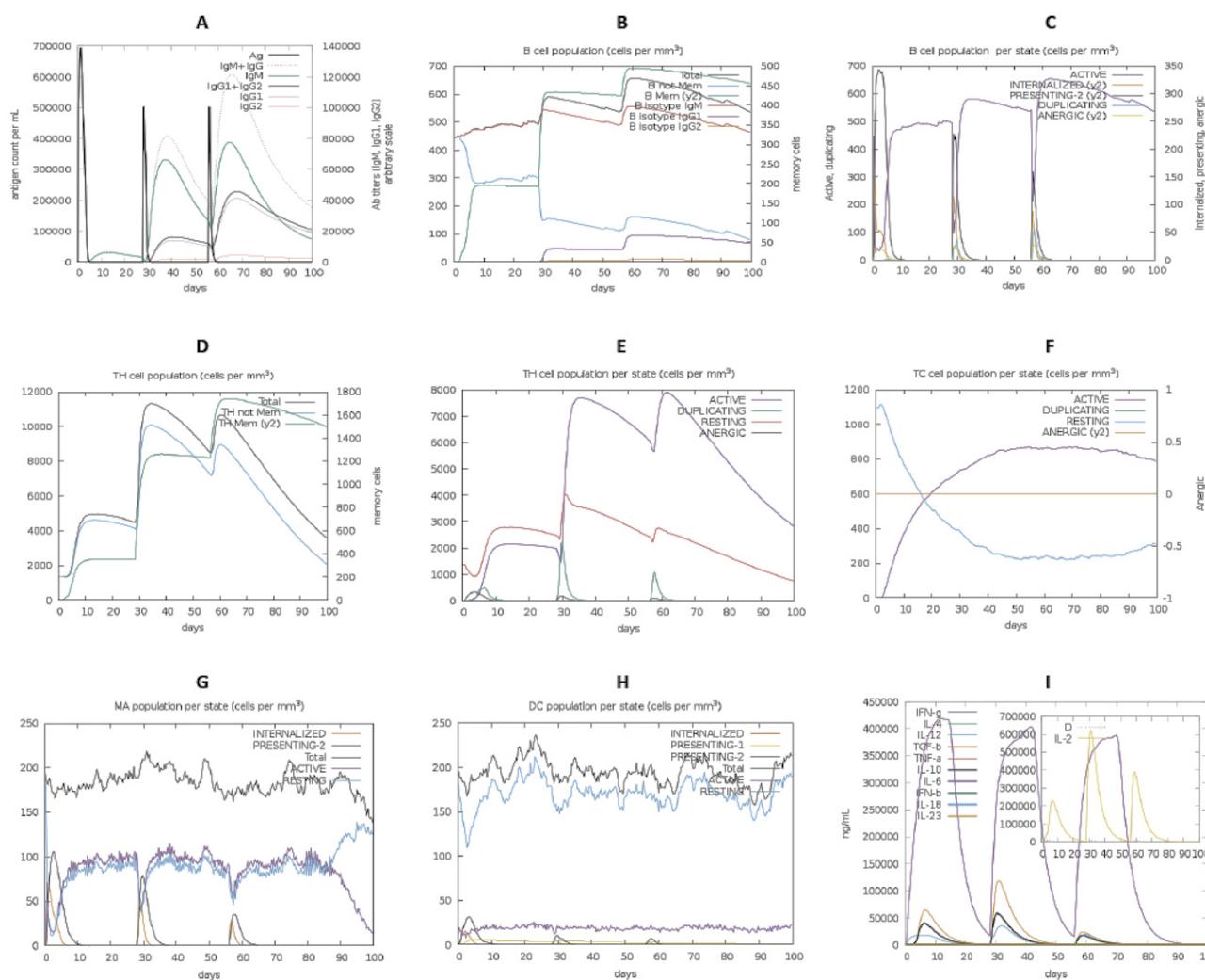


**Figure 11.** Molecular dynamic simulation of vaccine-receptor complex, showing **(A)** eigenvalue; **(B)** covariance matrix; and **(C)** elastic network analysis.

IgM, IgG1, IgG2), which coincides with the waning of antigen (Figure 12A). This demonstrates the development of immune memory cells and, as a result, intensified antigen neutralization upon successive exposure (Figure 12A). Additionally, several B-lymphocyte cell isotypes with prolonged life have been noticed, signifying potential class switching and memory B cell generation (Figure 12B–12C). Increased responses were also observed in helper T cells and cytotoxic T cells with memory generation (Figure 12D–12F). An increase in macrophage activity and engagement was perceived, with a vigorous proliferation of dendritic cells (Figure 12G–12H). An increased amount of IFN- $\gamma$  and IL-2 cytokines were also noticed (Figure 12I). These findings suggest the development of immune memory and, therefore, immunity against the virus.

## Discussion

RVFV causes higher mortality and fatality among animals than in humans. Since this is a zoonotic virus, it can be transmitted through mosquitos and many regions like Africa, the Arabian Peninsula and their neighbors are at risk of an endemic outbreak<sup>7</sup>. Thus, a prophylactic measure against RVFV is essential. Vaccination is an ideal approach to gain immunity against infectious agents like RVFV. Development and manufacturing vaccines using a live or attenuated virus takes a huge amount of time and money. Moreover, higher antigenic presence in a weakened vaccine may overstimulate the protective immune system and complicate the situation by causing hypersensitivity reactions<sup>21</sup>. In comparison with conventional vaccine strategies, multiepitope vaccines possess no such complications<sup>71</sup>.



**Figure 12.** *In-silico* immune response for vaccine construction as an antigen using C-IMMSIM server tool: **(A)** production of immunoglobulin due to injection of the vaccine; **(B)** B cell count following consecutive three injections; **(C)** population of B cell number per state; **(D)** helper T cell activation; **(E)** population of helper T cell number per state; **(F)** population of cytotoxic T cell number per state; **(G)** per state macrophage population; **(H)** dendritic cell population; and **(I)** level of cytokine and interleukins.

Epitope-based vaccines could be a great choice in terms of safety, viability and economic rationality. In addition, potency as well as immune responses of multiepitope vaccines can be enhanced through the deliberate engineering of targeted epitopes<sup>72</sup>. Researchers have been looking for a way to reduce the budget for vaccine development, as well as minimizing allergenicity and side-effects of vaccines for quite a long time. With the emergence of new computational technology and widely available databases, different strategies are now available for designing and developing epitope-based vaccines following immunoinformatics expedients<sup>73,74</sup>. Vaccines against multitudes of viruses such as SARS-CoV-2, Ebola, Lassa, hepatitis C, Oropouche, Dengue and many more in the pipeline are prime examples of a structural vaccinology approach applied to the design of a multiepitope vaccine model<sup>17,19,75-79</sup>. In this study, our primary focus was to devise a multiepitope vaccine with the ability to produce robust immunity against RVFV having considered all the parameters of a subunit vaccine.

After retrieving the 232 complete proteomes of RVFV from the ViPR server, screening for antigenic proteins was executed and glycoprotein sequences were selected based on higher antigenicity. A potent multiepitope vaccine should possess the potential to trigger B and T cell activation<sup>71</sup>. Therefore, potential CTL, HTL as well as LBL epitopic regions of RVFV glycoprotein were chosen for a candidate vaccine modeling. We were particularly interested in integrating B cell epitopes due to its function in antibody production and memory cell formation<sup>80</sup>. T cell-mediated immunity was also a concern in our vaccine design since plasma cells that give rise to humoral response reactions can easily be saturated by the deluge of antigens. Moreover, lifelong resistance can also be achieved by cell-mediated immunity or cytotoxic T cells<sup>81</sup>. Cytotoxic T cells can provide lifelong immunity through identifying and destroying infected cells<sup>82</sup>. Helper T cells, on the other hand, stimulate the release of IL-10, IL-4 and IFN- $\gamma$  to overcome pro-inflammatory responses and lessen the damage caused to tissues and cells. Besides, they help to produce IgG antibodies and neutralize RVFV infection from peripheral system<sup>83</sup>. Furthermore, CD4+ cell activate B cells, macrophages, CD8+ cells when they come into contact with an antigen. Thus, all these points were considered and examined during our pursuit of designing the RVFV vaccine.

In a previous study, Adhikari and Rahman predicted overlapping immunodominant T cell epitopes from both nucleocapsid and glycoprotein<sup>84</sup>. However, their study concluded without formulating the vaccine candidate. Therefore, while Adhikari and Rahman focused on finding overlapping conserved CD8+ and CD4+ epitopes, our work was more centered on selecting CD8+ and CD4+ epitopes with high antigenicity, immunogenicity, non-allergenicity, non-toxicity and eventually tailoring them into a rational and potent vaccine candidate. Herein, unique epitopes from helper and cytotoxic T cells and B cells were chosen not only based on their antigenicity but also on other factors including allergenicity, immunogenicity and toxicity. The complete vaccine design was assembled by attaching the chosen CTL, followed by HTL and finally B cell

peptides with the help of AAY, GPGPG, and KK linkers, respectively. Linkers were incorporated in vaccine construction as a part of an essential element to enhance stability, folding, and expression patterns of our vaccine protein<sup>85</sup>. The adjuvant L7/L12 ribosomal protein was attached to the first CTL epitope using the linker EAAAK. Multiepitope vaccines are less immunogenic when used alone due to a reduction in molecular weight compared to the protein, hence, it requires adjuvant to enhance its efficiency<sup>86</sup>. Adjuvants are components that help heighten cellular and humoral immunogenic responses for particular antigens as well as amplify the vaccine's stability and longevity<sup>87,88</sup>.

Due to the inclusion of immune dominant epitopes, the vaccine candidate was found to have higher antigenicity and immunogenicity while being devoid of allergenic feature. These are the prime features for a vaccine to be effective immune modulator in the first place. When it comes to peptide or subunit vaccines, the size of the vaccine candidate becomes an important matter. Interestingly, our designed multiepitope vaccine is only ~28 kD in size which makes it a suitable candidate vaccine. Solubility is another vital characteristic for any recombinant vaccine<sup>16</sup>. Luckily, our vaccine construct was predicted to be highly soluble inside the host *E. coli* system. Furthermore, physicochemical properties of the vaccine candidate were also in suitable range. For instance, the instability index suggested that the chimeric protein would remain stable after synthesis, while the GRAVY value and aliphatic index portrayed the vaccine to be hydrophilic and thermostable, respectively. However, the theoretical pI anticipated our vaccine as acidic which can be adjusted by modifying or adding some basic linkers or additional adjuvants.

The arrangement of crucial protein components was determined by 3D structure modeling, which was used for studying protein dynamics, functions, networking between residues and ligand interactions. The vaccine candidate that showed the best physio-chemical property was chosen and the vaccine construct was refined to significantly improve its expedient properties. The Ramachandran plot revealed 96.6% residues resided within the combined favored and acceptable regions along with only a couple of residues in disallowed region, which confirms the merit of the model. The GDT-HA and MolProbity score, along with clashscore, RMSD value and poor rotamer values were also indicative of sufficiency of the designed vaccine. The Z-score (-6.2) and Verify3D score (80.92%) also signal the satisfactory quality of the vaccine.

Molecular docking as well as molecular dynamics simulation of the TLR4 and vaccine complex were performed in order to gain knowledge about the binding strength, contact nature and structural integrity. The vaccine-receptor complex underwent energy minimization, resulting in boosting of the stability of the total combination. The eigenvalue suggests the stiffness of motion of the molecules within the system and energy needed for their distortion. In our study, eigenvalue increases with each mode passed by which indicates more rigid and compact structure given the time.

Serological evaluation to verify immuno-reactivity can be carried out for validating the candidate vaccine<sup>89</sup>. Thus, observation of the designed vaccine protein's expression inside a suitable host is necessary. *E. coli* has been considered to be a great host choice for determining a recombinant protein's expression patterns<sup>90,91</sup>. Hence, codon adaptation was accomplished for our recombinant vaccine to generate high levels of protein expression in *E. coli* K12. A CAI value of 1.0 and 49.75% GC content suggests the capability of optimum protein synthesis in the host. Moreover, two disulfide bridges were engineered into the vaccine for better stabilization, which is paramount for various biological and biotechnological applications. Animals immunized with glycoproteins of RVFV have been seen to produce protective antibodies against RVFV. Schmaljohn *et al.* have stated that mice inoculated with Sf9 cells expressing G1 and G2 glycoproteins of RVFV have produced antibodies that are adaptive in nature and neutralizes RVFV infection<sup>92</sup>. In a study conducted by Faburay *et al.* construction of a peptide vaccine consisting of two glycoprotein subunits Gn and Gc was carried out against RVFV, and injected into sheep to determine its neutralizing effects. Six sheep were immunized with their constructed subunit vaccine with a dose of 50 µg. The vaccine elicited a strong antibody response against RVFV which was confirmed using enzyme-linked immunosorbent assay (ELISA), suggesting recombinant glycoproteins can be used as a good subunit vaccine candidate. Faburay and his team further carried out a plaque reduction neutralization test to check the amount antibodies produced on primary and secondary dosages. Titer of first dose was low, but the second dosage increased the titer over 1280<sup>14</sup>.

In current study, we also carried out the immune simulation of the vaccine candidate in an *in silico* immune simulator which demonstrated a good pattern of immune responses. As repetitive vaccine doses enhanced immune responses, we administered three doses that produced a variety of B-cell isotypes and T cell-mediated immune reactions with a noticeably significant number of memory B cells having a half-life of several months. Our simulated immune response shows a stronger and active immune response on secondary and tertiary doses, compared to the initial primary dose. IgG and IgM antibodies had an arbitrary titer of 80000 on second dose and 120000 on third dose. Sustained generation of IFN-γ and IL-2 was observed after multiple exposures of the vaccine as a result of increased helper T cell activation and therefore, the vaccine effectively stimulated a humoral immune response to ramp up immunoglobulin production.

## Conclusion

RVFV is a pathogenic agent with potential to become more widespread in upcoming times. Multiepitope vaccines provide a modular and tunable approach to vaccine design compared to traditional vaccine efforts. Our study focused on designing an effective and potent multiepitope vaccine to provide immunity

against RVFV. The vaccine contained *in silico* assessed and predicted CTL, HTL, and LBL epitopes to produce an effective cellular and humoral immunologic response. Moreover, the vaccine simulation showed a potentially protective immune response and passed all molecular simulations with favorable results. However, further validation is required in *in vivo* and *in vitro* systems to guarantee the effectiveness of our designed vaccine.

## Data availability

### Underlying data

All data underlying the results are available as part of the article and no additional source data are required.

### Extended data

Harvard Dataverse: Extended DATA: Discovery of potential immune epitopes and peptide vaccine design - a prophylactic strategy against Rift Valley Fever Virus. <https://doi.org/10.7910/DVN/B9Y1EH><sup>23</sup>.

This project contains the following extended data in DOCX format:

- Supplementary Table 1 (All initially retrieved complete RVFV protein sequences)
- Supplementary Table 2 (Predicted all potential cytotoxic T-lymphocyte epitopes)
- Supplementary Table 3 (Predicted all potential helper T-lymphocyte epitopes)
- Supplementary Table 4 (Predicted all potential linear B-lymphocyte epitopes)
- Supplementary Table 5 (Physicochemical properties of all designed vaccine candidates)
- Supplementary Table 6 (Scores of top-10 best vaccine-TLR4 docked complexes)
- Supplementary Table 7 (Potential residue pairs for disulfide bridging)

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

## Author endorsement

Mohammad Minnatul Karim confirms that the author has an appropriate level of expertise to conduct this research, and confirms that the submission is of an acceptable scientific standard. Mohammad Minnatul Karim declares they have no competing interests. Affiliation: Department of Biotechnology and Genetic Engineering, Islamic University, Bangladesh.

## References

1. Davies FG, Linthicum KJ, James AD: Rainfall and epizootic Rift Valley fever. *Bull World Health Organ.* 1985; **63**(5): 941–3.  
[PubMed Abstract](#) | [Free Full Text](#)
2. Bird BH, Ksiazek TG, Nichol ST, et al.: Rift Valley fever virus. *J Am Vet Med Assoc.* 2009; **234**(7): 883–93.  
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Bouloy M, Weber F: Molecular Biology of Rift Valley Fever Virus. *Open Virol J.* 2010; **4**(1): 8–14.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Daubney R, Hudson JR, Garnham PC: Enzootic hepatitis or rift valley fever. An undescribed virus disease of sheep cattle and man from east africa. *J Pathol Bacteriol.* 1931; **34**(4): 545–79.  
[Publisher Full Text](#)
5. Ikegami T, Makino S: The Pathogenesis of Rift Valley Fever. *Viruses.* 2011; **3**(5): 493–519.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Coetzer JAW, Thomson GR, Tustin RC: Infectious diseases of livestock with special reference to Southern Africa. Cape Town; New York: Oxford University Press; 1994.  
[Reference Source](#)
7. LaBeaud AD, Muchiri EM, Ndovu M, et al.: Interepidemic Rift Valley Fever Virus Seropositivity, Northeastern Kenya. *Emerg Infect Dis.* 2008; **14**(8): 1240–6.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Antonis Afg, Kortekaas J, Kant J, et al.: Vertical Transmission of Rift Valley Fever Virus Without Detectable Maternal Viremia. *Vector-Borne Zoonotic Dis.* 2013; **13**(8): 601–6.  
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Hartman A: Rift Valley Fever. *Clin Lab Med.* 2017; **37**(2): 285–301.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Tigoi C, Sang R, Chepkorir E, et al.: High risk for human exposure to Rift Valley fever virus in communities living along livestock movement routes: A cross-sectional survey in Kenya. *PLoS Negl Trop Dis.* 2020; **14**(2): e0007979.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Anyangu AS, Gould LH, Sharif SK, et al.: Risk Factors for Severe Rift Valley Fever Infection in Kenya, 2007. *Am J Trop Med Hyg.* 2010; **83**(2 Suppl): 14–21.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Gerrard SR, Bird BH, Albariño CG, et al.: The NSm proteins of Rift Valley fever virus are dispensable for maturation, replication and infection. *Virology.* 2007; **359**(2): 459–65.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Besselaar TG, Blackburn NK: The effect of neutralizing monoclonal antibodies on early events in Rift Valley fever virus infectivity. *Res Virol.* 1994; **145**(1): 13–9.  
[PubMed Abstract](#) | [Publisher Full Text](#)
14. Faburay B, Lebedev M, McVey DS, et al.: A Glycoprotein Subunit Vaccine Elicits a Strong Rift Valley Fever Virus Neutralizing Antibody Response in Sheep. *Vector Borne Zoonotic Dis.* 2014; **14**(10): 746–56.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Nezafat N, ESLAMI M, Negahdaripour M, et al.: Designing an efficient multi-epitope oral vaccine against *Helicobacter pylori* using immunoinformatics and structural vaccinology approaches. *Mol Biosyst.* 2017; **13**(4): 699–713.  
[PubMed Abstract](#) | [Publisher Full Text](#)
16. Khatoon N, Pandey RK, Prajapati VK: Exploring *Leishmania* secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep.* 2017; **7**(1): 8285.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Samad A, Ahammad F, Nain Z, et al.: Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach. *J Biomol Struct Dyn.* 2020; **1**: 1–17.  
[PubMed Abstract](#) | [Publisher Full Text](#)
18. Nain Z, Karim MM, Sen MK, et al.: Structural basis and designing of peptide vaccine using PE-PGRS family protein of *Mycobacterium ulcerans*-An integrated vaccinomics approach. *Mol Immunol.* 2020; **120**: 146–63.  
[PubMed Abstract](#) | [Publisher Full Text](#)
19. Sayed SB, Nain Z, Khan MdSA, et al.: Exploring Lassa Virus Proteome to Design a Multi-epitope Vaccine Through Immunoinformatics and Immune Simulation Analyses. *Int J Pept Res Ther.* 2020; **1**: 1–19.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Nain Z, Abdulla F, Rahman MM, et al.: Proteome-wide screening for designing a multi-epitope vaccine against emerging pathogen *Elizabethkingia anophelis* using immunoinformatic approaches. *J Biomol Struct Dyn.* 2019; **1**: 1–18.  
[PubMed Abstract](#) | [Publisher Full Text](#)
21. Li W, Joshi MD, Singhania S, et al.: Peptide Vaccine: Progress and Challenges. *Vaccines (Basel).* 2014; **2**(3): 515–36.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Gaafar BBM, Ali SA, Abd-elrahman KA, et al.: Immunoinformatics Approach for Multiepitope Vaccine Prediction from H, M, F, and N Proteins of Peste des Petits Ruminants Virus. *J Immunol. Hindawi.* 2019; **2019**: e6124030.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Nain Z, Bhuiyan MA, Quayum ST, et al.: "Extended DATA: Discovery of Potential Immune Epitopes and Peptide Vaccine Design - A Prophylactic Strategy against Rift Valley Fever Virus". Harvard Dataverse, V2. 2020. <http://www.doi.org/10.7910/DVN/B9Y1EH>
24. Pickett BE, Sadat EL, Zhang Y, et al.: ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res.* 2012; **40**(Database issue): D593–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. Doychinova IA, Flower DR: Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics.* 2007; **8**(1): 4.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Larsen MV, Lundsgaard C, Lambeth K, et al.: Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics.* 2007; **8**(1): 424.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Moutaftsi M, Peters B, Pasquetto V, et al.: A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat Biotechnol.* 2006; **24**(7): 817–9.  
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Calis JJ, Maybeno M, Greenbaum JA, et al.: Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput Biol.* 2013; **9**(10): e1003266.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Dimitrov I, Bangov IP, Flower DR, et al.: AllerTOP v.2—a server for *in silico* prediction of allergens. *J Mol Model.* 2014; **20**(6): 2278.  
[PubMed Abstract](#) | [Publisher Full Text](#)
30. Gupta S, Kapoor P, Chaudhary K, et al.: *In Silico* Approach for Predicting Toxicity of Peptides and Proteins. *PLoS One.* 2013; **8**(9): e73957.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Alexander J, Fikes J, Hoffman S, et al.: The optimization of helper T lymphocyte (HTL) function in vaccine development. *Immunol Res.* 1998; **18**(2): 79–92.  
[PubMed Abstract](#) | [Publisher Full Text](#)
32. Wang P, Sidney J, Dow C, et al.: A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol.* 2008; **4**(4): e1000048.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. Dhanda SK, Vir P, Raghava GP: Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct.* 2013; **8**(1): 30.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
34. Dhanda SK, Gupta S, Vir P, et al.: Prediction of IL4 Inducing Peptides. *Clin Dev Immunol. Hindawi.* 2013; **2013**: e263952.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Nagpal G, Usmani SS, Dhanda SK, et al.: Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Sci Rep.* 2017; **7**(1): 42851.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Manavalan B, Govindaraj RG, Shin TH, et al.: iBCE-EL: A New Ensemble Learning Framework for Improved Linear B-Cell Epitope Prediction. *Front Immunol.* 2018; **9**: 1695.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
37. González-Galarza FF, Takeshita LYC, Santos EJM, et al.: Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res.* 2015; **43**(Database issue): D784–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
38. Bui HH, Sidney J, Dinh K, et al.: Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics.* 2006; **7**: 153.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
39. Olejnik J, Hume AJ, Mühlberger E: Toll-like receptor 4 in acute viral infection: Too much of a good thing. *PLoS Pathog.* 2018; **14**(12): e1007390.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Pandey RK, Bhatt TK, Prajapati VK: Novel Immunoinformatics Approaches to Design Multi-epitope Subunit Vaccine for Malaria by Investigating Anopheles Salivary Protein. *Sci Rep.* 2018; **8**(1): 1125.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
41. Wei XH, Wei X, Chen FY, et al.: The Upregulation of Translocator Protein (18 kDa) Promotes Recovery from Neuropathic Pain in Rats. *J Neurosci.* 2013; **33**(4): 1540–51.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
42. Arai R, Ueda H, Kitayama A, et al.: Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng Des Sel.* 2001; **14**(8): 529–32.  
[PubMed Abstract](#) | [Publisher Full Text](#)
43. Gu Y, Sun X, Li B, et al.: Vaccination with a Paramyosin-Based Multi-Epitope Vaccine Elicits Significant Protective Immunity against *Trichinella spiralis* Infection in Mice. *Front Microbiol.* 2017; **8**: 1475.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
44. Altschul SF, Gish W, Miller W, et al.: Basic local alignment search tool. *J Mol Biol.* 1990; **215**(3): 403–10.  
[PubMed Abstract](#) | [Publisher Full Text](#)

45. Wilkins MR, Gasteiger E, Bairoch A, et al.: **Protein identification and analysis tools in the ExPASy server.** *Methods Mol Biol.* 1999; **112**: 531–52.  
[PubMed Abstract](#) | [Publisher Full Text](#)
46. Magnan CN, Randall A, Baldi P: **SOLpro: accurate sequence-based prediction of protein solubility.** *Bioinformatics.* 2009; **25**(17): 2200–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
47. Buchan DWA, Minneci F, Nugent TCO, et al.: **Scalable web services for the PSIPRED Protein Analysis Workbench.** *Nucleic Acids Res.* 2013; **41**(W1): W349–57.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
48. Yang J, Yan R, Roy A, et al.: **The I-TASSER Suite: protein structure and function prediction.** *Nat Methods.* 2015; **12**(1): 7–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
49. Heo L, Park H, Seok C: **GalaxyRefine: protein structure refinement driven by side-chain repacking.** *Nucleic Acids Res.* 2013; **41**(Web Server issue): W384–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. Wiederstein M, Sippl MJ: **ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.** *Nucleic Acids Res.* 2007; **35**(Web Server issue): W407–10.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
51. Eisenberg D, Lüthy R, Bowie JU: **VERIFY3D: assessment of protein models with three-dimensional profiles.** *Methods Enzymol.* 1997; **277**: 396–404.  
[PubMed Abstract](#) | [Publisher Full Text](#)
52. Lovell SC, Davis IW, Arendall WB, et al.: **Structure validation by Ca geometry:  $\phi, \psi$  and  $\text{C}\beta$  deviation.** *Proteins Struct Funct Bioinforma.* 2003; **50**(3): 437–50.  
[Publisher Full Text](#)
53. Laskowski RA, MacArthur MW, Moss DS, et al.: **PROCHECK: a program to check the stereochemical quality of protein structures.** *J Appl Crystallogr.* 1993; **26**(2): 283–91.  
[Publisher Full Text](#)
54. Hodder AN, Crewther PE, Matthew ML, et al.: **The disulfide bond structure of Plasmodium apical membrane antigen-1.** *J Biol Chem.* 1996; **271**(46): 29446–52.  
[PubMed Abstract](#) | [Publisher Full Text](#)
55. Craig DB, Dombrowski AA: **Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins.** *BMC Bioinformatics.* 2013; **14**(1): 346.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
56. Grote A, Hiller K, Scheer M, et al.: **JCat: a novel tool to adapt codon usage of a target gene to its potential expression host.** *Nucleic Acids Res.* 2005; **33**(Web Server issue): W526–31.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
57. Sharp PM, Li WH: **The codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications.** *Nucleic Acids Res.* 1987; **15**(3): 1281–95.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
58. "SnapGene software (from Insightful Science; available at snapgene.com)".
59. Lengauer T, Rarey M: **Computational methods for biomolecular docking.** *Curr Opin Struct Biol.* 1996; **6**(3): 402–6.  
[PubMed Abstract](#) | [Publisher Full Text](#)
60. Hise AG, Taylor Z, Hall NB, et al.: **Association of Symptoms and Severity of Rift Valley Fever with Genetic Polymorphisms in Human Innate Immune Pathways.** *PLoS Negl Trop Dis.* 2015; **9**(3): e0003584.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
61. Berman HM, Battistuz T, Bhat TN, et al.: **The Protein Data Bank.** *Acta Crystallogr D Biol Crystallogr.* 2002; **58**(Pt 6 No 1): 899–907.  
[PubMed Abstract](#) | [Publisher Full Text](#)
62. Kozakov D, Hall DR, Xia B, et al.: **The ClusPro web server for protein–protein docking.** *Nat Protoc.* 2017; **12**(2): 255–78.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
63. Aalten DMV, Groot BLD, Findlay JBC, et al.: **A comparison of techniques for calculating protein essential dynamics.** *J Comput Chem.* 1997; **18**(2): 169–81.  
[Publisher Full Text](#)
64. Wüthrich K, Wagner G, Richarz R, et al.: **Correlations between internal mobility and stability of globular proteins.** *Biophys J.* 1980; **32**(1): 549–60.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
65. López-Blanco JR, Aliaja JJ, Quintana-Ortí ES, et al.: **iMODS: internal coordinates normal mode analysis server.** *Nucleic Acids Res.* 2014; **42**(Web Server issue): W271–6.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
66. López-Blanco JR, Garzón JJ, Chacón P: **iMod: multipurpose normal mode analysis in internal coordinates.** *Bioinformatics.* 2011; **27**(20): 2843–50.  
[PubMed Abstract](#) | [Publisher Full Text](#)
67. Rapin N, Lund O, Bernaschi M, et al.: **Computational Immunology Meets Bioinformatics: The Use of Prediction Tools for Molecular Binding in the Simulation of the Immune System.** *PLOS ONE.* 2010; **5**(4): e9862.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
68. Castiglione F, Mantile F, De Berardinis P, et al.: **How the Interval between Prime and Boost Injection Affects the Immune Response in a Computational Model of the Immune System.** *Comput Math Methods Med.* Hindawi; 2012; **2012**: 842329.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
69. Chen X, Jensen PE: **The role of B lymphocytes as antigen-presenting cells.** *Arch Immunol Ther Exp (Warsz).* 2008; **56**(2): 77–83.  
[PubMed Abstract](#) | [Publisher Full Text](#)
70. Kovacs JA, Chacón P, Abagyan R: **Predictions of protein flexibility: First-order measures.** *Proteins.* 2004; **56**(4): 661–8.  
[PubMed Abstract](#) | [Publisher Full Text](#)
71. Zhang L: **Multi-epitope vaccines: a promising strategy against tumors and viral infections.** *Cell Mol Immunol.* 2018; **15**(2): 182–4.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
72. Shey RA, Ghogomu SM, Esoh KK, et al.: **In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases.** *Sci Rep.* 2019; **9**(1): 4409.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
73. Seib KL, Zhao X, Rappuoli R: **Developing vaccines in the era of genomics: a decade of reverse vaccinology.** *Clin Microbiol Infect.* 2012; **18**(Suppl 5): 109–16.  
[PubMed Abstract](#) | [Publisher Full Text](#)
74. María RAR, Arturo CJV, Alicia JA, et al.: **The Impact of Bioinformatics on Vaccine Design and Development.** *Vaccines.* 2017.  
[Publisher Full Text](#)
75. Dash R, Das R, Junaid M, et al.: **In silico-based vaccine design against Ebola virus glycoprotein.** *Adv Appl Bioinform Chem.* Dove Press; 2017; **10**: 11–28.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
76. Ikram A, Zaheer T, Awan FM, et al.: **Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches.** *Sci Rep.* 2018; **8**(1): 16107.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
77. Adhikari UK, Tayebi M, Rahman MM: **Immunoinformatics Approach for Epitope-Based Peptide Vaccine Design and Active Site Prediction against Polyprotein of Emerging Oropouche Virus.** *J Immunol Res.* Hindawi; 2018; **2018**: 6718083.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
78. Ali M, Pandey RK, Khatoon N, et al.: **Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection.** *Sci Rep.* 2017; **7**(1): 9232.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
79. Callaway E: **The race for coronavirus vaccines: a graphical guide.** *Nature.* 2020; **580**(7805): 576–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
80. Cooper NR, Nemerow GR: **The Role of Antibody and Complement in the Control of Viral Infections.** *J Invest Dermatol.* 1984; **83**(1, Supplement): S121–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
81. Bacchetta R, Gregori S, Roncarolo MG: **CD4+ regulatory T cells: Mechanisms of induction and effector function.** *Autoimmun Rev.* 2005; **4**(8): 491–6.  
[PubMed Abstract](#) | [Publisher Full Text](#)
82. Garcia KC, Teyton L, Wilson IA: **Structural Basis of T Cell Recognition.** *Annu Rev Immunol.* 1999; **17**: 369–97.  
[PubMed Abstract](#) | [Publisher Full Text](#)
83. Dodd KA, McElroy AK, Jones MEB, et al.: **Rift Valley fever virus clearance and protection from neurologic disease are dependent on CD4+ T cell and virus-specific antibody responses.** *J Virol.* 2013; **87**(11): 6161–71.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
84. Adhikari UK, Rahman MM: **Overlapping CD8+ and CD4+ T-cell epitopes identification for the progression of epitope-based peptide vaccine from nucleocapsid and glycoprotein of emerging Rift Valley fever virus using immunoinformatics approach.** *Infect Genet Evol.* 2017; **56**: 75–91.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
85. Shamriz S, Ofoghi H, Moazami N: **Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application.** *Comput Biol Med.* 2016; **76**: 24–9.  
[PubMed Abstract](#) | [Publisher Full Text](#)
86. Meza B, Ascencio F, Sierra-Beltrán AP, et al.: **A novel design of a multi-antigenic, multistage and multi-epitope vaccine against Helicobacter pylori: An in silico approach.** *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis.* 2017; **49**: 309–17.  
[PubMed Abstract](#) | [Publisher Full Text](#)
87. Lee S, Nguyen MT: **Recent Advances of Vaccine Adjuvants for Infectious Diseases.** *Immune Netw.* 2015; **15**(2): 51–7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
88. Bonam SR, Partidos CD, Halmuthur SKM, et al.: **An Overview of Novel Adjuvants Designed for Improving Vaccine Efficacy.** *Trends Pharmacol Sci.* 2017; **38**(9): 771–93.  
[PubMed Abstract](#) | [Publisher Full Text](#)
89. Gori A, Longhi R, Peri C, et al.: **Peptides for immunological purposes: design, strategies and applications.** *Amino Acids.* 2013; **45**(2): 257–68.  
[PubMed Abstract](#) | [Publisher Full Text](#)
90. Chen R: **Bacterial expression systems for recombinant protein production: *E. coli* and beyond.** *Biotechnol Adv.* 2012; **30**(5): 1102–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
91. Rosano GL, Ceccarelli EA: **Recombinant protein expression in Escherichia coli: advances and challenges.** *Front Microbiol.* 2014; **5**: 172.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
92. Schmaljohn CS, Parker MD, Ennis WH, et al.: **Baculovirus expression of the M genome segment of Rift Valley fever virus and examination of antigenic and immunogenic properties of the expressed proteins.** *Virology.* 1989; **170**(1): 184–92.  
[PubMed Abstract](#) | [Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status: ? ?

---

## Version 1

Reviewer Report 28 January 2021

<https://doi.org/10.5256/f1000research.27552.r77290>

© 2021 Brun A et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Sandra Moreno

National Institute for Agriculture and Food Research and Technology (INIA), Madrid, Spain

Alejandro Brun

Animal Health Research Centre (CISA), National Institute for Agriculture and Food Research and Technology (INIA), Madrid, Spain

In this article, Maruf Ahmed Bhuiyan and colleagues use an exhaustive in silico approach to design a polypeptide sequence that includes an array of epitopes predicted that should theoretically induce both cell and antibody mediated immune responses in humans. In addition, they complete the study approach by adding predictions on populations immune responses based on HLA distributions, allergenicity responses, molecular docking to Toll like receptor 4, as well a simulation of immune responses upon vaccination. The authors conclude that this could be a promising vaccine candidate against RVF. As they mention, further investigations both in vitro and in vivo are warranted to ascertain its true potential to fight against RVF.

In our opinion the tools provided in this work constitutes an ideal in silico procedure. It remains however to be investigated further. For example, if this theoretical protein once expressed in a *E. coli* system can be produced in sufficient amounts to be easily purified.

On the other hand and, unfortunately, the article lacks novelty since it is a replication of other highly similar articles published by the same authors dealing with Lassa fever and, more recently, with SARS-CoV-2. Again and sadly, no data is yet available confirming that this methodology could help aiding vaccine design, since no in vivo experimental data has been yet provided.

Not all the results are clearly presented, they make references to some computational scores and values that are not well explained. I would like to see some more references to other peptide vaccines for RVF and other similar viruses.

The study is very interesting but not as much as it could be if after modeling the vaccine, they had generated the vaccine and carried out *in vivo* tests.

Some of their conclusions seem too categorical since all their data is based on informatics modeling (e.g. "there is good coverage in the majority of regions", "These findings suggest the development of immune memory and, therefore, immunity against the virus"). To draw the proposed conclusions they should carry some experimental assays. Otherwise the authors should indicate the limitations of this approach.

**Specific comments****Abstract:**

Background:

1. "Any form of direct contact with body fluids, blood or tissue of infected animals is the mode of transmission of this pathogen." It should be included mosquito bites as a possible mode of transmission.
2. "Despite being an emerging virus, no proper vaccinations are yet available for the public". I would rather say no proper vaccines (rather than vaccinations) are available for human use or prophylaxis

Results:

1. "A glycoprotein with the highest antigenicity was selected. "Please explain whether Gn or Gc was selected. What is the criteria used to define high antigenicity?
2. "We selected T cell epitopes based on their immunological potencies" What is the meaning for immunological potency?
3. "The chimera protein was modelled, refined, and validated prior to docking against toll-like receptor 4." Why against TLR-4? This is a membrane receptor and RVFV should be preferentially detected by endosomal TLRs. Explain if carbohydrate binding was considered for selecting TLR-4 or include references if available.
4. "Finally, codon optimization and cloning conducted with Escherichia coli K12 suggests high translation efficiency" I do not understand this sentence. Would you mean that codon optimization was performed to increase translation efficiency?

Conclusion:

1. "We believe that our designed multiepitope vaccine is a promising prophylactic candidate against RVF pathogenesis." This is a belief not supported by experimental in vivo testing. I would rather say this bioinformatics approach represents a strategy for rationally design of potential RVF immunogens.

**Introduction**

1. "RVF is an arthropod-borne zoonotic infectious viral disease caused by RVF" Probably this sentence here is redundant.
2. ".....and has a viral genome encoding four proteins:" This is only partially true. At least two additional proteins can be encoded by ribosomal frameshifts, including NSm and a 78kDa protein corresponding to NSm-Gn polypeptide. Please modify accordingly.
3. "Although the non-structural proteins facilitate RVF to survive inside its host by inhibiting first-line immunogenic responses, its glycoproteins are essential and highly crucial for invasion, entry and viral replication inside the host cell" This is probably not accurate. Please rephrase. The role of glycoproteins in viral replication after entry has not been described.
4. ".....to find the highest antigenic glycoprotein to predict T and B cell epitopes using a computational approach....." Should you mean "the highest antigenic glycoprotein T and B cell epitopes predicted using a computational approach".

5. "cytotoxic T-lymphocyte (CTL), helper T-lymphocyte (HTL) and linear B-lymphocyte (LBL) epitopes predicted" What is a linear B lymphocyte (LBL)? Should you mean "cytotoxic T-lymphocyte (CTL), helper T-lymphocyte (HTL) and B-lymphocyte (BL) linear epitopes predicted"

**Methods**

1. "TLR4 can recognize viral glycoprotein". The provided reference does not support this claim.  
Revise for accuracy

**Results**

1. "The glycoprotein sequence was selected for further analysis." Explain why the NSs protein was excluded from further analysis in spite of being more antigenic (table 1) and clarify which glycoprotein they refer to when selecting the different epitopes. (Note that RVFV encodes two glycoproteins upon posttranslational processing.)
2. The data shown in tables 2-4 is confusing. They shortlist different amount of epitopes in each table, refer to others in the text and then they only use some to assemble the vaccine construct.
3. I think it would be useful to see more details in the inserted multiepitope in the map of plasmid construct (fig 9).
4. The units of the *in silico* immune response lack precision, sometimes units are not indicated. The antibody titers analyzed have a confusing scale and the immunization protocol is not clear. It would be interesting to know if these antibody levels could be related to a specific *in vitro* test.

**Is the work clearly and accurately presented and does it cite the current literature?**

Partly

**Is the study design appropriate and is the work technically sound?**

No

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

I cannot comment. A qualified statistician is required.

**Are all the source data underlying the results available to ensure full reproducibility?**

Yes

**Are the conclusions drawn adequately supported by the results?**

No

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Virology, Immune responses, Vaccines, Rift Valley fever virus

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.**

Reviewer Report 14 January 2021

<https://doi.org/10.5256/f1000research.27552.r75911>

© 2021 Emran T et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **Firzan Nainu** 

Faculty of Pharmacy, Hasanuddin University, Makassar, Sulawesi Selatan-90245, Indonesia

**Trina E. Tallei** 

Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Sam Ratulangi, Manado, North Sulawesi-95115, Indonesia

**Talha Bin Emran** 

Department of Pharmacy, BGC Trust University Bangladesh, Chittagong, Bangladesh

The author appraised this paper by predicting epitope- and a peptide-based vaccine against glycoprotein of Rift Valley fever virus using immunoinformatics approaches. This manuscript has limited novelty, and most of the results presented by them are already published.

However, your article is inadequately presented. Furthermore, there are many grammatical mistakes and spelling mistakes, as well.

Although the article has scientific rigor, several major flaws need to be corrected before indexing.

There are several limitations to this study listed in the comprehensive comments to the authors hereinafter. The presentation of findings and data interpretation in this manuscript could have been advanced compared to this existing design.

#### Major comments:

#### General Comments

1. The authors just have written several issues haphazardly. Many sentences/information throughout the manuscript have serious flaws that withdrawn my attention from it.
2. Many non-scientific and incorrect/wrong information/sentences are there, which may mislead the readers.
3. Every section of the manuscript must be written scientifically according to the published literature with appropriate references.

4. The authors need to improve their writing style. The whole manuscript needs to undergo an English language edit.
5. The current manuscript will add no new information/recommendation for the readers.
6. The work seems preliminary. The study problem is not apparent.
7. Spacing, punctuation marks, grammar, and spelling errors should be reviewed wholly.
8. The research questions could be defined more precisely. This is probably due to language concerns.
9. If the same work was performed on the SARS-CoV-2 (specific nCoV-2019), this would be wonderful.

**Title:**

1. The present title of the article is verbose. Authors need to shorten the title precisely for a clearer understanding.
2. Also, the current title indicates a consequence (e.g., immune epitopes and peptide vaccine design), which could be written as a statement instead.

**Abstract:**

1. The purpose and significance of this research requisite to be explained in the abstract more clearly.
2. The abstract section is unsuitable—no focus point in the abstract section.
3. Change the sentence "To identify immunodominant epitopes and design a potent vaccine candidate, we applied a series of immunoinformatic approaches with molecular dynamics and immune response simulation frameworks". What does the series of immunoinformatic approaches mans?
4. Increase the sentences in the conclusion section.

**Introduction:**

1. The introduction section is inapplicable. Need to change the introduction considerably. Try to include the existing research limitations also, how the present research unravels those limits.
2. The study's gaps should be clearly defined in the introduction section with the applicable references.
3. The flow of the introduction is not perfect and unspecific. My advice is to make the sentences more lucid and legible for more productive comprehension.
4. In the introduction section, there are lots of redundant sentences repeating in the whole paper, making me feel the paper like a "cut and paste" from several other resources. For example, "RVFV epidemics have killed hundreds of thousands of animals, more than a thousand humans, and caused significant economic losses." The sentence is not a way of scientific description. I suggest deleting sentences like these (there are more throughout

the paper).

5. The paragraphs are not logically arranged; there are unnecessary repeats. Some of the explanation should be marked on their first appearance (for example, These reports unanimously concluded that coordinated efforts are needed to prepare for preventive measures against the recurrent emergence of RVFV), this makes me feel the authors didn't read through their paper after they finished it, and the abrupt terminologies make the paper like a Plagiarism.
6. Arrange the sentences (no meanings): "Epidemic alarms have persuaded several national and international health organizations to issue cautions about the rising risk of infection in Rift Valley fever (RVF) uninfected countries, like Europe and USA, due to the existence of vectors of transmission which are highly permissive, further compounded by global animal trade."
7. There should be at least a brief discussion about peptide vaccines in epidemic situations in the introduction and references to any such vaccines proposed or underdevelopment for Rift Valley fever virus published materials, conference reports, etc.

**Materials and Methods:**

1. In silico prediction of immunogenic epitopes is unlikely fitting with the wet-lab data. Experimental validation of the predicted epitopes is highly required to recommend any predicted epitope as a potential vaccine.
2. The authors should need to add a workflow as a pictorial form, not the text form. It will help to comprehend easily for readers.
3. Many of the tools used are not cutting-edge or represent the best available tools. For example, PSIPRED is not competitive in protein secondary structures. Authors are suggested to utilize the Rosetta or several other tools to predict protein 3D structures. The binding energy derived from it would be meaningful.
4. Some tools used do not represent state of the art, and hence, the quality and confidence of the results may be limited.
5. The experimental section is poorly designed for "Molecular docking between TLR4 and vaccine". They could use them to analyze the active site instead of other servers that do not have proper validation or cross-checking with other tools.
6. In the method sections, the author's elaborate details regarding the software/website they have used—which need to describe briefly.
7. To get the hit or lead through virtual screening along with current finding, I strongly recommend to the authors to calculate the Binding Free Energy Calculation and molecular dynamics (MD) simulation (at least for 200 ns) to investigate the dynamic stability, the mechanism of binding, the behavior of structural conformation. These general strategies will support to get the best hit in terms of interaction, their stability, and to explore the mechanism of action. Authors also need to describe and illustrate how selectively their antigens will interact with the target to avoid nonspecific binding.

8. Need references for the following sentences "Determining the stability and firmness of the protein-receptor docked structure by molecular dynamics is essential for carrying out an in silico study" also for "The NMA tool surfaced the extent of motion of the vaccine-receptor complex in terms of covariance and ability to deform; along with eigenvalues and B-factor results".
9. The author here mentioned the identification of an active site by using a web server. However, the author needs to add some previous experimental data regarding the literature review to consolidate the viral proteome sequences by comparing them with the published articles.
10. Major weakness of the manuscript is the lack of controls (both positive and negative) for selected peptides/sequences. The authors only presented data for selected peptides, which makes score output hard to understand.
11. A 1-2 sentence explanation of why some method was selected, how it works, and what the score means will make the manuscript easier to follow.
12. Finally, the output of this work should be a list of potential epitopes to be tested experimentally. A pleasant addition would be a comparison with previously published findings.
13. Several published works predict potential peptides for CTL, HTL, and LBL epitopes. It seems that those studies do not support the authors' proposal because NTKCRLSGTALIRAG, STAHEVVPF, or PNDYQSAHYLNN has not been nominated as a candidate peptide.
14. This is because it is hard to judge the algorithm adopted in the current study feasible or not.
15. It would add to the article's value if an MD (molecular dynamics at least 200ns) is included.

**Results:**

1. The topic is interesting; however, the study is not solid enough to deserve indexing as it lacks any experimental proof. In fact, in the absence of any experimental work, this study represents just a hypothesis. The authors themselves keep repeating, "this work needs experimental validation"!
2. Result section is poorly designed. Nothing was explained, although very little explanation can be found in the discussion section. Why do they want to design a peptide vaccine? The antigenicity and allergenicity evaluation of the vaccine construct they design lacks explanation.
3. The Vaxijen, which was developed over ten years ago, has been recently benchmarked against other methods (Dalsass *et al.*, *Front. Immunol.* 2019). The recently published Vaxign-ML (Ong *et al.*, *Bioinformatics*, 2020) work was reported to out-perform other prediction methods, including Vaxijen. The Vaxign-ML was extended to predict viral vaccine candidate prediction, specifically the whole SARS-CoV-2 proteome, and the SGP was predicted to have the highest score (Ong *et al.*, *bioRxiv*, 2020). The authors should carefully compare their

method and prediction results to the previous works (mainly focus on SARS-CoV-2; e.g., Abdelmageed *et al.*, *Biomed Res Int.*, 2020, etc.).

4. I suggest the authors discuss the results obtained from the "Estimated population coverage" more in-depth.
5. What is the need for "Disulfide bridging in the vaccine structure" and "In silico immune responses"?
6. It is not clear whether the current study's statistical methods are valid and correctly applied.
7. It is better to as a review to a bioinformatician, instead of molecular biologists.
8. Some of the linear B cell epitopes might be too short to be considered. The authors could filter out epitopes based on length (e.g., seven amino acids).
9. Limited novelty, and most of the results presented by them are already published.

**Discussion:**

1. Many text repetitions found in the discussion section. It is highly recommended to emphasize findings and assumptions that support or disagree with other work(s).
2. Addendum, repetition of the results should be avoided in the discussion section.
3. A sound discussion includes principal, relationship, and generalizations supported by the results.
4. The discussion section should include a summary of why peptide/epitope vaccines are not licensed yet for human use.
5. The author should discuss more docking simulation studies in the discussion section.
6. In the docking part, more texts are needed regarding the interaction between the residues and the receptor.

**Conclusions:**

1. Novelty of the work should be supplemented by the author (in the conclusion section).
2. This section should be supported by the results/insights. Conclusively, it will confer a distinct idea of the study.
3. The conclusion needs to address future perspectives.

**Figures:**

1. Figure legends are self-explanatory. But the repetition of the results and discussion should be avoided in the figure legends.
2. The resolution of figures can be improved.

3. The figures lack precision, units, and axis are not well-labeled. The units are not standardized (and incorrect in some cases). Most of the figures do not have the proper resolution.

**References:**

1. For better understanding, I feel this manuscript needs more detailed background information and precise explanation of their study in the Introduction and Discussion section.
2. Several published articles related to SARS-CoV-2 must be included within the relevant text part of the manuscript.

**Is the work clearly and accurately presented and does it cite the current literature?**

Partly

**Is the study design appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

Not applicable

**Are all the source data underlying the results available to ensure full reproducibility?**

Partly

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Bioinformatics

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.**

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**