

Immunoinformatics based designing a multi-epitope vaccine against pathogenic *Chandipura vesiculovirus*

Debashrito Deb¹ | Srijita Basak¹ | Tamalika Kar¹ | Utkarsh Narsaria¹ | Filippo Castiglione² | Abhirup Paul¹ | Ashutosh Pandey³ | Anurag P. Srivastava¹ 

¹Department of Life Sciences, Garden City University, Bangalore, Karnataka, India

²Institute for Applied Computing, National Research Council of Italy, Via dei Taurini, Rome, Italy

³Plant Metabolic Engineering, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi, India

Correspondence

Ashutosh Pandey, Plant Metabolic Engineering, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi, India.

Email: ashutosh@nipgr.ac.in

Anurag P. Srivastava, Department of Life Sciences, Garden City University, Bangalore 560067, Karnataka, India.
 Email: anuittk@gmail.com

Abstract

Chandipura vesiculovirus (CHPV) is a rapidly emerging pathogen responsible for causing acute encephalitis. Due to its widespread occurrence in Asian and African countries, this has become a global threat, and there is an urgent need to design an effective and nonallergenic vaccine against this pathogen. The present study aimed to develop a multi-epitope vaccine using an immunoinformatics approach. The conventional method of vaccine design involves large proteins or whole organism which leads to unnecessary antigenic load with increased chances of allergenic reactions. In addition, the process is also very time-consuming and labor-intensive. These limitations can be overcome by peptide-based vaccines comprising short immunogenic peptide fragments that can elicit highly targeted immune responses, avoiding the chances of allergenic reactions, in a relatively shorter time span. The multi-epitope vaccine constructed using CTL, HTL, and IFN- γ epitopes was able to elicit specific immune responses when exposed to the pathogen, *in silico*. Not only that, molecular docking and molecular dynamics simulation studies confirmed a stable interaction of the vaccine with the immune receptors. Several physicochemical analyses of the designed vaccine candidate confirmed it to be highly immunogenic and nonallergic. The computer-aided analysis performed in this study suggests that the designed multi-epitope vaccine can elicit specific immune responses and can be a potential candidate against CHPV.

KEY WORDS

Chandipura vesiculovirus, immune simulation, molecular docking, molecular dynamic simulation, multi-epitope

1 | INTRODUCTION

Chandipura vesiculovirus (CHPV), associated with an encephalitic illness in humans, is a member of the Rhabdoviridae family, first discovered by Bhatt and

Rodrigues.¹ It was reported from two febrile cases during an outbreak of Dengue and Chikungunya in Nagpur, Maharashtra, India.¹ Though maximum cases are reported in India,^{2,3} studies have also shown its presence in other Asian³ and African countries.^{3,4} It also infects

Debashrito Deb, Srijita Basak, Tamalika Kar, and Utkarsh Narsaria contributed equally to this study.

many mammalian species around the globe.⁵ Due to its worldwide occurrence, it has gained global attention as an emerging neurotropic pathogen, imposing a mortality rate ranging from 55% to 77%.^{2,6} The complete genome sequencing makes it evident that the uniqueness of the virus has become a new threat to humanity and what seems to be a tropical outbreak can very soon be a global one.⁷

The viral genome is ~11 kb consisting of single-stranded RNA coding for five polypeptides namely, Nucleocapsid protein N, Phosphoprotein P, Matrix protein M, Glycoprotein G, and Large protein L.^{3,7} L and P proteins together act as a viral RNA-dependent RNA polymerase (RdRp) and matrix protein links the encapsidated genome RNA with the phospholipid bi-layer.⁸ Spike-like glycoprotein present in the membrane act as the major antigenic determinant.⁹ Thus, glycoprotein becomes a potent choice for vaccine designing. A recombinant vaccine against CHPV has been developed in mice using the entire G gene,¹⁰ but it lacks human trials.⁶ Hence, a licensed vaccine is still not available.⁶ Considering the global threat imposed by the virus, there is an urgent need to develop a highly effective vaccine against CHPV.

The conventional method of vaccine design involves large proteins or whole organism. The major drawback of this approach is that it leads to unnecessary antigenic load and also increases the chances of allergenic responses.¹¹ This limitation can be overcome by peptide-based vaccines comprising short immunogenic peptide fragments that can elicit highly targeted immune responses, avoiding the chances of allergenic reactions.¹¹ Recent advancements in computational biology are very effective in designing various approaches for vaccine synthesis with one of these advancements being the emerging role of immunoinformatics tools in effective vaccine designing.^{12,13} Several antigenic epitopes are used for the construction of a multi-epitope vaccine which can activate both humoral and adaptive immune responses.¹⁴ The immunoinformatic approach is time-saving as well as cost-effective and has the potential to ensure a successful vaccine design.¹⁵ The techniques used in this approach have the potential to identify candidate proteins that might be overlooked by conventional experimentation.¹⁶ Hence, in this study the immunoinformatic approach is used to develop a multi-epitope, prophylactic vaccine against CHPV consisting cytotoxic T lymphocyte (CTL), helper T lymphocyte (HTL), and interferon- γ inducing (IFN- γ) epitopes (Figure 1).

CHPV is a fast-growing pathogen that causes acute encephalitis in humans, especially children.

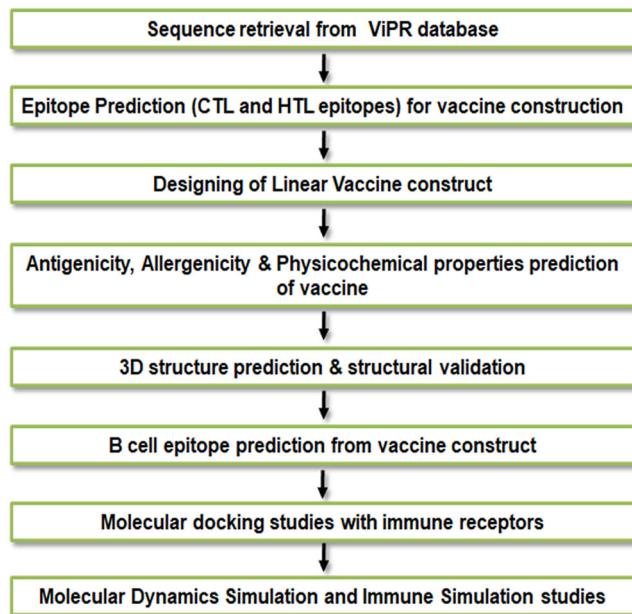


FIGURE 1 A systematic representation of all the steps followed in the designed study. CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte

This study utilized the immunoinformatics approach to construct a novel and effective multi-epitope vaccine, which can deliver both humoral and cell-mediated immune response. After the application of various immune filters and a range of immunoinformatics techniques, the final vaccine was constructed with 6 CTL, 7 HTL, and 3 IFN- γ epitopes. The vaccine was predicted to be immunogenic and nonallergenic. The study also presented an *in silico* model of the three-dimensional structure of the engineered protein that has the capacity to induce immune responses. The molecular docking analysis confirmed strong interaction of the vaccine with immune cell receptors (TLR4/TLR2/TLR1), indicating the possibility of generation of immune response once the vaccine interacts with them. In addition, the molecular dynamics simulation (MDS) studies suggested the vaccine to be stable under varied physiological condition. To check the efficacy of the vaccine candidate an *in silico* immune simulation was performed which gave a clear indication of the fact that the vaccine candidate has the ability to generate a strong immune response when introduced into the body. A control simulation was also performed which consisted of an injection of the live virus at time zero. The results indicated that though a naive immune response is elicited, it is not sufficient to eradicate the virus in absence of vaccination.

2 | MATERIALS AND METHODS

2.1 | Selection and retrieval of target protein

The VIPR database (<https://www.viprbrc.org/brc/home.sp?decorator=vipr>) was used to retrieve the glycoprotein sequences of the viruses belonging to the genus *Vesiculovirus* of the Rhabdoviridae family to investigate their evolutionary relationship. These amino acid sequences were aligned implementing the MUSCLE program in MEGA X.¹⁷ The alignment file obtained was further used for the phylogenetic analysis. The phylogenetic tree was constructed using Maximum likelihood algorithm with bootstrap resampling of 1000 replicates in MEGA X.¹⁸ The evolutionary history was inferred by using the Maximum Likelihood method and JTT (Jones-Taylor-Thornton) matrix-based model. The initial tree for the heuristic search was obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using JTT model, and then selecting the topology with superior log likelihood value. In addition, Clustal Omega was used for checking the sequence similarity between the CHPV glycoproteins obtained from different strains. Clustal Omega is one of the most widely used package for carrying out multiple sequence alignment.¹⁹ It is also fast enough to make very large alignments, and the accuracy of protein alignments is high when compared to other alternative packages available for the same purpose.^{19,20} Not only Clustal Omega is most widely used, fast and accurate but also consumes least RAM making it computationally less demanding.²⁰ Since, using Clustal Omega one can perform fast and accurate MSAs with the consumption of very little memory in the system's processor, it was the desired software chosen for sequence alignment in this study.

2.2 | Homology modeling, refinement, and validation of glycoprotein

The structure of the glycoprotein was modeled using Raptor X server (<http://raptordx.uchicago.edu/>), a homology modeling tool. Using GalaxyRefine server (<http://galaxy.seoklab.org/refine>), which is based on CASP10 tested refinement method, the distortions in the structures recovered after homology modeling were reduced, and refined models were obtained. The best-refined structure of the glycoprotein was validated by Ramachandran plot analysis using RAMPAGE (<http://mordred.bioc.cam.ac.uk/%7Errapper/rampage>.

php), Z-score generated by ProSA (<https://prosa.services.came.sbg.ac.at/prosa.php>), and ERRAT score (<https://servicesn.mbi.ucla.edu/ERRAT/>).

2.3 | Analysis of the physicochemical properties of the target proteins

The ExPASy PortParam server (<https://web.expasy.org/protparam/>) was used for the physicochemical analysis of the refined structure of the target glycoprotein. In addition, VaxiJen v2.0 server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) was used to determine the antigenicity of the glycoprotein. In addition, AA Prop server was also used to support the findings of ExPASy ProtParam server (<http://www.biogem.org/tool/aa-prop>).

2.4 | T-cell epitope prediction

2.4.1 | CTL epitopes prediction

The NetCTL1.2 server (<http://www.cbs.dtu.dk/services/NetCTL/>) predicted 9-mer long T cell epitopes identified in the human population by the frequently occurring HLA Class I supertypes, that is, A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62.²¹ In NetCTL1.2 server, the thresholds for distinctive parameters such as proteasomal C-terminal cleavage, Transporter Associated with Antigen Processing (TAP) transport capacity, and epitope recognition were set at 0.15, 0.05, and 0.75, respectively. NetCTL provides prediction of epitopes with sensitivity of 54%–89% and specificity of 94%–99%. Furthermore, the epitopes identified by other HLA Class I alleles were also detected by Immune Epitope Consensus (IEDB) tool (<http://tools.iedb.org/mhci/>).

2.4.2 | HTL epitopes prediction

Net MHC II pan 3.2 server (<http://www.cbs.dtu.dk/services/NetMHCIIpan/>) was used to predict the 15-mer long T cell epitopes, which have been identified by HLA Class II alleles. The peptides had been classified as strong, intermediate, and non-binders based on the idea of percentile rank as given by Net MHC II pan 3.2 server. The threshold value for strong, weak, and nonbinding epitopes were set at 2, 10, and more than 10%, respectively. NetMHC II pan 3.2 server predicts peptide binding on human HLA-DR, HLA-DQ, and HLA-DP alleles using Artificial Neural Network (ANN) system.

2.5 | B-cell epitope prediction

Linear and discontinuous/conformational epitopes were predicted using the ElliPro tool (<http://tools.iedb.org/ellipro/>) from IEDB server with default parameters.

2.5.1 | Prediction of IFN- γ epitopes

IFN- γ plays a key role in provoking antitumour, anti-viral, and immune regulation activities. Thus, it is important to identify epitopes with the capacity to induce IFN- γ to model an efficient multi-epitope vaccine. IFNepitope server (<http://crdd.osdd.net/raghava/ifnepitope/>) was used to predict out the IFN- γ epitopes from the target protein.²² The server functionality is focused on a data set which includes IFN- γ inducing and non-inducing MHC class II binders. Predictions were made by the use of a combination of various approaches, such as machine learning strategy, motive-based analysis, and accuracy hybrid approach having maximum accuracy of 81.39%.²²

2.5.2 | Population coverage

The population coverage of the screened-out epitopes based on HLA genotypic frequencies was calculated using IEDB population coverage analysis (<http://tools.iedb.org/population/>) (Table S7).

2.6 | Multi epitope vaccine construct, structural modeling, refinement, and validation

An immunological adjuvant, Cholera Toxin subunit B (CTB) was added to the N-terminal of the vaccine construct by EAAAK linker as it has the ability to induce regulatory responses. It shows affinity to monosialotetrahexosylganglioside (GM1) which is distributed in all immune cells.²³ The screened CTL, HTL, and IFN- γ epitopes were linked together by glycine-proline-rich GPGPG linkers. 3D structure of the linear construct was generated using trRosetta (<https://yanglab.nankai.edu.cn/trRosetta/>) which produced a final model after refinement. The model was then evaluated using ERRAT score, Verify3D score (<https://services.mbi.ucla.edu/Verify3D/>), RAMPAGE for Ramachandran Plot analysis, and ProSA generated Z-score for determining the overall quality of the vaccine construct.

2.7 | Antigenicity, allergenicity, and physicochemical properties of the vaccine construct

The antigenicity of the final vaccine construct was checked by VaxiJen v2.0²⁴ and ANTIGENpro (<http://scratch.proteomics.ics.uci.edu/>).²⁵ VaxiJen v2.0, an antigen-free alignment method, focuses on auto-cross covariance (ACC) transformation of protein sequences into standardized vectors with key amino acid properties. This server uses viral databases to derive predictive models where each data set constituted of 100 known and 100 non-antigenic antigens. Internal leave-one-out cross-validation and external validation using data sets is used to check the generated models. These models behave well in both validations with predictive accuracy varying from 70% to 89% (Dyotchinova and Flower 2007). ANTIGENpro is an alignment-free, sequence-based, and pathogen-independent server which utilizes protein antigenicity microarray data to predict the antigenicity. It correctly classifies 82% of the known protective antigens when trained using only the protein microarray datasets. The accuracy on the combined data set is estimated at 76% by cross-validation experiments.²⁵ AllerTOP v2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>)²⁶ and AllergenFP (<http://ddg-pharmfac.net/AllergenFP/>)²⁷ were used to determine the allergenicity of the vaccine construct. AllerTOP server utilizes ACC sorting of protein sequences into standardized vectors of equal length. This was extended to analyses of peptides of various lengths of quantitative structure–activity relations (QSAR). AllerTOP uses K-nearest neighbour algorithm (kNN, $k = 1$) to classify the proteins based on a training set consisting 2427 recognized allergens from different species and 2427 non-allergens.²⁷ AllergenFP server uses five E-descriptors which defines the amino acids in the protein sequences into data sets, and converts the strings into standardized vectors by ACC transformation. The E-descriptors of the amino acids were extracted from the key component analysis of a data matrix comprising of 237 physicochemical properties. AllergenFP categorizes a protein as an allergen or non-allergen based on the maximum Tanimoto coefficient, ratio of the intersecting ratio set to the union set as a similarity indicator.²⁷ Isoelectric point, molecular weight, and other parameters like half-life, instability index, aliphatic index, and GRAVY score of the vaccine construct was calculated from ExPASy server (<https://web.expasy.org/protparam/>). TMHMM server v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) and SignalP4.1 (<http://www.cbs.dtu.dk/services/SignalP-4.1/>) were used to search any potential transmembrane helices or any peptide signals in the final vaccine construct.

2.8 | Modeling of epitopes with HLA alleles

PEPFOLD 3 server (<https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/>) was used for generating the 3D structures of the epitopes used in the final vaccine construct. X-ray crystallographic structure of two predominant allele HLA-DRB1*15:01 (HLA class II) and HLA-A*02:01 (HLA class I allele) were retrieved from Protein Data Bank (PDB) using PDB IDs 1BX2 and 1QEW, respectively. Then the HTL epitopes were docked against HLA-DRB1*15:01 and the CTL epitopes were docked against HLA-A*02:01 (HLA Class I allele) using HADDOCK 2.4. (<https://wenmr.science.uu.nl/haddock2.4/>).

2.9 | Docking of TLR 1, TLR 2, and TLR 4 with the vaccine

Molecular docking is one of the most important tools to understand the protein-protein interaction. For the generation of an appropriate immune response, interaction of the antigenic molecule or vaccine with the target immune cell receptor is necessary. Hence, the binding pattern of TLR1, TLR2, and TLR4 with the multi-epitope vaccine was analysed as they localize at the cell surface and gets activated.^{28,29} Both TLR1 and TLR2 structures were obtained from the protein data bank (ID 2Z7X). Similarly, PDB ID 3FXI was used to obtain the TLR4/MD2 heterotetramer structure from protein data bank.

Molecular docking analysis of the multi-epitope vaccine with TLR1, 2, and 4 was carried out using HADDOCK 2.4 (<http://www.bonvinlab.org/software/haddock2.4/>). Easy interface level was used to carry out the docking analysis. Based on the lowest HADDOCK score, docked clusters were formed, and the best cluster was identified. The best representative structure from the clusters was subjected to molecular refinement using HADDOCK Refinement Interface. Active and passive residues for the interaction were predicted by CPoRT (<http://alcazar.science.uu.nl/services/CPoRT/>). PyMOL (Schrodinger) was used for visualization of the result. PDBsum (<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>) was used to map the interacting residues between the vaccine and the TLRs.

2.10 | Energy minimization and MDS

2.10.1 | Vaccine construct

GROMACS (GROningen MAchine for Chemical Simulations) a command-line Linux-based program was used

for the MDS and energy minimization.³⁰ The vaccine structure was subjected to MDS to mimic the experimental conditions. In this study, the physical conditions like the temperature and the pressure were mimicked using the canonical ensemble, NVT and the isobaric and isothermal ensemble, NPT. The temperature used for the simulation was 300 K and the pressure used was of 1 bar to mimic the experimental conditions. The topology file was generated using OPLS-AA (Optimized Potential for Liquid Simulation-All Atom) force field constraint for energy minimization and equilibration. The structure was placed at a distance of 1 nm from the cube's edge which was filled with water molecules to generate its periodic image 2 nm apart. An equilibrated three-point water model, spc216 was used as the solvent to simulate the vaccine with periodic boundary conditions. The net charge of the vaccine construct was evaluated, and the system was neutralized by the addition of charged ions. The energy minimization process was performed, and the energy minimized structure was obtained. The NVT and NPT equilibration was conducted for 1000 picoseconds (ps) consisting of 500 000 steps to stabilize the temperature and pressure of the system. A 50 ns MDS run was performed for the energy minimized structure to find the root mean square deviation (RMSD) of backbone and root mean square fluctuation (RMSF) of side chain. The radius of gyration was plotted to check the compactness of the protein structure. Xmgrace, a linux-based software, was used to visualize the graphs generated from the simulation.³¹

2.10.2 | Vaccine-TLR4 complex

The energy minimization process of the vaccine-TLR4 complex was also performed similar to that of the vaccine construct and accordingly, the energy minimized structure was obtained. The NVT and NPT equilibration was performed for 1000 ps comprising of 500 000 steps for stabilizing the temperature and pressure of the system. A MDS run of 50 ns of the energy minimized structure was performed to find the RMSD of backbone and RMSF of side chain.

2.11 | Reverse translation and codon optimization and in silico cloning

Reverse translation and codon optimization was performed by the Java Codon Adaptation Tool (JCat) (<http://www.jcat.de/>) to attain the reverse translated complementary DNA (cDNA) sequence.³² This cDNA is optimized for the vaccine to express in *Escherichia*

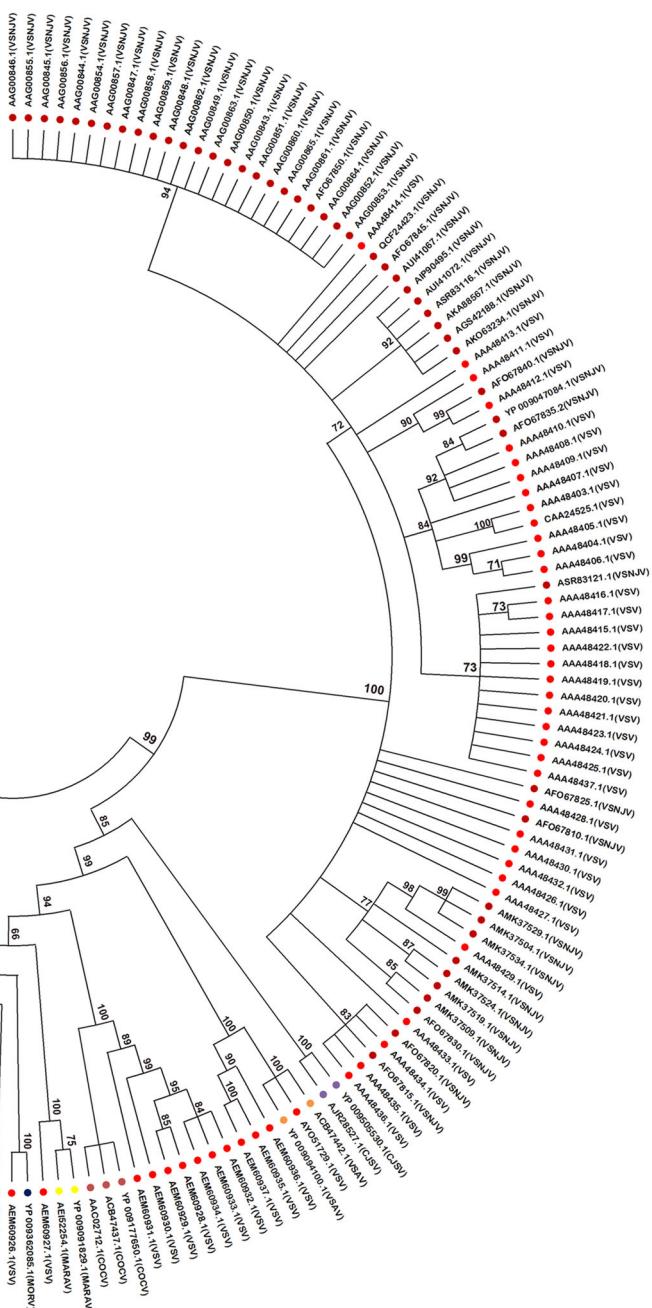


FIGURE 2 Phylogenetic analysis of the *vesiculovirus* glycoprotein sequences. A total of 203 glycoprotein sequences, taken from different species of *vesiculoviruses*, were used to generate the phylogenetic tree using the Maximum likelihood algorithm of MEGA X, with default parameters and 1000 bootstrap replicates. The CHPV glycoproteins are marked with a blue arc. CHPV, *Chandipura vesiculovirus*

coli K-12 strain. The results include GC content and the codon adaptation index (CAI). CAI score is ideal when it is 1, though scores more than 0.8 is also considerable.³³ GC content of the vaccine should range from 30% to 70%, which indicates good translation and transcription efficiencies in *E. coli*.³² Eventually, the SnapGene tool was used to insert the optimized multi-epitope vaccine sequence into the pET-28a (+) vector.

2.12 | Immune simulation

To further characterize the chimeric peptide's immunogenicity and immune response profile, the C-ImmSim server (<https://www.iac.cnr.it/%7Efilippo/projects/c-immsim-online.html>) was used to perform in silico immune simulations.³⁴ C-ImmSim is an agent-based model for immune response prediction that uses position-specific scoring matrices (PSSM) derived from

machine learning techniques for predicting immune interactions.³⁴ For most of the vaccines currently in use, the minimum recommended time between the dose 1 and dose 2 is 4 weeks.³⁵ The entire simulation ran for 1400 time steps which are about 15 months (a time step is about 8 h). Twelve peptide injections were given four weeks apart at time step 10, 94, 178, 262, 346, 430, 514, 598, 682, 766, 850, 934. Then a live virus was injected at time step 1100, which is about 12 months after the simulation starts.

3 | RESULTS

3.1 | Selection of virulent protein

The VIPR database was utilized for the retrieval of the amino acid sequence of the target glycoprotein of CHPV (Accession number—YP_007641380.1). The phylogenetic tree was constructed using all the glycoproteins of different viruses of the genus Vesiculovirus belonging to Rhabdoviridae family (Figure 2). The CHPV glycoproteins clustered together in a single clade (Figure 2). A phylogeny tree was also constructed using all the glycoproteins of different strains of CHPV (Figure S1). The multiple sequence alignment by Clustal Omega generated a percentage similarity table which showed that all the glycoproteins from different strains of CHPV have high percentage of similarity amongst each other, the highest being 100% and lowest being 88.55%. The accession numbers and the percentage similarity table of all the strains of CHPV glycoprotein are shown in Table S1. In addition, another multiple sequence alignment was performed using MUSCLE suite for the CHPV glycoproteins obtained from different strains. The results from MUSCLE alignment also indicated high percentage similarity between the glycoproteins of different strains of CHPV, which was also suggested by Clustal Omega. In this case, the highest and lowest percentage similarity observed between the glycoproteins was 100% and 95.92%, respectively. Hence, the results from Clustal Omega and MUSCLE are both similar and indicative of the same fact that the protein sequences are highly conserved with high percentage identity (Supporting Information File S2).

The target glycoprotein was also checked for human proteome hit recognition to assess the risk of autoimmune reactions. The database reference proteins (refseq_protein) using BLASTp (protein–protein BLAST) were used for the identification of hits in the human proteome, which confirmed the glycoprotein showed no significant similarity with any of the proteins in human.

3.2 | Homology modeling, refinement, and validation of the glycoprotein

Since there is no complete structure of Chandipura glycoprotein involving entire 530 amino acids and the PDB structure available for the glycoprotein of this virus involves only 419 amino acids,³⁶ some of our predicted epitopes couldn't be located in the same. Therefore, a 3D model was constructed using the entire glycoprotein of CHPV by Raptor X homology modeling tool (Figure 3), and the 3D model was used to locate the epitopes selected in this study (Figure 4). GalaxyRefine server predicted five refined models of the glycoprotein from which Model 5 was selected (Table S2). This model was chosen based on different scores obtained from GalaxyRefine server like, RMSD, clash score, poor rotamers score, and Ramachandran plot analysis. Model 5 had a low RMSD of 0.393 (lower the RMSD better is the structure), Clash score of 10.6 (lower the value of clash score, better is the structure), poor rotamers score of 0.2 (lower the score for poor rotamers, better is the structure) and 98.7% of residues in Ramachandran plot favored region (a structure having more than 90% of the residues in the favored region indicates the structure to be of good quality). Hence, the results for Model 5 indicated it to be the best model when compared to the other refined models and was chosen for further analyses.

In the chosen model, the favored regions in the Ramachandran plot were 98.7% (Figure S2), ERRAT score was 89.03 (Figure S2) and Z-score was -6.07 (Figure S2). The molecular weight of the glycoprotein was found to be 59.03 kDa by ExPASy ProtParam server.

3.3 | Prediction of CTL and HTL epitopes

For the prediction of CTL epitopes, the NetCTL 1.2 and IEDB consensus methods were used (Tables 1 and S3), whereas the HTL epitopes were predicted using NetMHC



FIGURE 3 Modeled structure of CHPV glycoprotein with sequence length of 530 amino acids as predicted by RaptorX. CHPV, *Chandipura vesiculovirus*

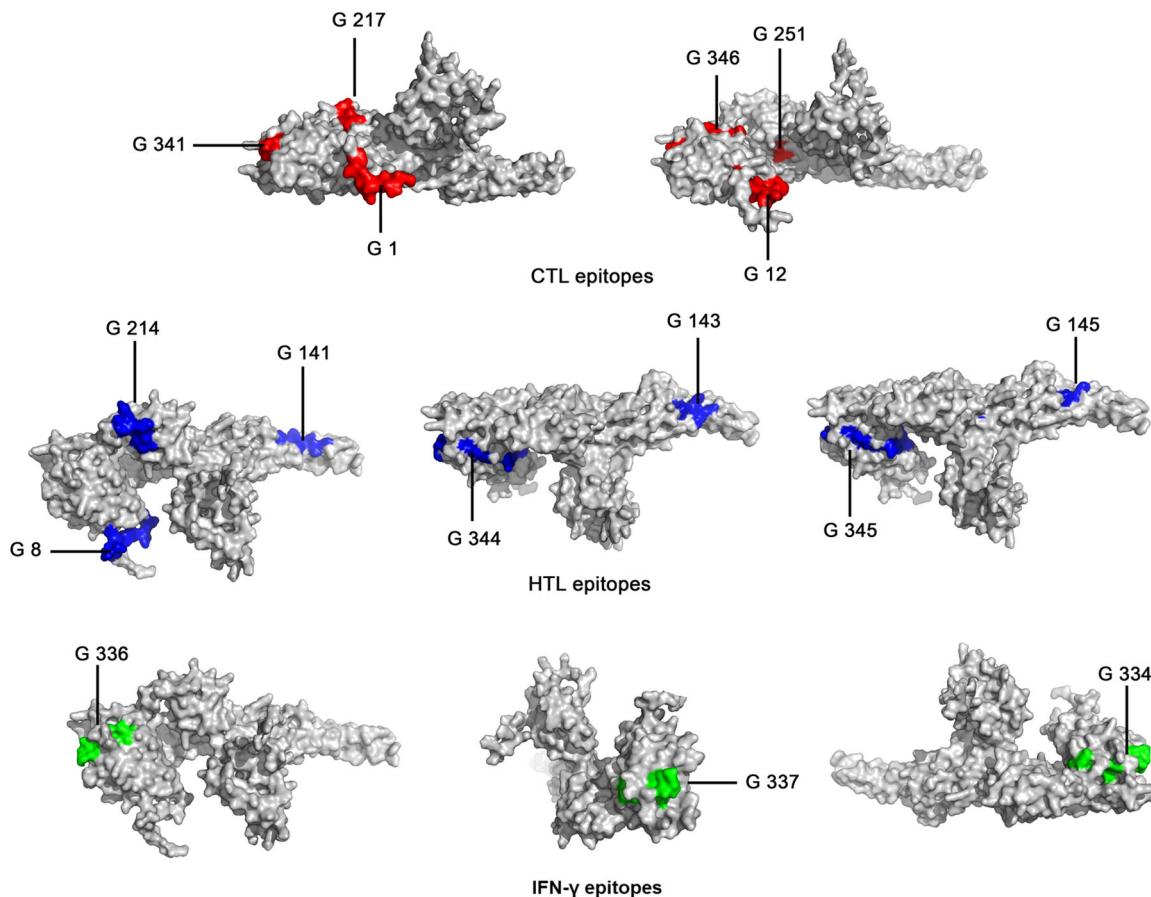


FIGURE 4 Glycoprotein 3D models showing the surface position of the epitopes included in the multi-epitope vaccine construct. CTL epitopes marked by red color, HTL epitopes are marked by blue color and IFN- γ epitopes are marked by green color. 3D, three-dimensional; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte; IFN- γ , interferon- γ

II pan 3.2 server (Tables 2 and S4). In addition, the identified epitopes that were expected to have high binding affinities with HLA class I and class II alleles were subjected to specific immune filters to screen out the strongest possible epitopes. The parameters for selecting the best possible epitopes were: (i) epitopes should be immunogenic, (ii) should be promiscuous, (iii) should be antigenic, and (iv) should have high population coverage. This study led to the identification of promiscuous and overlapping epitopes (Table S5) based on all the above-mentioned criteria. The overlapping epitopes were antigenic, as predicted using VaxiJen v2.0 server.

3.4 | Multi epitope vaccine construct, structural properties, and refinement

The final multi-epitope vaccine was designed by combining 6 CTL, 7 HTL (Tables 1 and 2), and 3 IFN- γ (Table S6) epitopes via GPGPG linkers that prevented the chance of any junctional epitope formation.³⁷ To the final vaccine construct, the CTB adjuvant was linked at the

N-terminal of the vaccine with EAAAK linker (Figure 5A). The criteria for the incorporation of the epitopes in the multi-epitope vaccine construct were: (i) they should have overlapping HTL and CTL epitopes (Table S5), (ii) should be immunogenic, (iii) should have high affinity to HLA alleles, (iv) they should be promiscuous, and (v) should have high population coverage. The final vaccine construct thus designed, consisted of 388 amino acids. The 3D model of the final vaccine construct was made using trRosetta webserver (Figure 5B). Since all of the models were already refined by trRosetta, further refinement of the models was not required. All the models predicted by trRosetta were subjected to various structural analyses to select the best-suited model. The final model chosen revealed a Z score of -7.69 that was within the range of scores of comparable-sized proteins (Figure 5D). Ramachandran plot analysis showed 93.5%, 3.9%, and 2.6% residues in favored, allowed, and outlier regions, respectively (Figure 5E), which verified the overall quality of the final multi-epitope vaccine construct. Verify3D predicted that 81.44% of the amino acids scored ≥ 0.2 , which further

TABLE 1 CTL epitopes showing promiscuity were predicted using NetCTL 1.2

CTL binding with HLA Class I						
Epitope	Position	Supertype	Allele Class I	Consensus percentile rank	IC ₅₀	Antigenicity
TLSFAHTRY	346	A1, A3, B62, B58	HLA-A*30:02	1.135	201.82	1.6573
			HLA-A*01:01	1.9	2279.76	
MTSSVTISV	1	A1, A2	HLA-A*68:02	0.1	1.75	0.6600
			HLA-A*02:01	0.5	33.86	
			HLA-B*58:01	0.87	393.97	
			HLA-A*02:06	0.965	17.55	
			HLA-A*02:03	1.28	18.39	
			HLA-A*01:01	1.5	2595.97	
LISFIAPSY	12	A1, A2, B58, B62	HLA-A*30:02	0.215	20.17	0.5292
			HLA-B*15:01	0.9	131.06	
			HLA-A*32:01	1.8	1886.46	
			HLA-A*01:01	1.8	3309.76	
			HLA-B*35:01	1.9	48.97	
TVINGTLSF	341	A26, B7, B58, B62	HLA-A*26:01	0.165	22.66	0.8432
			HLA-B*15:01	0.3	16.96	
			HLA-A*32:01	0.4	40.36	
			HLA-A*23:01	0.62	210.0	
			HLA-B*35:01	1.6	42.36	
			HLA-B*58:01	2.0	1601.19	
EIAAGAIVF	217	A26, B62	HLA-A*26:01	0.4	385.11	0.9015
			HLA-B*15:01	1.3	154.75	
			HLA-B*35:01	1.7	68.68	
FPNGEWVSL	251	B7, B39	HLA-B*07:02	0.6	43.26	0.7309
			HLA-B*35:01	0.7	14.14	
			HLA-B*53:01	0.7	48.62	
			HLA-B*51:01	1.8	2825.12	

Note: Epitopes with IC₅₀ value <500 nm are considered as good binders towards specific alleles. Antigenicity scores were predicted using VaxiJen v2.0 keeping a threshold of 0.4.

supported the high-quality structure of the vaccine model (Figure 5C). ERRAT of the already refined vaccine construct projected a score of 59.5819 which further verifies the overall quality of the vaccine construct (Figure S3).

3.5 | Physicochemical properties of the final vaccine construct

The linear construct of the multi-epitope vaccine was predicted to be antigenic by VaxiJen v2.0 and AntigenPro with a score of 0.6334 and 0.857163, respectively. It was found to be nonallergenic as predicted by AllerTOP and

AllergenFP. Evaluation of the physicochemical properties predicted by ExPASy (Supporting Information Material SM1) showed that the multi-epitope vaccine construct has a molecular weight of around 40 kDa. The calculated instability index was found to be 13.11 (i.e., <40) which implies that the vaccine is stable. The theoretical pI and aliphatic index were found to be 8.73 and 79.95, respectively. The vaccine is projected to be thermostable at different temperatures due to high aliphatic index value. The estimated half-life of the vaccine is 30 h in mammalian reticulocytes, >20 h in yeast, and >10 h in *E. coli*. The grand average hydropathicity (GRAVY) value was recorded as 0.083, indicating the polar nature of the

Epitopes (position)	HLA Alleles II	Score	Antigenicity
SEFLVIMITPHHGV(143)	DRB1*01:01 DRB1*04:05 DRB1*04:01 DRB1*07:01 DRB1*08:02 DRB1*11:01 DRB1*12:01 DRB1*15:01 DRB4*01:01 DRB5*01:01	0.40 0.07 0.17 1.20 0.03 0.80 0.05 0.08 0.04 0.90	0.8996
TDSEFLVIMITPHHV(141)	DRB1*01:01 DRB1*04:05 DRB1*04:01 DRB1*07:01 DRB1*08:02 DRB1*11:01 DRB1*12:01 DRB1*15:01 DRB4*01:01 DRB5*01:01	1.00 0.08 0.25 1.60 0.12 1.80 0.12 0.30 0.12 1.90	0.7730
FLVIMITPHHGVDD(145)	DRB1*01:01 DRB1*04:01 DRB1*04:05 DRB1*08:02 DRB1*11:01 DRB1*12:01 DRB1*15:01 DRB4*01:01 DRB5*01:01	1.00 1.00 0.60 0.09 1.40 0.15 0.25 0.12 1.70	0.9586
SVILLISFIAPSYSS(8)	DRB1*04:05 DRB1*08:02 DRB1*12:01 DRB1*15:01 DQA1*01:01-DQB1*05:01	1.10 1.60 0.60 0.70 0.80	0.5617
GTLSFAHTRYVRMWI(345)	DPA1*02:01-DPB1*01:01 DPA1*01:03-DPB1*02:01 DPA1*01:03-DPB1*04:01 DPA1*02:01-DPB1*05:01 DRB1*07:01 DRB1*09:01	1.20 1.40 1.20 1.50 0.50 1.20	0.5170

TABLE 2 HLT epitopes showing promiscuity were predicted using NetMHC II pan 3.2 server

Epitopes (position)	HLA Alleles II	Score	Antigenicity
NGTLSFAHTRYVRMW (344)	DPA1*02:01-DPB1*01:01	1.90	0.5929
	DPA1*01:03-DPB1*04:01	2.00	
	DRB1*07:01	0.50	
	DRB1*09:01	1.20	
KTKEIAAGAIVFKSK (214)	DPA1*02:01-DPB1*14:01	1.80	1.6029
	DQA1*01:02-DQB1*06:02	0.17	
	DQA1*05:01-DQB1*03:01	0.40	

Note: Antigenicity scores were predicted using VaxiJen v2.0 keeping a threshold of 0.4.

vaccine construct and there was no signal peptide detected which hence will prevent protein localization (Figure S4). Since no transmembrane helix has been identified in the designed vaccine construct, no expression difficulties are anticipated in the vaccine development (Figure S5). In addition, AA Prop server was also used for validating the results provided by ExPASy protparam server. The results for the prediction were similar to the prediction made by ExPASy server which has been provided in Table S8.

3.6 | Prediction of B-cell epitopes

The ElliPro server was used to predict linear/continuous as well as conformational/discontinuous B-cell epitopes with the default parameters as shown in Tables 3 and 4, respectively. The predicted B-cell epitopes within the vaccine construct were visualized using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) (Figure S6). Furthermore, the IFN- γ inducing epitopes were predicted using IFNepitope server from the target glycoprotein (Table S6).

3.7 | Modeling of epitopes with HLA allele

The CTL and HTL epitopes were independently docked with commonly occurring alleles: HLA-A*02:01 (HLA Class I allele) and HLA-DRB1*15:01 (HLA Class II allele), respectively to evaluate their binding patterns (Figure 6).

3.8 | Docking of TLR1, TLR2, and TLR4 with the vaccine

The HADDOCK (High Ambiguity Driven protein-protein Docking) 2.4 server performed the molecular docking of

the energy minimized vaccine construct with TLR1, TLR2, and TLR4. In addition, ClusPro docking software was also used for further validating the docking results. The results obtained from ClusPro and HADDOCK were further compared using binding affinity analysis with the help of PRODIGY server.

3.9 | Docking with TLR4-

HADDOCK clustered 63 structures in 6 cluster(s), which represents 31.5% of the water refined HADDOCK generated models. The top cluster with the lowest HADDOCK score is the most reliable cluster of all. Therefore, a representative model from this top cluster has been subjected to refinement. The HADDOCK refining server grouped the resulting 20 structures into one cluster, accounting 100% of the water refined HADDOCK generated models. The statistics of the refined cluster is shown in Table 5A, and the details of the structural analysis are provided in Figure S9. The docked structure along with amino acid interaction is shown in Figure 7. The detailed overview and the interaction with amino acid sequences are given in Figure S10 and Supporting Information Material SM2, respectively. The secondary structure of the docked complex was predicted (Figure S8), and further Ramachandran plot and Z-score analysis were carried out for structural validation of the docked complex (Figure S7).

3.10 | Docking with TLR2-

HADDOCK clustered 147 structures in 20 cluster(s), which represents 73% of the water refined HADDOCK generated models. The top cluster with the lowest HADDOCK score is the most reliable cluster of all. Therefore, a representative model from this top cluster has been subjected to refinement. The HADDOCK

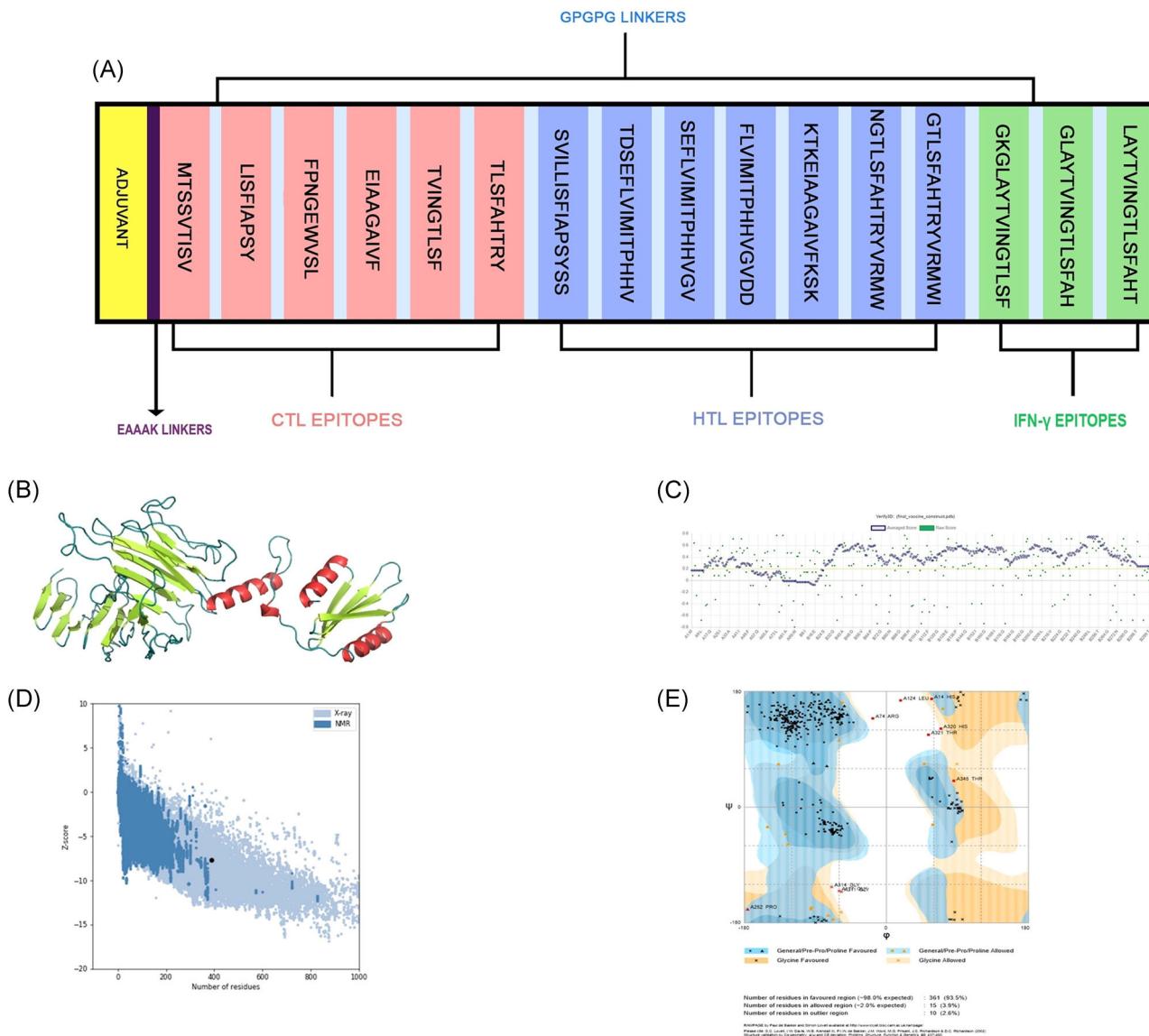


FIGURE 5 (A) Diagrammatic representation of the final multi-epitope vaccine. The pink, blue, and green boxes represent the CTL, HTL, and IFN- γ epitopes, respectively. Violet color represents the EAAAK linker and pale blue color represents GPGPG linkers. (B) Tertiary structure of the final vaccine construct. Helices, sheets, and loops are represented by red, limon, and deep teal colors, respectively. (C) Verify 3D plot of the vaccine having a score of 81.44%. (D) Z score plot of vaccine having a score of -7.69. (E) Ramachandran plot showing 93.5%, 3.9%, and 2.6% in the favored, allowed, and outlier region, respectively. CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte; IFN- γ , interferon- γ

refining server grouped the resulting 20 structures into one cluster, accounting 100% of the water refined HADDOCK generated models. The statistics of the refined cluster is shown in Table 5B, and the details of the structural analysis are provided in Figure S13. The docked structure along with amino acid interaction is shown in Figure 8. A detailed overview and the interaction with amino acid sequences are given in Figure S14 and Supporting Information Material SM3, respectively. The secondary structure of the docked complex was predicted (Figure S12), and further, Ramachandran plot and Z-score analysis were carried

out for structural validation of the docked complex (Figure S11).

3.11 | Docking with TLR1-

In case of TLR1 docking, HADDOCK clustered 161 structures in 15 cluster(s), which represents 80% of the water refined HADDOCK generated models. The top cluster with the lowest HADDOCK score is the most reliable cluster of all. Therefore, a representative model from this top cluster has been subjected to refinement.

TABLE 3 Conformational B-cell epitopes in the vaccine construct as predicted by ElliPro server

Discontinuous epitopes	Score
MT(1-2)	0.782
LCAEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGATFQVEVPGSQHIDSQQKAIERMKDTLRIAYL (9-78)	
APSYSSGPGPGTDSE(203-217)	0.726
GPGPGL(119-124)	0.694
GPGPGFFPNGEWV(147-158)	0.64
NGTLS(169-173)	
VGVGPGPGF(246-254)	0.624
DDGPGPGKTK(267-276)	
V(284)	
KSKKGPGPNGT(286-296)	
S(298)	
HTR(301-303)V(305)M(307)	0.621
GPGPGGTLTLSFAHTRYV(309-324)GPGPGKG(329-336)	
VINGTLSFGPGPGLAYTVINGTLSFAHGPGPGLAY(341-376)	
VINGTLSFAHT(378-388)	
CVWNNKT(87-93)	0.568
FGPGPG(174-179)	0.544

Note: Positions of the epitopes are indicated in the brackets.

TABLE 4 Linear B-cell epitopes in the vaccine construct as predicted by ElliPro server

Linear epitopes	Position	Score
LCAEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGATFQVEVPGSQHIDSQQKAIERMKDTLRIAY	9-77	0.797
APSYSSGPGPGTDS	203-216	0.750
DDGPGPGKTKEI	267-278	0.724
GPGPGFFPNGEWV	147-158	0.717
GPGPGL	119-124	0.694
FGPGPGLAYTVINGTLSFAHGPGPGLAYTVINGTLSFAHT	348-388	0.675
WGPGPGTLSFAHTRY	308-323	0.64
FKSKKGPGPNG	285-295	0.623
VWNNKT	88-93	0.594
GPGPGKG	329-336	0.589
GVGPGPG	247-253	0.567
GTLSFGPGPG	170-179	0.520

The HADDOCK refinement server grouped the resulting 20 structures into one cluster, accounting 100% of the water refined HADDOCK generated models. The refined cluster statistics are shown in Table 5C, and the details of the structural analysis are provided in Figure S17. The docked structure along with amino acid interaction is shown in Figure 9. A detailed overview and the

interaction with amino acid sequences are given in Figure S18 and Supporting Information Material SM4, respectively. The secondary structure of the docked complex was predicted (Figure S16), and further Ramachandran plot and Z-score analysis were carried out for structural validation of the docked complex (Figure S15).

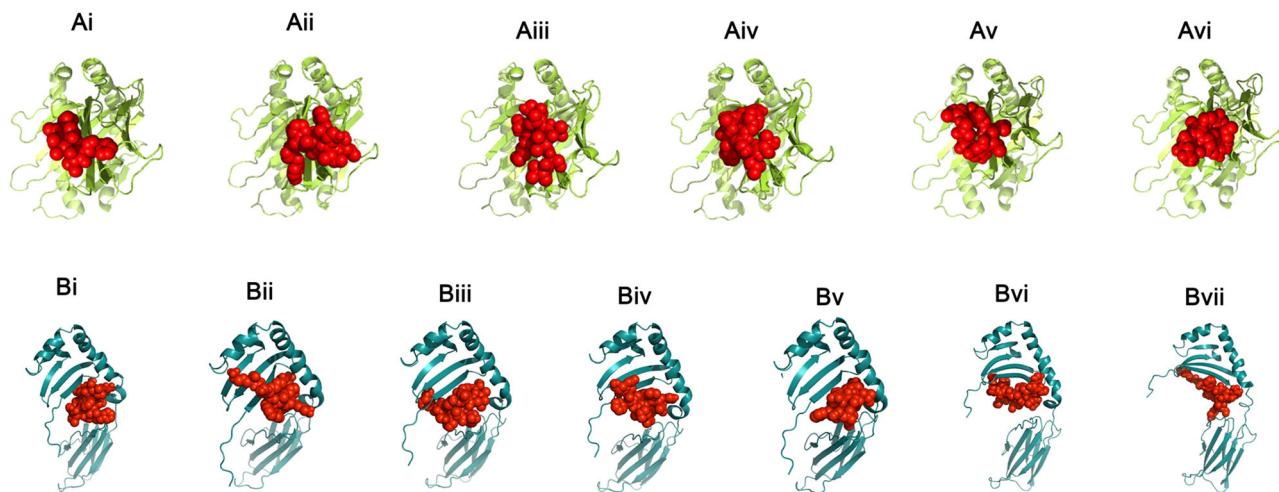


FIGURE 6 Docking of the selected epitopes included in the vaccine with HLA alleles. Ai–Avi shows the CTL epitopes docked with HLA class I allele, HLA-A*01:01 (limon), and Bi–Bvii shows the HTL epitopes docked with HLA class II allele, DRB1*15:01 (teal). All the epitopes are shown as red spheres

TABLE 5A Table showing statistics of docked vaccine and TLR4 cluster

Vaccine-TLR4	
HADDOCK score (a.u)	-168.7 ± 4.6
Cluster size	20
RMSD from the overall lowest-energy structure (Å)	0.3 ± 0.2
Van der Waals energy (kcal mol ⁻¹)	-144.6 ± 2.2
Electrostatic energy (kcal mol ⁻¹)	-236.9 ± 33.4
Desolvation energy (kcal mol ⁻¹)	23.2 ± 5.9
Restraints violation energy (kcal mol ⁻¹)	1.2 ± 0.24
Buried surface area (Å ²)	4246.9 ± 34.4

Note: Statistics are shown for the top-ranked best refined docked complex. Smaller HADDOCK score represents strong protein interaction which is expressed in arbitrary units (a.u).

3.12 | Docking using ClusPro

The vaccine construct was also docked with the TLR receptors using ClusPro to validate the findings of HADDOCK.³⁸ ClusPro gave 10 best docked structures from which the models with lowest energy scores were chosen. The scores for TLR1, TLR2, and TLR4 were predicted to be -1156.5 , -1085.7 , and -1078.6 kCal/mol, respectively (Supporting Information Material SM5–SM7). These were further compared with models obtained from the HADDOCK server, using binding affinity analysis (Table S9). The binding affinity analyses for the docked complexes obtained using HADDOCK and ClusPro showed similar energies which further verifies that the vaccine can form stable

interaction with the chosen immune cell receptors (TLR1, TLR2, and TLR4).

3.13 | Energy minimization and MDS of vaccine and vaccine-TLR4 complex

MDS is used for studying the thermodynamic properties and time-dependent phenomena.^{39,40} In this study MDS was conducted to study the stability of the vaccine and vaccine-TLR docked complex at varying thermo-baric conditions. The energy components, density, pressure, temperature, and potential energy were studied using varying parameters.

The 3D structure of the multi-epitope vaccine was subjected to MDS using GROMACS (GROningen MAchine for Chemical Simulations). The OPLS-AA force field was applied, and the vaccine construct mass was found to be 40012.04 a.m.u. “gmx solvate” command in GROMACS, was used to add 59685 solvent molecules to the system. Since the overall charge on the protein was calculated to be +3, three chloride ions were added to the system for neutralization. Energy minimization was conducted for 1979 steps where the force reached <1000 kJ/mol. The potential energy of the vaccine was evaluated during the energy minimization step. The plot for the same demonstrates nice and steady convergence of potential energy during this step (Figure S19). The energy minimized structure with a consistent low energy level of $-3.2e + 06$ kJ/mol supported it to be a reasonable starting structure for the further steps of MDS. The structure was then subjected to an equilibration phase for 1000 ps, to study its various thermodynamic properties

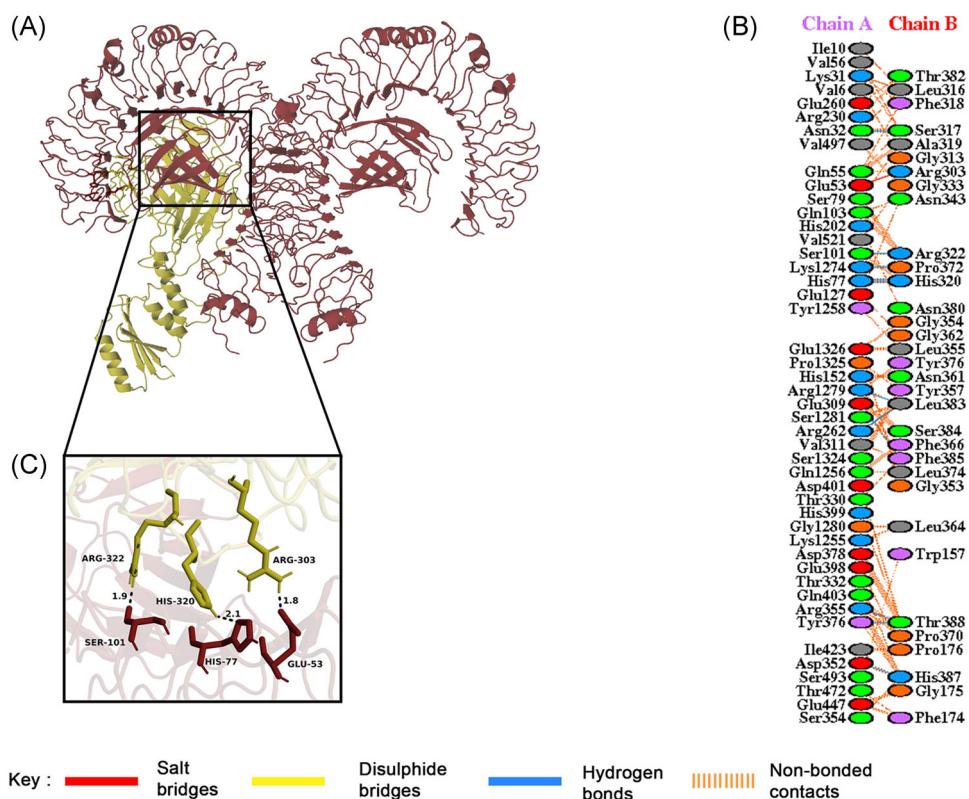


FIGURE 7 (A) Molecular docking of vaccine-TLR4/MD2 complex. Vaccine construct and TLR4 receptor are shown in yellow and red color, respectively. (B) Amino acid interactions between docked TLR4 receptor (chain A) and vaccine (chain B). (C) Overview of few hydrogen bonds within vaccine-TLR4/MD2 complex are shown

TABLE 5B Table showing statistics of docked vaccine and TLR2 cluster

Vaccine-TLR2	
HADDOCK score (a.u)	-146.8 ± 8.1
Cluster size	20
RMSD from the overall lowest-energy structure (Å)	0.3 ± 0.2
Van der Waals energy (kcal mol ⁻¹)	-65.5 ± 1.6
Electrostatic energy (kcal mol ⁻¹)	-125.2 ± 7.2
Desolvation energy (kcal mol ⁻¹)	-56.2 ± 7.6
Restraints violation energy (kcal mol ⁻¹)	0.0 ± 0.00
Buried Surface Area (Å ²)	1707.6 ± 38.3

Note: Statistics are shown for the top-ranked best refined docked complex. Smaller HADDOCK score represents strong protein interaction which is expressed in arbitrary units (a.u).

and the effect of temperature and pressure on the system. The first equilibration phase was conducted under the NVT ensemble to check if the system remained stable at the desired temperature of 300 K. The temperature of the system quickly reached 300 K and was maintained for the equilibration period with minimal fluctuations,

indicating the stability of the system at 300 K (Figure S19). The system was then subjected to a second equilibration phase under the NPT ensemble to check the stability of the system at a pressure of 1 bar. The pressure of the system was maintained at 1 bar and the negligible fluctuations observed in the plot (Figure S19) indicated that the structure also remains stable at a pressure of 1 bar. The average density of the system computed was 1007.4 kg/m³ with a total drift of 0.3 kg/m³.

Similarly, energy minimization was performed for the vaccine-TLR4 complex. The potential energy of the system was found to be $-1.1e+07$ kJ/mol (Figure S20). The consistent low energy level of the energy minimized structure supported it to be a reasonable stable structure for the next steps of the simulation. The structure was also subjected to an equilibration phase for 1000 ps to study the effect of temperature and pressure on the system. The equilibration was conducted using the same method as described above. The result for the NVT equilibration of the vaccine-TLR4 complex indicated that the temperature of the system very quickly reached 300 K and was maintained for most of the equilibration phase, indicating its stability (Figure S20). Similarly, the NPT equilibration of the vaccine-TLR4 complex suggested that

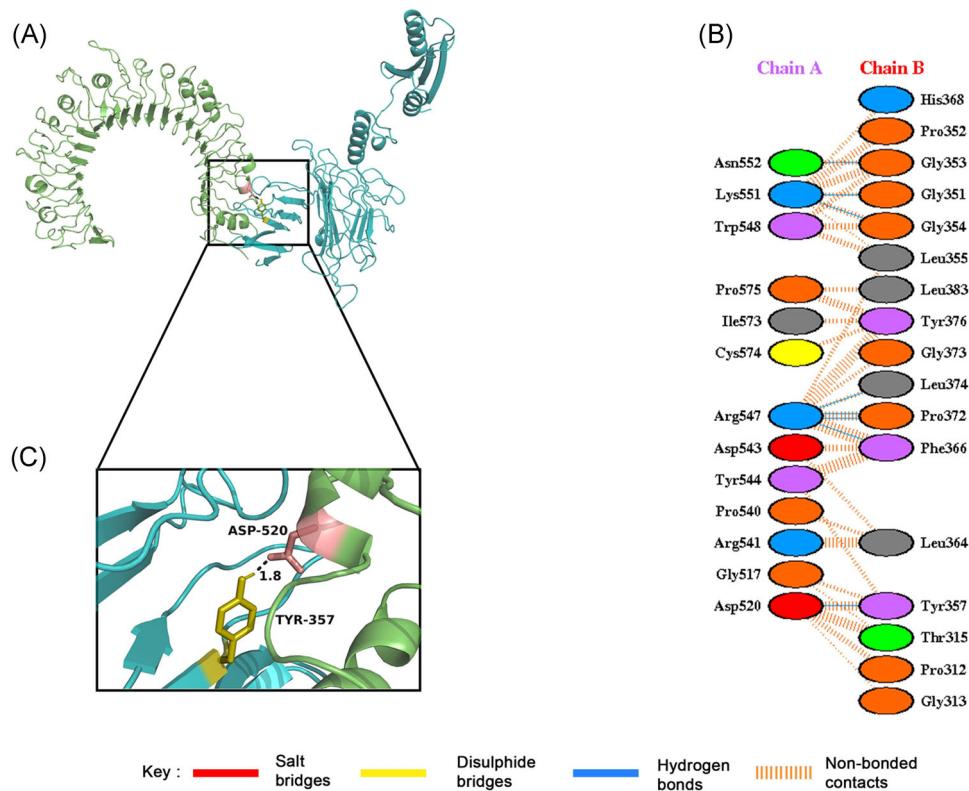


FIGURE 8 (A) Molecular docking of vaccine-TLR2 complex. Vaccine construct and TLR2 receptor are shown in blue and green color, respectively. (B) Amino acid interactions between docked TLR2 receptor (chain A) and vaccine (chain B). (C) Overview of few hydrogen bonds within vaccine-TLR2 complex are shown

TABLE 5C Table showing statistics of docked vaccine and TLR1 cluster

Vaccine-TLR1	
HADDOCK score (a.u)	-197.6 ± 3.2
Cluster size	20
RMSD from the overall lowest-energy structure (Å)	0.3 ± 0.2
Van der Waals energy (kcal mol ⁻¹)	-80.9 ± 2.7
Electrostatic energy (kcal mol ⁻¹)	-153.6 ± 13.4
Desolvation energy (kcal mol ⁻¹)	-86.0 ± 4.4
Restraints violation energy (kcal mol ⁻¹)	0.0 ± 0.00
Buried surface area (Å ²)	1998.9 ± 26.5

Note: Statistics are shown for the top-ranked best refined docked complex. Smaller HADDOCK score represents strong protein interaction which is expressed in arbitrary units (a.u).

the system remains stable at an atmospheric pressure of 1 bar which is evident from the very little fluctuations as observed in the plot (Figure S20). The computed density for the vaccine-TLR4 complex was found to be 1011.6 kg/m³ with a total drift of 0.6 kg/m³.

In the next step, a trajectory analysis was performed for 50 ns to monitor the overall stability of the vaccine and vaccine-TLR4 complex. The RMSD plot of the vaccine obtained after the simulation period of 50 ns showed successive fluctuation up to 20 ns, and then it reached a steady-state for the rest of the simulation (Figure 10A). The RMSD plot of the vaccine-TLR4 complex indicated that the complex endured fluctuations until 20 ns and then reached a steady-state which remained stable for the rest of the simulation (Figure 10C). This trend arises from the efforts of TLR4 and the vaccine molecules to get the best position relative to each other so that they can obtain the most appropriate interaction during the MDS run. The RMSD value for the vaacie-TLR4 complex was much lower when compared to the RMSD value of the vaccine alone. This is a clear indication of the fact that the vaccine and TLR4 has formed strong interactions among themselves which led to the greater stability of the complex.

RMSF was another parameter which was evaluated to check the regions with greater flexibility as well as stability. The RMSF plot of the vaccine showed that various parts of the vaccine molecule had different fluctuations (Figure 10B). In particular, Pro372, Leu383, Phe385,

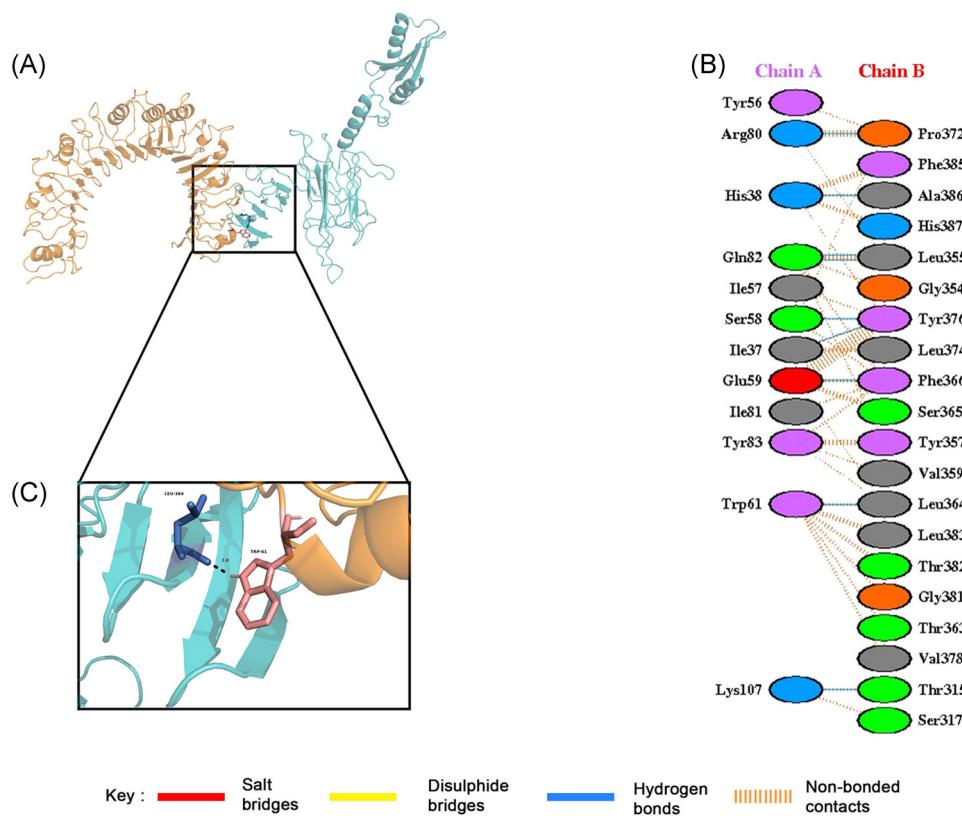


FIGURE 9 (A) Molecular docking of vaccine-TLR1 complex. Vaccine construct and TLR1 receptor are shown in blue and orange color, respectively. (B) Amino acid interactions between docked TLR1 receptor (chain A) and vaccine (chain B). (C) Overview of few hydrogen bonds within vaccine-TLR1 complex are shown

His387, and Thr388 of the vaccine showed high degree of fluctuation during the MDS run. The RMSF plot for the vaccine-TLR4 complex was almost stable with minimal fluctuation (Figure 10D). The highly fluctuating residues (Pro372, Leu383, Phe385, His387, and Thr388) obtained from the RMSF data of the vaccine showed very little fluctuation when compared to the RMSF data of the vaccine-TLR4 complex. This indicates that these residues tried to modify their interaction with the TLR4 and this modification led to the reinforcement of appropriate interactions between the vaccine and TLR4 protein, which led to the stability of the system. This result is further supported from our data obtained from studying the docking interactions between the vaccine-TLR4 complex, where the same residues of the vaccine (Pro372, Leu383, Phe385, His387, and Thr388) have formed hydrogen bonds with TLR4 protein.

The changes in the radius of gyration of vaccine and vaccine-TLR4 was monitored to evaluate the compactness of the protein structure during the MDS run of 30 ns. The plot for radius of gyration of both vaccine and vaccine-TLR4 complex indicated the compactness of the structure and it can be concluded that the intra and intermolecular interactions between the vaccine and TLR4

led to the compaction of both TLR4 and the designed vaccine (Figures S19 and S20).

3.14 | Reverse translation and codon optimization

GC content in the reverse translated cDNA was found to be 54.89%, which is admirable as it falls between expected ranges of 30%–70%.³² The CAI value was calculated to be 1, meaning the multi-epitope vaccine construct ventures high-level expression in *E. coli* K-12 strain.³³ The cDNA of the vaccine was then inserted into pET-28a (+) vector for restriction cloning (Figure 11).

3.15 | Immune simulation

C-ImmSim immune server was used for generating an in silico immune response. The generated response was compatible with the actual immune responses as shown in Figure 12. The graphs showed significantly higher secondary and tertiary responses as compared to the primary response. Elevated levels of immunoglobulin

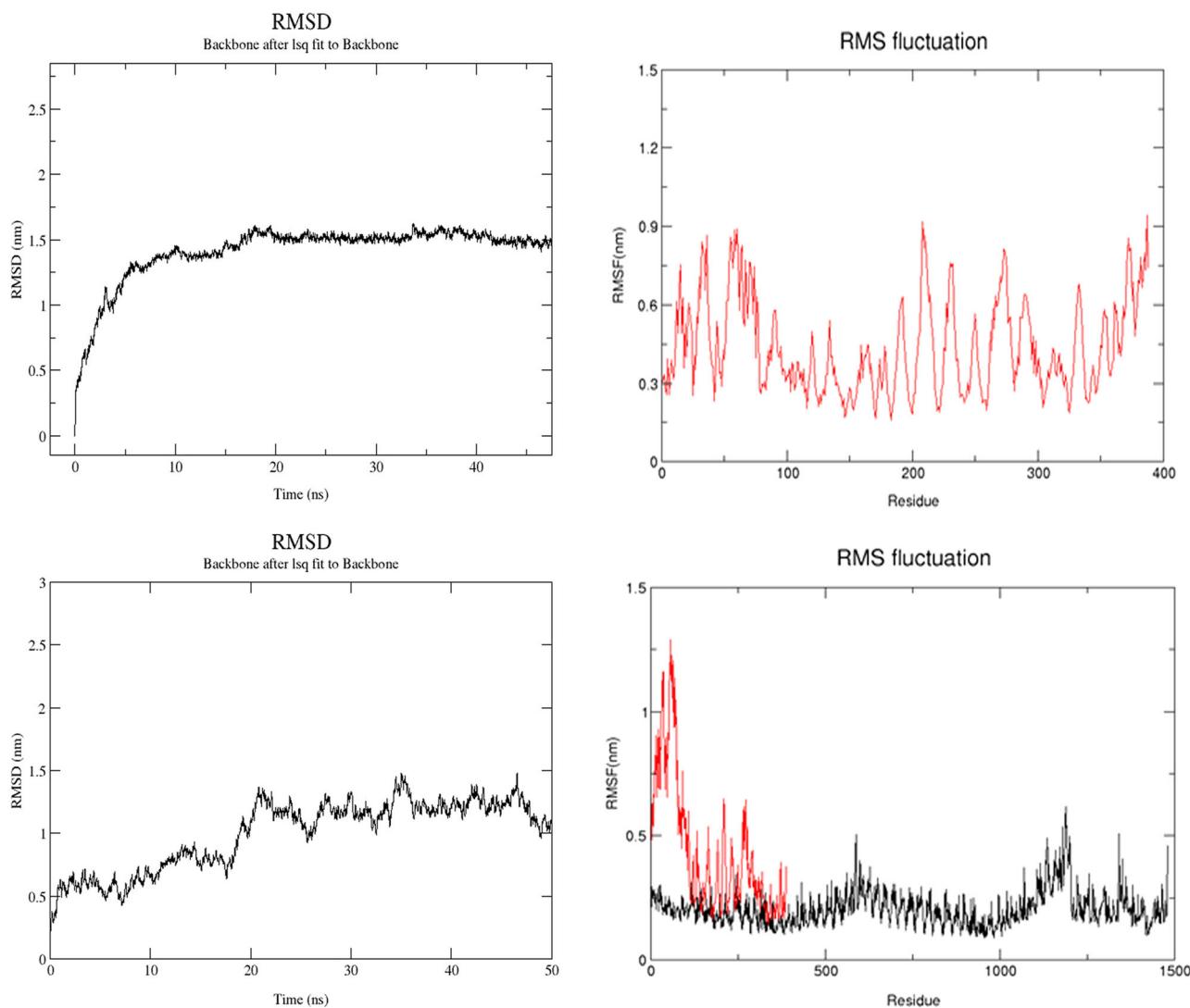


FIGURE 10 (A) RMSD plot of the vaccine backbone indicating stability. (B) RMSF plot of the vaccine (shown in red color) with high fluctuations indicating high flexibility. (C) RMSD plot of the vaccine-TLR4/MD2 complex backbone with both the chains of TLR4 along with the coreceptor MD2, indicating stability. (D) RMSF plot of vaccine-TLR4/MD2 complex with both the chains of TLR4 along with the coreceptor MD2, indicating stability. Vaccine is shown in red color and TLR4 shown in black color. RMSD, root mean square deviation; RMSF, root mean square fluctuation

activity (IgG1 + IgG2, IgM, and IgG + IgM antibodies) were observed in both secondary and tertiary responses with a subsequent decline in antigenic concentration. Additionally, multiple B-cell isotypes were found with long-lasting behavior, suggesting possible class switching and memory development (Figure 12A,B) (Figure S21). A relatively high response was observed in populations of TH (helper) and TC (cytotoxic) cells with preactivation of TCs during vaccination (Figure 12C,D) (Figure S21). Higher macrophage activity was demonstrated during exposure, while dendritic and NK cell activity was found to be consistent (Figure S21). High levels of IFN- γ and IL-2 were also evident which says that a good immune response has been generated

(Figure 12E). To check the efficacy of the vaccine, a live replicating virus was injected at around day 366. The results from the antigen graph (Figure 12A) show that when a live replicating virus is injected after vaccination, the antigenic surge is virtually absent. This is a clear indication of the fact that an effective immune response has been generated mainly because of the high concentration of specific antibodies. To check the effectiveness of the vaccination, a control simulation was also performed which consisted of an injection of the live virus at time zero. The results indicate that though a naive immune response is elicited, it is not sufficient to eradicate the virus in absence of vaccination (Figure S22).

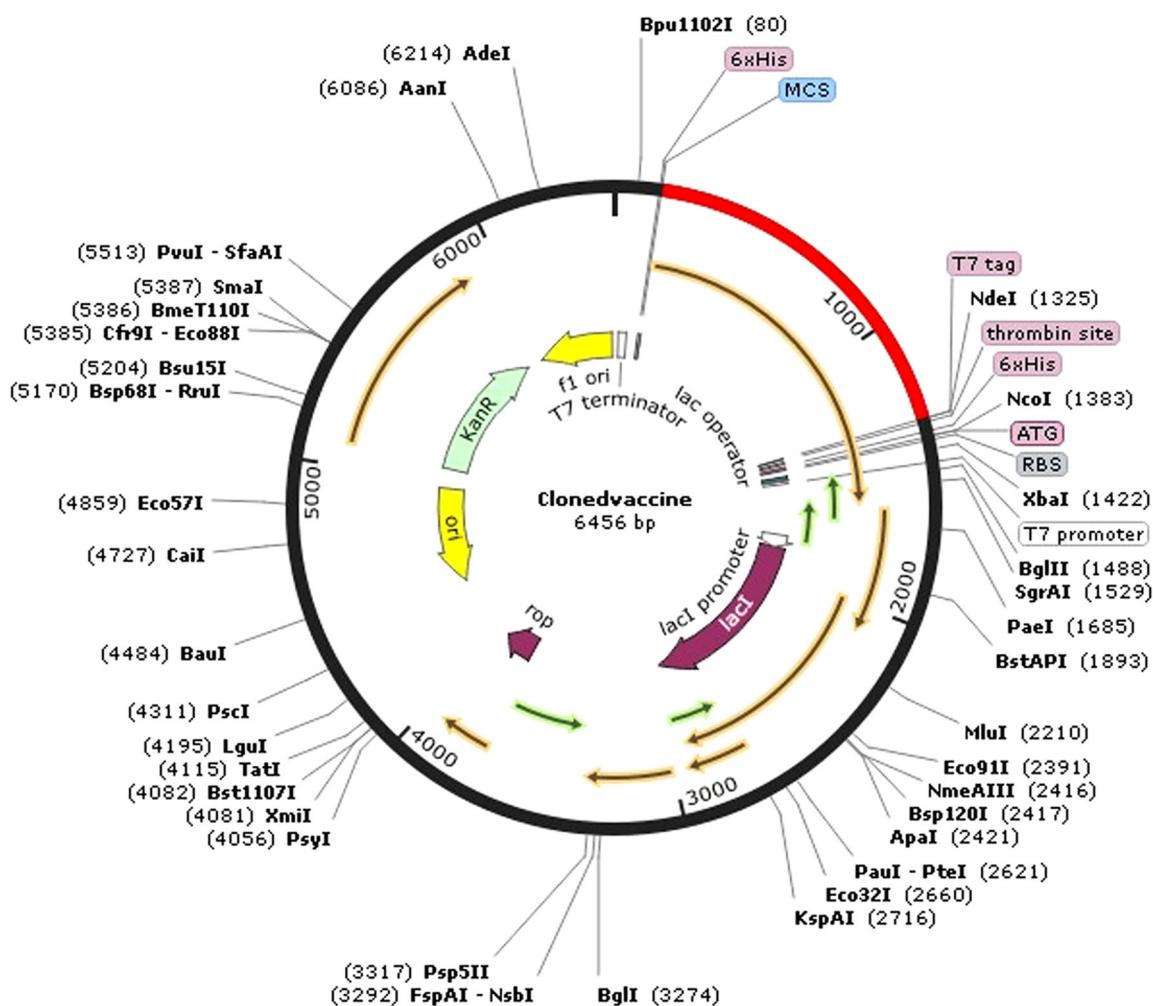


FIGURE 11 In silico restriction cloning. The red part represents the codon-optimized multi-epitope vaccine insert into the pET-28a (+) expression vector (shown in black)

4 | DISCUSSION

CHPV, which causes encephalitis, has been regarded as an emerging tropical pathogen with fatality rate of 55%–77%, predominantly affecting children of age group 2–16 years.⁴¹ There is no specific treatment available for the disease till date but a symptomatic treatment is done using mannitol to reduce the brain edema.⁴² Hence, development of an effective vaccine has become the need of the hour. A recombinant vaccine has been developed using the complete G gene of the CHPV isolate.¹⁰ A β-propio lactone (BPL) inactive tissue culture-based vaccine has also been developed.⁴³ Unfortunately, for both types of vaccines developed so far, clinical trials have been performed only in mice but not in humans.⁶ So far no vaccine is available against CHPV. Thus, aim of the present study was to design a multi-epitope, prophylactic vaccine targeting the CHPV glycoprotein which is primarily responsible for the virus envelopment, budding,

and antigenicity. Epitope-based vaccines represent a new strategy, developing immune response only against the selected epitopes, thus, avoiding the side effects of other unfavorable epitopes unlike the case of using complete antigen in recombinant vaccine technology.¹¹

Designing of epitope-driven vaccine is a modern idea, which is successfully applied in several immunological studies for the development of vaccines.⁴⁴ Similarly, the present study is centered on designing of a multi-epitope vaccine because these vaccines have advantages over traditional and single-epitope vaccines due to the following unique features: (i) TCRs from various T-cell subsets can recognize the multiple MHC Class I and Class II epitopes; (ii) CTL, HTL, and B-cell epitopes may overlap, thus, it has the capacity to induce humoral and cellular immune responses simultaneously; (iii) linking an adjuvant to the vaccine enhances the immunogenicity and provides long-lasting immune response; (iv) the likelihood of pathological immune responses or adverse

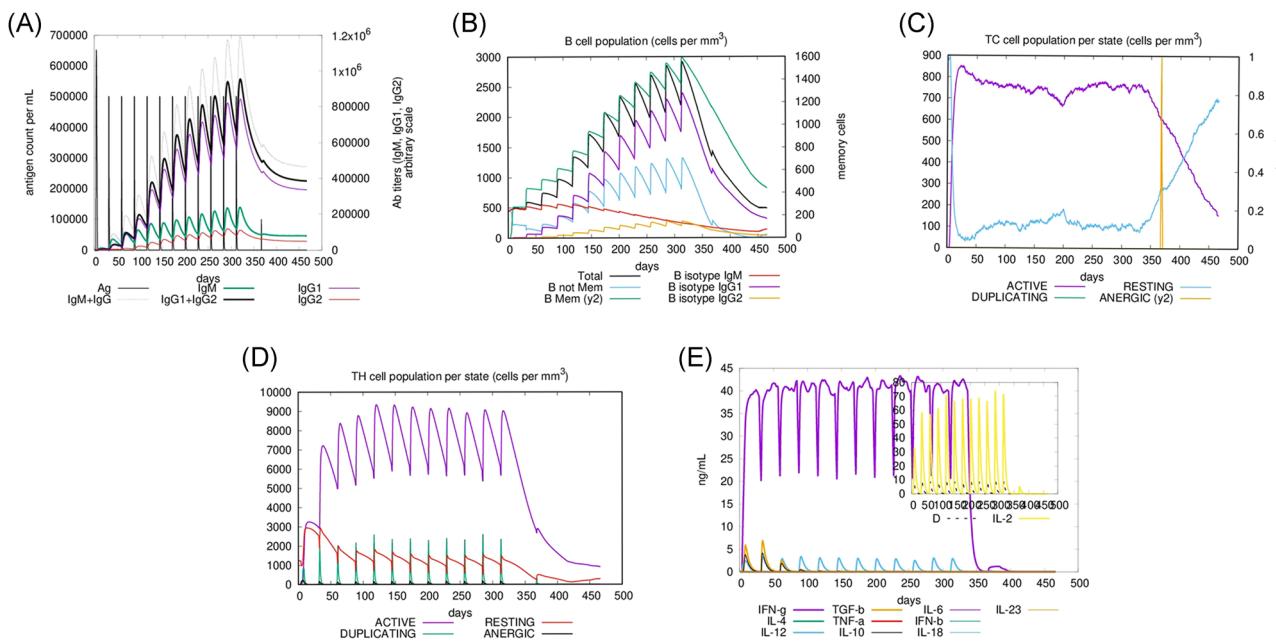


FIGURE 12 In silico simulation of immune response using vaccine as antigen. (A) Antigen and immunoglobulins. (B) B-cell population. (C) TC cell population per state. (D) TH cell population per state. (E) Production of cytokine and interleukins

effects is lowered because it is less likely to contain unwanted components.^{14,37,45–48} Further, it has been demonstrated that multi-epitope combined vaccine induce stronger CTL responses compared to those induced by a single-epitope vaccine by enhancing cellular immunity and releasing immune tolerance.⁴⁹ The cellular and humoral responses generated by the multi-epitope vaccines are highly specific with increased cytokines production.⁵⁰ A multi-epitope vaccine developed with such precautions can thus become a powerful tool in the battle against viral infections.⁵¹

One of the troubles with the conventional approach of vaccine discovery is that many of the proteins expressed during infection are not always expressed in vitro, that is, good candidate antigens might be overlooked.¹⁶ In silico methods whereas, screens for all the probable candidate antigens, as predicted by various tools and algorithms which might otherwise remain undetected.¹⁶ It is extremely important to pick the correct antigenic epitopes of the target proteins to be used in the building of multi-epitope vaccine through in silico methods.⁵² The CTL, HTL, and IFN- γ epitopes selected for the study were screened against various immunological filters (Tables 1 and 2). The applied filters were: the epitopes should be promiscuous (bind with multiple MHC class I and MHC class II alleles), must be present on the surface of the target protein, must be immunogenic as well as antigenic. The promiscuous epitopes are those with sensitivity towards several HLA alleles. These epitopes are of paramount importance in

vaccine formulation, as they are capable of developing an efficient immune response in the host, as they have affinity to several forms of HLA allele. Thus, the filtered out HLA class I and class II T-cell epitopes were further evaluated for the study of promiscuity. In the present study, the epitopes expected to have a strong binding affinity to multiple HLAs were screened out and identified as promiscuous epitopes. The overlapping CTL and HTL epitopes have the potential to trigger cytotoxic T cells and helper T cells, which in turn generate immune responses. Allergenicity is one of the key issues faced during the production of vaccines. Hence, evaluation of allergenicity is necessary at an early stage of designing the vaccine. While developing the final vaccine model, the screened out epitopes were first subjected to an allergenicity assessment followed by the whole vaccine construct. The vaccine construct designed in this study was observed to be nonallergenic. Other physicochemical features like molecular weight, instability index, theoretical pI, aliphatic index, GRAVY score, and half-life of the vaccine were also checked. The molecular weight of the vaccine was found to be 40 kDa and the instability index calculated was 13.11 which indicated that the designed vaccine is quite stable. Generally, a protein whose instability index is smaller than 40 is predicted to be stable and values above 40 predicts that the protein may be unstable.⁵³ Moreover, the vaccine has exhibited a fair percentage of solubility in overexpressed conditions. The GRAVY index of the vaccine was 0.083 indicating the vaccine's polar nature and its effective interaction with

water.⁵⁴ The high aliphatic index of 79.95 signified that the vaccine is a thermostable protein. The aliphatic index is commonly defined as the relative volume of a protein which is occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine).⁵⁵ The half-life of the vaccine is 30 h in mammalian reticulocytes, >20 h in yeast, and >10 h in *E. coli*. The half-life of a protein is defined as the time it takes for the concentration of the protein to be reduced by 50% after its synthesis in the cell. Similarly, Foroutan and his colleagues used the same array of in silico analysis to assess the allergenicity and physicochemical properties of their designed vaccine candidate against *Toxoplasma gondii*.⁵⁶ They have also performed laboratory validation of their vaccine candidate, which revealed that the multi-epitope vaccine was able to trigger strong humoral and cellular responses in mice.⁵⁶ The physicochemical properties predicted in our study were comparable to those predicted by Foroutan et al., in their recently published work.⁵⁶ In fact, the instability index, aliphatic index of our vaccine candidate was found to be better when compared to the values reported by Foroutan et al.⁵⁶ The Ramachandran Plot, ERRAT score, Verify3D score, and Z score analysis validated the overall quality of the vaccine construct (Figure 5). Thus, after rigorous in silico analysis the final vaccine construct was designed. A similar approach was used by Bazhan and his co-workers, where they have designed a T-cell multi-epitope vaccine against Ebola virus. The T-cell epitopes were predicted using IEDB—Immune Epitope Database and the vaccine candidate constructed using the suitable epitopes were found to be immunogenic when expressed in mice.⁵⁷ These studies strengthen the fact that the vaccine candidates designed *in silico* using computational tools can be a successful strategy for designing an efficient vaccine candidate against diseases.

As the CHPV glycoprotein is an envelope protein, the toll-like receptor-4 (TLR4) and toll-like receptor-2 (TLR2) expressed in the plasma membrane of the cells should primarily recognize the structural components of the virus.^{58,59} CHPV regulates TLR4, which leads to the secretion of proinflammatory cytokines and nitric oxide (NO) in monocytes-macrophage cell line of mouse.⁶⁰ In humans, TLR4 is expressed in immune cells such as monocytes, macrophages, granulocytes, and immature dendritic cells.⁶¹ Cholera toxin B (CTB) has been proven to act as a potential viral adjuvant.⁶² Activation of TLR4 by CTB is presumably due to the direct interaction between TLR4 and CTB.⁶³ This conclusion was supported by the evidence that the capacity of CTB to induce inflammatory response is lost in TLR4-deficient macrophages. CTB is able to induce NF-κB activation in TLR4 receptor cells by direct binding, which has been demonstrated using ELISA-based assays.⁶³ Further, TLR2

has been associated with the recognition of viral envelope glycoprotein.⁵⁹ The core TLR2 signaling pathway utilizes myeloid differentiation factor 88 (MyD88) as the primary adaptor, and results in NF-κB and mitogen-activated protein kinase (MAPK) activation as well as secretion of a core panel of cytokines.⁵⁹ It has also been reported that TLR2 acts as heterodimer with TLR1 for the generation of innate immune response and has been shown to recognize viral proteins.^{29,59} TLR1/TLR2 dimer generates intracellular signaling via IRAK4 mediated activation of IRAK1/2, which results in the activation of NF-κB, p38, and JNK proteins in the cytoplasm, involved in triggering innate immune response.⁶⁴

Hence, the interaction pattern of multi-epitope vaccine against TLR4/TLR2/TLR1 was analyzed using molecular docking analysis (Figures 7–9). The docking analysis of TLR4 and the vaccine construct showed that there is 1 salt bridge and 12 hydrogen bonds formed during this interaction. The docked complex showed that the salt bridge was formed between Glu53 of TLR4 and Arg303 of the vaccine. Out of the 12 hydrogen bonds, 10 bonds were formed between TLR4 and the vaccine, and remaining 2 bonds were formed between the MD-2 coreceptor protein and the vaccine. Thus, docking studies indicate that both TLR4 and MD-2 are responsible for a stable interaction of the vaccine. MDS for both the vaccine and vaccine-TLR4 complex was performed to assess the stability of the vaccine and the complex at various thermo-baric conditions. MDS results indicated that both the structures remained stable at a temperature of 300 K and a pressure of 1 bar. A trajectory analysis of 30 ns revealed that both the structures remained stable during the simulation run of 30 ns (Figure 10). However, the RMSD value obtained for the vaccine-TLR4 complex was much lower when compared to the RMSD value for the vaccine, indicating stable interaction between the vaccine and TLR4 protein. The RMSF plot for the multi-epitope vaccine showed various regions of high flexibility for the vaccine, whereas the RMSF plot for the vaccine-TLR4 complex was almost stable with very little fluctuation. The highly fluctuating residues (Pro372, Leu383, Phe385, His387, and Thr388) obtained from the RMSF data of the vaccine showed very little fluctuation when compared to the RMSF data of the vaccine-TLR4 complex. This indicated that these residues tried to modify their interaction with the TLR4 and this modification led to the reinforcement of appropriate interactions between the vaccine and TLR4 protein, which led to the stability of the system. The plots for radius of gyration also showed little variation which indicated the compactness of the protein structures due to inter and intramolecular interactions between the vaccine and

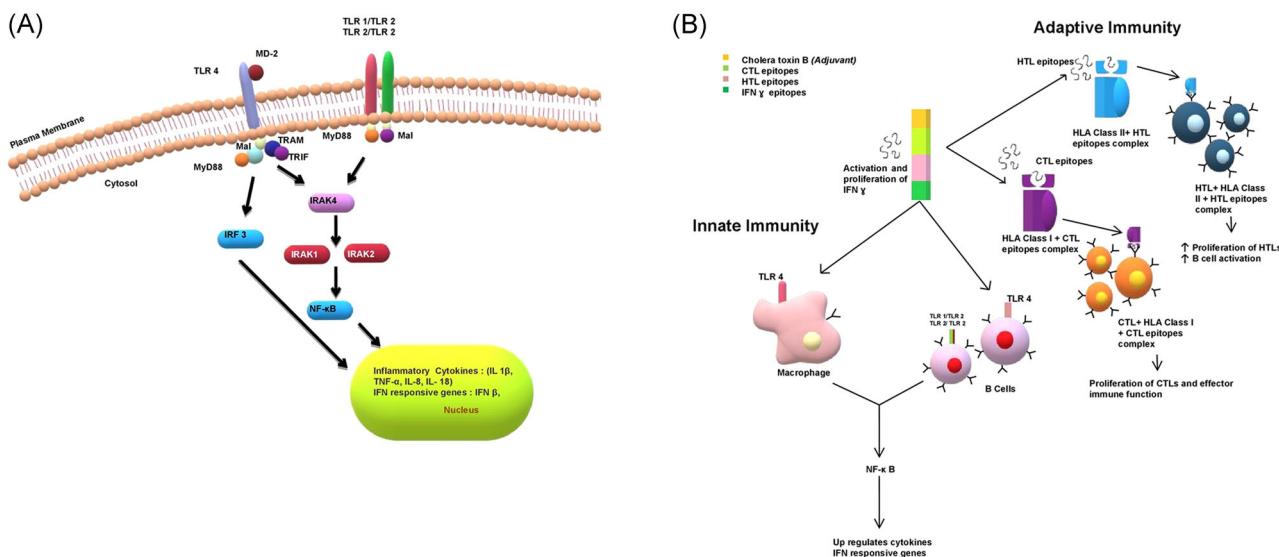


FIGURE 13 (A) TLR signal transduction pathway: TLR1/TLR2 heterodimer or TLR 2/2 homodimer utilizes MyD88 and MAL as primary adapters to activate NF- κ B that triggers inflammatory cytokine secretion. TLR4 uses four primary adapters namely MyD88, MAL, TRIF, and TRAM for NF- κ B secretion which in turn induce inflammatory cytokine secretion activating IFN pathway. (B) The CTB activates and interacts with TLR4, expressed on macrophages, B cells, and monocytes which upregulate the cytokine secretion. The other immune cells, such as NK cells, T cells, or other human monocytes, will also indirectly be stimulated by CTB. Furthermore, the CTL and HTL epitopes interact with HLA class I and HLA class II and thus form epitope-HLA complexes which in turn interact with CTLs and HTLs, activate them, and induce their proliferation. The IFN- γ will induce IFN genes. The proposed vaccine is thus capable of stimulating both adaptive and innate immunity. CTB, Cholera Toxin subunit B; CTL, cytotoxic T lymphocyte; HTL, helper T lymphocyte; IFN- γ , interferon- γ ; NF- κ B, nuclear factor- κ B

TLR4 protein. The efficient cloning and expression of such a multi-epitope vaccine in a suitable expression vector is again a very important step. Hence in the present study, an *in silico* cloning was performed to assure effective expression and translation of the designed multi-epitope vaccine construct into an expression vector: pET-28a (+) (Figure 11). Several research groups have recently applied similar strategy to design a multi-epitope vaccine against *Klebsiella pneumoniae*,⁵³ *Kaposi sarcoma*,¹¹ *Pseudomonas aeruginosa*,⁶⁵ Epstein Barr virus,⁶⁶ Malaria,⁶⁷ and Nipah virus.⁶⁸ Similar approaches have also been used for developing vaccine against cancer antigens.^{69–71} The proposed mechanism of action was also predicted for the final vaccine model (Figure 13). Since, the vaccine comprised of CTL, HTL, and IFN- γ epitopes, it could trigger the stimulation of the respective immune cells in the host, which could further activate other immune cells by complex signaling. The vaccine candidate proposed in this study can be further used for biochemical validation since it showed promising results *in silico*.

ACKNOWLEDGMENTS

The authors thank Dr. Joseph V.G., Chancellor Garden City University for his constant support to carry out this study work.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

Debashrito Deb designed and performed experiments and helped in the writing. Srijita Basak designed and performed epitope selection prediction experiment, and helped in the writing. Tamalika Kar performed the docking experiments and helped in the writing. Utkarsh Narsaria performed molecular dynamics simulation experiment and helped in the writing. Filippo Castiglione performed immune simulation experiment and helped in writing. Abhirup Paul helped in phylogenetic analysis. Ashutosh Pandey and Anurag P. Srivastava devised and supervised experiments, analyzed data, and helped in writing the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the Supporting Information Material of this paper.

ORCID

Anurag P. Srivastava <http://orcid.org/0000-0003-3493-1375>

REFERENCES

- Bhatt P, Rodrigues F. Chandipura: a new Arbovirus isolated in India from patients with febrile illness. *Indian J Med Res.* 1967;55(12):1295-1305.
- Rao B, Basu A, Wairagkar NS, et al. A large outbreak of acute encephalitis with high fatality rate in children in Andhra Pradesh, India, in 2003, associated with Chandipura virus. *The Lancet.* 2004;364(9437):869-874. doi:10.1016/S0140-6736(04)16982-1
- Sapkal GN, Sawant PM, Mourya DT. Suppl-2, M2: Chandipura viral encephalitis: a brief review. *Open Virol J.* 2018;12(suppl 2, M2):44-51. doi:10.2174/1874357901812010044
- Traoré-Lamizana M, Fontenille D, Diallo M, et al. Arbovirus surveillance from 1990 to 1995 in the Barkedji area (Ferlo) of Senegal, a possible natural focus of Rift Valley fever virus. *J Med Entomol.* 2001;38(4):480-492. doi:10.1603/0022-2585-38.4.480
- Bisen PS, Raghuvanshi R. *Emerging Epidemics: Management and Control.* John Wiley & Sons; 2013.
- Sudeep A, Gurav Y, Bondre V. Changing clinical scenario in Chandipura virus infection. *Indian J Med Res.* 2016;143(6):712-721. doi:10.4103/0971-5916.191929
- Cherian SS, Gunjikar RS, Banerjee A, Kumar S, Arankalle VA. Whole genomes of Chandipura virus isolates and comparative analysis with other rhabdoviruses. *PLOS One.* 2012;7(1):e30315. doi:10.1371/journal.pone.0030315
- Basak S, Mondal A, Polley S, Mukhopadhyay S, Chattopadhyay D. Reviewing Chandipura: a vesiculovirus in human epidemics. *Biosci Rep.* 2007;27(4-5):275-298. doi:10.1007/s10540-007-9054-z
- Neumann G, Whitt MA, Kawaoka Y. A decade after the generation of a negative-sense RNA virus from cloned cDNA—what have we learned? *J Gen Virol.* 2002;83(11):2635-2662. doi:10.1099/0022-1317-83-11-2635
- Venkateswarlu C, Arankalle V. Recombinant glycoprotein based vaccine for Chandipura virus infection. *Vaccine.* 2009;27(21):2845-2850. doi:10.1016/j.vaccine.2009.02.089
- Chauhan V, Rungta T, Goyal K, Singh MP. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Sci Rep.* 2019;9(1):1-15. doi:10.1038/s41598-019-39299-8
- Dorosti H, Eslami M, Negahdaripour M, et al. Vaccinomics approach for developing multi-epitope peptide pneumococcal vaccine. *J Biomol Struct Dyn.* 2019;37(13):3524-3535. doi:10.1080/07391102.2018.1519460
- Hajighahramani N, Eslami M, Negahdaripour M, et al. Computational design of a chimeric epitope-based vaccine to protect against *Staphylococcus aureus* infections. *Mol Cell Probes.* 2019;46:101414. doi:10.1016/j.mcp.2019.06.004
- Lu I-N, Farinelle S, Sausy A, Muller CP. Identification of a CD4 T-cell epitope in the hemagglutinin stalk domain of pandemic H1N1 influenza virus and its antigen-driven TCR usage signature in BALB/c mice. *Cell Mol Immunol.* 2017;14(6):511-520. doi:10.1038/cmi.2016.20
- Khan A, Junaid M, Kaushik AC, et al. Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PLOS One.* 2018;13(5):e0196484. doi:10.1371/journal.pone.0196484
- Davies MN, Flower DR. Harnessing bioinformatics to discover new vaccines. *Drug Discovery Today.* 2007;12(9-10):389-395. doi:10.1016/j.drudis.2007.03.010
- Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics.* 2004;5(1):113. doi:10.1186/1471-2105-5-113
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018;35(6):1547-1549. doi:10.1093/molbev/msy096
- Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. *Prot Sci.* 2018;27(1):135-145. doi:10.1002/pro.3290
- Pais FS-M, de Cássia Ruy P, Oliveira G, Coimbra RS. Assessing the efficiency of multiple sequence alignment programs. *Algorithms Mol Biol.* 2014;9(1):1-8. doi:10.1186/1748-7188-9-4
- Larsen MV, Lundsgaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics.* 2007;8:424. doi:10.1186/1471-2105-8-424
- Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct.* 2013;8(1):30. doi:10.1186/1745-6150-8-30
- Stratmann T. Cholera toxin subunit B as adjuvant—an accelerator in protective immunity and a break in autoimmunity. *Vaccines.* 2015;3(3):579-596. doi:10.3390/vaccines3030579
- Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics.* 2007;8(1):4. doi:10.1186/1471-2105-8-4
- Magnan CN, Zeller M, Kayala MA, et al. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics.* 2010;26(23):2936-2943. doi:10.1093/bioinformatics/btq551
- Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v. 2—a server for in silico prediction of allergens. *J Mol Model.* 2014;20(6):2278. doi:10.1186/1471-2105-14-S6-S4
- Dimitrov I, Naneva L, Doytchinova I, Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics.* 2014;30(6):846-851. doi:10.1093/bioinformatics/btt619
- Chen J, Ng MM-L, Chu JJH. Activation of TLR2 and TLR6 by dengue NS1 protein and its implications in the immunopathogenesis of dengue virus infection. *PLOS Pathog.* 2015;11(7):e1005053. doi:10.3390/vaccines7030088
- Patel MC, Shirey KA, Pletneva LM, et al. Novel drugs targeting Toll-like receptors for antiviral therapy. *Future Virol.* 2014;9(9):811-829. doi:10.2217/fvl.14.70
- Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1:19-25. doi:10.1016/j.softx.2015.06.001
- Turner P. XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology; 2005.
- 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(suppl_2):W526-W531. doi:10.1093/nar/gki376
33. Sharp PM, Li W-H. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 1987;15(3):1281-1295. doi:10.1093/nar/15.3.1281
34. Rapin N, Lund O, Bernaschi M, Castiglione F. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLOS One.* 2010;5(4):e9862. doi:10.1371/journal.pone.0009862
35. Castiglione F, Mantile F, De Berardinis P, Prisco A. How the interval between prime and boost injection affects the immune response in a computational model of the immune system. *Comput Math Methods Med.* 2012;2012:842329. doi:10.1155/2012/842329
36. Baquero E, Albertini AA, Raux H, et al. Structure of the low pH conformation of Chandipura virus G reveals important features in the evolution of the vesiculovirus glycoprotein. *PLOS Pathog.* 2015;11(3):e1004756. doi:10.1371/journal.ppat.1004756
37. Saadi M, Karkhah A, Nouri HR. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infect Genet Evol.* 2017;51:227-234. doi:10.1016/j.meegid.2017.04.009
38. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein–protein docking. *Nat Protoc.* 2017;12(2):255-278. doi:10.1038/nprot.2016.169
39. Adcock SA, McCammon JA. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev.* 2006;106(5):1589-1615. doi:10.1021/cr040426m
40. McQuarrie D. *Statistical Mechanics, Chap. 21.* Harper and Row; 1976.
41. Ghosh S, Dutta K, Basu A. Chandipura virus induces neuronal death through Fas-mediated extrinsic apoptotic pathway. *J Virol.* 2013;87(22):12398-12406. doi:10.1128/JVI.01864-13
42. Menghani S, Chikhale R, Raval A, Wadibhasme P, Khedekar P. Chandipura virus: an emerging tropical pathogen. *Acta Trop.* 2012;124(1):1-14. doi:10.1016/j.actatropica.2012.06.001
43. Jadi R, Sudeep A, Barde P, Arankalle V, Mishra A. Development of an inactivated candidate vaccine against Chandipura virus (Rhabdoviridae: Vesiculovirus). *Vaccine.* 2011;29(28):4613-4617. doi:10.1016/j.vaccine.2011.04.063
44. Gaafar B, Ali SA, Abd-elrahman KA, Almofti YA. Immunoinformatics approach for Multiepitope vaccine prediction from H, M, F, and N proteins of peste des petits ruminants virus. *J Immunol Res.* 2019;2019:2019. doi:10.1155/2019/6124030
45. He R, Yang X, Liu C, et al. Efficient control of chronic LCMV infection by a CD4 T cell epitope-based heterologous prime-boost vaccination in a murine model. *Cell Mol Immunol.* 2018;15(9):815-826. doi:10.1038/cmi.2017.3
46. Jiang P, Cai Y, Chen J, et al. Evaluation of tandem Chlamydia trachomatis MOMP multi-epitopes vaccine in BALB/c mice model. *Vaccine.* 2017;35(23):3096-3103. doi:10.1016/j.vaccine.2017.04.031
47. Lennerz V, Gross S, Gallerani E, et al. Immunologic response to the survivin-derived multi-epitope vaccine EMD640744 in patients with advanced solid tumors. *Cancer Immunol Immunother.* 2014;63(4):381-394. doi:10.1007/s00262-013-1516-5
48. Zhu S, Feng Y, Rao P, et al. Hepatitis B virus surface antigen as delivery vector can enhance Chlamydia trachomatis MOMP multi-epitope immune response in mice. *Appl Microbiol Biotechnol.* 2014;98(9):4107-4117. doi:10.1007/s00253-014-5517-x
49. Lu C, Meng S, Jin Y, et al. A novel multi-epitope vaccine from MMSA-1 and DKK 1 for multiple myeloma immunotherapy. *Br J Haematol.* 2017;178(3):413-426. doi:10.1111/bjh.14686
50. Li Y, Zheng K, Tan Y, et al. A recombinant multi-epitope peptide vaccine based on MOMP and CPSIT_p6 protein protects against Chlamydia psittaci lung infection. *Appl Microbiol Biotechnol.* 2019;103(2):941-952. doi:10.1007/s00253-018-9513-4
51. Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. *Cell Mol Immunol.* 2018;15(2):182-184. doi:10.1038/cmi.2017.92
52. Dar HA, Zaheer T, Shehroz M, et al. Immunoinformatics-aided design and evaluation of a potential multi-epitope vaccine against *Klebsiella pneumoniae*. *Vaccines.* 2019;7(3):88.
53. Walker JM. *The Proteomics Protocols Handbook.* Springer; 2005.
54. Kyte J, Doolittle RF. A simple method for displaying the hydrophobic character of a protein. *J Mol Biol.* 1982;157(1):105-132. doi:10.1016/0022-2836(82)90515-0
55. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem.* 1980;88(6):1895-1898. doi:10.1093/oxfordjournals.jbchem.a133168
56. Foroutan M, Ghaffarifard F, Sharifi Z, Dalimi A. Vaccination with a novel multi-epitope ROP8 DNA vaccine against acute Toxoplasma gondii infection induces strong B and T cell responses in mice. *Comp Immunol Microbiol Infect Dis.* 2020;69:101413. doi:10.1016/j.cimid.2020.101413
57. Bazhan SI, Antonets DV, Karpenko LI, et al. In silico designed ebola virus T-cell multi-epitope DNA vaccine constructions are immunogenic in mice. *Vaccines.* 2019;7(2):34. doi:10.3390/vaccines7020034
58. Boehme KW, Compton T. Innate sensing of viruses by toll-like receptors. *J Virol.* 2004;78(15):7867-7873. doi:10.1128/JVI.78.15.7867-7873.2004
59. Cartly M, Bowie AG. Recent insights into the role of Toll-like receptors in viral infection. *Clin Exp Immunol.* 2010;161(3):397-406. doi:10.1111/j.1365-2249.2010.04196.x
60. Anukumar B, Shahir P. Chandipura virus infection in mice: the role of toll like receptor 4 in pathogenesis. *BMC Infect Dis.* 2012;12(1):125. doi:10.1186/1471-2334-12-125
61. Vaure C, Liu Y. A comparative review of toll-like receptor 4 expression and functionality in different animal species. *Front Immunol.* 2014;5:316. doi:10.3389/fimmu.2014.00316
62. Hou J, Liu Y, Hsi J, Wang H, Tao R, Shao Y. Cholera toxin B subunit acts as a potent systemic adjuvant for HIV-1 DNA vaccination intramuscularly in mice. *Hum Vaccines Immunother.* 2014;10(5):1274-1283. doi:10.4161/hv.28371
63. Phongsisay V, Iizasa Ei, Hara H, Yoshida H. Evidence for TLR4 and FcR γ -CARD9 activation by cholera toxin B subunit and its direct bindings to TREM2 and LMIR5 receptors. *Mol Immunol.* 2015;66(2):463-471. doi:10.1016/j.molimm.2015.05.008

64. Saeed U, Mazoor S, Jalal N, Piracha ZZ. Contemplating the importance of Toll-like receptors I and II regarding human viral pathogenesis. *Jundishapur J Microbiol*. 2015;8(1):13348. doi:10.5812/jjm.13348
65. Solanki V, Tiwari M, Tiwari V. Prioritization of potential vaccine targets using comparative proteomics and designing of the chimeric multi-epitope vaccine against *Pseudomonas aeruginosa*. *Sci Rep*. 2019;9(1):1-19. doi:10.1038/s41598-019-41496-4
66. Ali A, Khan A, Kaushik AC, et al. Immunoinformatic and systems biology approaches to predict and validate peptide vaccines against Epstein-Barr virus (EBV). *Sci Rep*. 2019;9(1):1-12. doi:10.1038/s41598-018-37070-z
67. 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):1-11. doi:10.1038/s41598-018-19456-1
68. Ojha R, Pareek A, Pandey RK, Prusty D, Prajapati VK. Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus zoonotic infection. *ACS Omega*. 2019;4(8):13069-13079. doi:10.1021/acsomega.9b00944
69. Mishra S, Sinha S. Immunoinformatics and modeling perspective of T cell epitope-based cancer immunotherapy: a holistic picture. *J Biomol Struct Dyn*. 2009;27(3):293-306. doi:10.1080/07391102.2009.10507317
70. Nezafat N, Ghasemi Y, Javadi G, Khoshnoud MJ, Omidinia E. A novel multi-epitope peptide vaccine against cancer: an in silico approach. *J Theor Biol*. 2014;349:121-134. doi:10.1016/j.jtbi.2014.01.018
71. Parvizpour S, Razmara J, Pourseif MM, Omidi Y. In silico design of a triple-negative breast cancer vaccine by targeting cancer testis antigens. *BioImpacts*. 2019;9(1):45-56. doi:10.15171/bi.2019.06

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Deb D, Basak S, Kar T, et al. Immunoinformatics based designing a multi-epitope vaccine against pathogenic *Chandipura vesiculovirus*. *J Cell Biochem*. 2021; 1-25. doi:10.1002/jcb.30170