

A Candidate Multi-Epitope Vaccine Against Pathogenic Chandipura Vesiculovirus Identified using Immunoinformatics.

Debashrito Deb

Garden City University <https://orcid.org/0000-0002-8704-3534>

Srijita Basak

Garden City University <https://orcid.org/0000-0001-6947-128X>

Tamalika Kar

Garden City University <https://orcid.org/0000-0001-6241-9681>

Utkarsh Narsaria

Garden City University <https://orcid.org/0000-0001-6550-1998>

Filippo Castiglione

Garden City University <https://orcid.org/0000-0002-1442-3552>

Abhirup Paul

Garden City University <https://orcid.org/0000-0003-2143-7511>

Ashutosh Pandey

National Institute for Plant Genome Research <https://orcid.org/0000-0002-5018-9628>

Anurag Prakash Srivastava (✉ anuiitkgp@gmail.com)

Garden City University <https://orcid.org/0000-0003-3493-1375>

Research article

Keywords: Chandipura vesiculovirus (CHPV) , CTL, HTL, IFN- γ

Posted Date: August 12th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-56434/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

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 non-allergenic vaccine against this pathogen. 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 labour intensive. These limitations can be overcome by peptide-based vaccines comprising of 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 non-allergic. 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.

Introduction

Chandipura vesiculovirus (CHPV), associated with an encephalitic illness in humans, is a member of the *Rhabdoviridae* family, first discovered by Bhatt and Rodrigues in 1966. 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 [2–5], studies have also shown its presence in Asian [5, 6] and African countries [5, 7, 8]. It also infects many mammalian species around the globe [9]. Due to its worldwide occurrence, it has gained global attention as an emerging neurotropic pathogen, imposing a mortality rate ranging from 55–77% [2, 10]. 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 [11].

CHPV was isolated from sand-flies, which are the potential vectors for transmitting the disease [12, 13]. The viral genome is ~ 11 kb [14] consisting of single stranded RNA coding for five polypeptides namely, Nucleocapsid protein N, Phosphoprotein P, Matrix protein M, Glycoprotein G and Large protein L [5, 11, 15, 16]. 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 [17]. Spike like glycoprotein present in the membrane act as the major antigenic determinant [18]. Thus, glycoprotein becomes a potent choice for vaccine designing [19]. A recombinant vaccine against CHPV has been developed in mice using the entire G gene [19], but it lacks human trials [10]. Hence, a licensed vaccine is still not available [10]. 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 [20]. 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 [20]. Recent advancements in computational biology are very effective in designing various approaches for vaccine synthesis [21–24]. One of these advancements is the emerging role of immunoinformatics tools in effective vaccine designing [25–28]. Several antigenic epitopes predicted using various immunoinformatics tools have potential translational implications [25, 26]. These epitopes are used for the construction of a multi-epitope vaccine which can activate both humoral and adaptive immune responses [29, 30]. The immunoinformatic approach is time saving as well as cost effective and has the potential to ensure a successful vaccine design [31]. The techniques used in this approach have the potential to identify candidate

proteins that might be overlooked by conventional experimentation [32]. 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 (Fig. 1).

Methodology

Selection and retrieval of target protein

The VIPR database (<https://www.viprbrc.org/brc/home.spg?decorator=vipr>) was used to retrieve the glycoprotein sequences of the viruses belonging to the genus *Vesiculovirus* of the *Rhabdoviridae* family in order to investigate their evolutionary relationship. These amino acid sequences were aligned implementing the MUSCLE program in MEGA X [33]. 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 [34]. The evolutionary history was inferred by using the Maximum Likelihood method and JTT (Jones-Taylor-Thornton) matrix-based model [35]. 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.

Homology modelling, refinement and validation of glycoprotein

The structure of the glycoprotein was modelled using Raptor X server (<http://raptorx.uchicago.edu/>), a homology modelling tool [36]. Using GalaxyRefine server (<http://galaxy.seoklab.org/refine>), which is based on CASP10 tested refinement method, the distortions in the structures recovered after homology modelling were reduced, and refined models were obtained [37]. The best refined structure of the glycoprotein was validated by Ramachandran plot analysis using RAMPAGE (https://github.com/kepbod/rampage_alu) [38], Z-score generated by ProSA (<https://prosa.services.came.sbg.ac.at/prosa.php>) [39] and ERRAT score (<https://servicesn.mbi.ucla.edu/ERRAT/>) [40].

Analysis of the physicochemical properties of the target proteins.

The ExPASy ProtParam server (<https://web.expasy.org/protparam/>) was used for the physicochemical analysis of the refined structure of the target glycoprotein [41]. In addition, VaxiJen v2.0 server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) was used to determine the antigenicity of the glycoprotein.

T cell epitope prediction.

CTL epitopes prediction: The NetCTL.1.2 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, i.e., A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58 and B62 [42]. In NetCTL.1.2 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/>) [43].

HTL epitopes prediction

Net MHC II pan 3.2 server (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 [44]. 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 non-binding epitopes were set at 2, 10 and more than 10%, respectively. NetMHC II pan 3.2 server predicts peptide binding on human HLADR, HLA-DQ and HLA-DP alleles using Artificial Neural Network (ANN) system.

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 [45].

Prediction of IFN- γ epitopes- IFN- γ plays a key role in provoking anti-tumour, antiviral and immune regulation activities. Thus, it is important to identify epitopes with the capacity to induce IFN- γ in order 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 [46]. The server functionality is focused on a dataset 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% [46].

Population coverage- The population coverage of the selected epitopes based on HLA genotypic frequencies was calculated using IEDB population coverage analysis (<http://tools.iedb.org/population/>) (Supplementary Table 7) [47].

Multi epitope vaccine construct, Structural Modelling, 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 [48]. 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 [49] which produced a final model after refinement. The model was then evaluated using ERRAT score [40], Verify3D score [50], RAMPAGE for Ramachandran Plot analysis [38] and ProSA generated Z-score [39] for determining the overall quality of the vaccine construct.

Antigenicity, allergenicity and physicochemical properties of the vaccine construct

The antigenicity of the final vaccine construct was checked by VaxiJen v2.0 [51] and ANTIGENpro (<http://scratch.proteomics.ics.uci.edu/>) [52]. 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 dataset 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 70–89% [51]. 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 dataset is estimated at 76% by cross-validation experiments [52]. AllerTOP v2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>) [53] and AllergenFP (<http://ddg-pharmfac.net/AllergenFP/>) [54] were used to determine the allergenicity of the vaccine construct. AllerTOP server utilizes auto-cross-covariance (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 [53]. 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 auto-cross covariance

(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 [54]. 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 [63]. 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 [55] or any peptide signals in the final vaccine construct [56].

Modelling 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 in the final vaccine construct [57]. 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 [58].

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 localise at the cell surface and gets activated [59, 60]. 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 TLR 1, 2 and 4 was carried out using HADDOCK 2.4 (<http://www.bonvinlab.org/software/haddock2.4/>) [58]. 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 [61]. PyMOL (Schrodinger) was used for visualization of the result. PDBsum (<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>) was used in order to map the interacting residues between the vaccine and the TLRs [62].

Energy minimization and Molecular dynamics simulation

Vaccine construct- GROMACS (GRONingen MAchine for Chemical Simulations) a command line Linux-based program was used for the Molecular Dynamics Simulation (MDS) and energy minimization [63]. The vaccine structure was subjected to MDS in order to mimic the experimental conditions. In this study, the physical conditions like the temperature and the pressure was mimicked using the canonical ensemble, NVT and the isobaric and isothermal ensemble, NPT. The temperature used for the simulation was 300K and the pressure used was of 1 bar in order 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 in order 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 pico-seconds (ps) consisting of 500,000 steps in order to stabilize the temperature and pressure of the system. A 30 ns Molecular Dynamics (MD) simulation run was performed for the energy minimized structure in order 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 visualise the graphs generated from the simulation [64].

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 pico-seconds comprising of 500,000 steps for stabilizing the temperature and pressure of the system. A molecular dynamics simulation run of 30 ns of the energy minimized structure was performed to find the Root Mean Square Deviation (RMSD) of backbone and Root Mean Square Fluctuation (RMSF) of side chain.

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 cDNA sequence [65]. This cDNA is optimized for the vaccine to express in *E. 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 [66]. GC content of the vaccine should range from 30–70%, which indicates good translation and transcription efficiencies in *E. coli* [65]. Eventually, the SnapGene tool was used to insert the optimized multi-epitope vaccine sequence into the pET-28a (+) vector.

Immune simulation

To further characterize the chimeric peptide's immunogenicity and immune response profile, the C-ImmSim server (<https://www.iac.cnr.it/~filippo/projects/c-immsim-online.html>) was used to perform *in silico* immune simulations [67]. 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 [67]. For most of the vaccines currently in use, the minimum recommended time between the dose 1 and dose 2 is 4 weeks [68]. The entire simulation ran for 1400 time steps which are about 15 months (a time step is about 8 hours). 12 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.

Results

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). A phylogenetic tree was constructed using all the glycoproteins of different viruses of the genus *Vesiculovirus* belonging to *Rhabdoviridae* family in order to check the evolutionary relationship of CHPV glycoprotein with the glycoprotein of other *vesiculovirus* (Fig. 2). The CHPV glycoproteins clustered together in a single clade indicating that all the CHPV glycoproteins are closely related to one another with minimum variations (Fig. 2). In addition, a phylogeny tree was also constructed using all the glycoproteins of different strains of CHPV (Supplementary Fig. 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 results suggest that a vaccine designed against one strain of CHPV would be effective against all the other strains of the virus. The accession numbers and the percentage similarity table of all the strains of CHPV glycoprotein are shown in Supplementary Table 1. The target glycoprotein was also checked for human proteome hit recognition in order 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.

Homology modelling, 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 [69], 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 modelling tool (Fig. 3), and the 3D model was used to locate the epitopes selected in this study (Fig. 4). GalaxyRefine server predicted 5 refined models of the glycoprotein from which Model 5 was selected (Supplementary Table 2). In the chosen model, the favoured regions in the Ramachandran plot were 98.7% (Supplementary Fig. S2), ERRAT score was 89.03 (Supplementary Fig. S2) and Z-Score was - 6.07 (Supplementary Fig. S2). The molecular weight of the glycoprotein was found to be 59.03 kDa by ExPASy ProtParam server.

Prediction of CTL and HTL epitopes

For the prediction of CTL epitopes, the NetCTL 1.2 [42] and IEDB consensus methods [43] were used (Table 1) (Supplementary Table 3), whereas the HTL epitopes were predicted using NetMHC II pan 3.2 server (Table 2) (Supplementary Table 4) [44]. 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 (Supplementary Table 5) based on all the above mentioned criteria. The overlapping epitopes were antigenic, as predicted using VaxiJen v2.0 server.

Table 1

CTL epitopes showing promiscuity were predicted using NetCTL 1.2. 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.

CTL BINDING WITH HLA CLASS I						
EPITOPE	POSITION	SUPERTYPE	ALLELE CLASS I	CONSENSUS PERCENTILE RANK	IC50	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	

CTL BINDING WITH HLA CLASS I						
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	

Table 2

HTL epitopes showing promiscuity were predicted using NetMHC II pan 3.2 server.
Antigenicity scores were predicted using VaxiJen v2.0 keeping a threshold of 0.4.

EPITOPES (POSITION)	HLA ALLELES II	SCORE	ANTIGENICITY
SEFLVIMITPHHVG(143)	DRB1*01:01	0.40	0.8996
	DRB1*04:05	0.07	
	DRB1*04:01	0.17	
	DRB1*07:01	1.20	
	DRB1*08:02	0.03	
	DRB1*11:01	0.80	
	DRB1*12:01	0.05	
	DRB1*15:01	0.08	
	DRB4*01:01	0.04	
	DRB5*01:01	0.90	
TDSEFLVIMITPHHV(141)	DRB1*01:01	1.00	0.7730
	DRB1*04:05	0.08	
	DRB1*04:01	0.25	
	DRB1*07:01	1.60	
	DRB1*08:02	0.12	
	DRB1*11:01	1.80	
	DRB1*12:01	0.12	
	DRB1*15:01	0.30	
	DRB4*01:01	0.12	
	DRB5*01:01	1.90	
FLVIMITPHHVG(145)	DRB1*01:01	1.00	0.9586
	DRB1*04:01	1.00	
	DRB1*04:05	0.60	
	DRB1*08:02	0.09	
	DRB1*11:01	1.40	
	DRB1*12:01	0.15	
	DRB1*15:01	0.25	
	DRB4*01:01	0.12	
	DRB5*01:01	1.70	
	SVILLISFIAPSYSS(8)	DRB1*04:05	
DRB1*08:02		1.60	

EPITOPES (POSITION)	HLA ALLELES II	SCORE	ANTIGENICITY
	DRB1*12:01	0.60	
	DRB1*15:01	0.70	
	DQA1*01:01-DQB1*05:01	0.80	
GTLTSAFHTRYVRMWI(345)	DPA1*02:01-DPB1*01:01	1.20	0.5170
	DPA1*01:03-DPB1*02:01	1.40	
	DPA1*01:03-DPB1*04:01	1.20	
	DPA1*02:01-DPB1*05:01	1.50	
	DRB1*07:01	0.50	
	DRB1*09:01	1.20	
NGTLTSAFHTRYVRMW (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	
KTKEIAAGAI VFKSK (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	

Multi epitope vaccine construct, structural properties and refinement.

The final multi-epitope vaccine was designed by combining 6 CTL, 7 HTL (Tables 1 & 2) and 3 IFN- γ (Supplementary Table 6) epitopes via GPGPG linkers that prevented the chance of any junctional epitope formation [70]. To the final vaccine construct, the CTB adjuvant was linked at the N-terminal of the vaccine with EAAAK linker (Fig. 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 (Supplementary Table 5), (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 web server [49] (Fig. 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 in order 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 (Fig. 5D) [39]. Ramachandran plot analysis showed 93.5%, 3.9% and 2.6% residues in favoured, allowed and outlier regions, respectively (Fig. 5E), which verified the overall quality of the final multi-epitope vaccine construct [38]. Verify3D predicted that 81.44% of the amino acids scored ≥ 0.2 , which further supported the high quality structure of the vaccine model (Fig. 5C) [50]. ERRAT of the already refined vaccine construct projected a score of 59.5819 which further verifies the overall quality of the vaccine construct (Supplementary Fig. S3) [40].

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 colour represents the EAAAK linker and pale blue colour represents GPGPG linkers. (B) Tertiary structure of the final vaccine construct. Helices, sheets and loops are represented by red, limon and deep teal colour, respectively. (C) Verify3D 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.

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 non-allergenic as predicted by AllerTOP and AllergenFP. Evaluation of the physicochemical properties predicted by ExPASy (Supplementary 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 was 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 hours in mammalian reticulocytes, > 20 hours in yeast and > 10 hours in *Escherichia coli*. The grand average hydropathicity (GRAVY) value was recorded as 0.083, indicating the polar nature of the vaccine construct and there was no signal peptide detected which hence will prevent protein localization (Supplementary Fig. S4). Since no transmembrane helix has been identified in the designed vaccine construct, no expression difficulties are anticipated in the vaccine development (Supplementary Fig. S5).

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 Table 3 and Table 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.) (Supplementary Fig. S6). Furthermore, the IFN- γ inducing epitopes were predicted using IFNepitope server from the target glycoprotein (Supplementary Table 6).

Table 3

Conformational B cell epitopes in the Vaccine construct as predicted by ElliPro server. Positions of the epitopes are indicated in the brackets.

DISCONTINUOUS EPITOPES	SCORE
MT(1-2) LCAEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGATFQVEVPGSQHIDSQKKAIERMKDTRLRIAYL (9-78)	0.782
APSYSSGPGPGTDSE(203-217)	0.726
GPGPGL(119-124)	0.694
GPGPGFFPNGEWV(147-158) NGTLS(169-173)	0.64
VGVGPGPGF(246-254) DDGPGPGKTK(267-276) V(284) KSKGPGPGNGT(286-296) S(298)	0.624
HTR(301-303)V(305)M(307) GPGPGGTLFAHTRYV(309-324)GPGPGGKG(329-336) VINGTSLFGPGPGGLAYTVINGTSLFAHGPGPGLAY(341-376) VINGTSLFAHT(378-388)	0.621
CVWNNKT(87-93)	0.568
FGPGPG(174-179)	0.544

Table 4
Linear B cell epitopes in the Vaccine construct as predicted by ElliPro server.

LINEAR EPITOPES	POSITION	SCORE
LCAEYHNTQIHTLNDKIFSYTESLAGKREMAITFKNGATFQVEVPGSQHIDSQKKAIERMKDTRLRIAY	9–77	0.797
APSYSSGPGPGTDS	203–216	0.750
DDGPGPGKTKEI	267–278	0.724
GPGPGFFPNGEWV	147–158	0.717
GPGPGL	119–124	0.694
FGPGGGLAYTVINGTLSFAHGPGPGLAYTVINGTLSFAHT	348–388	0.675
WPGPGGGLTSLFAHTRY	308–323	0.64
FKSKGPGPGNG	285–295	0.623
VWNNKT	88–93	0.594
GPGPGGKG	329–336	0.589
GVGPGPG	247–253	0.567
GTLSFGPGPG	170–179	0.520

Modelling 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 in order to evaluate their binding patterns (Fig. 6).

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.

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 supplementary figure S9. The docked structure along with amino acid interaction is shown in Fig. 7. The detailed overview and the interaction with amino acid sequences are given in supplementary figure S10 and supplementary material SM2, respectively. The secondary structure of the docked complex was predicted (Supplementary Fig. S8), and further Ramachandran plot and Z-score analysis were carried out for structural validation of the docked complex (Supplementary Fig. S7).

Table 5

A. Table showing statistics of docked vaccine and TLR4 cluster. 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).

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

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 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 the supplementary figure S13. The docked structure along with amino acid interaction is shown in Fig. 8. A detailed overview and the interaction with amino acid sequences are given in supplementary figure S14 and supplementary material SM3, respectively. The secondary structure of the docked complex was predicted (Supplementary Fig. S12), and further, Ramachandran plot and Z-score analysis were carried out for structural validation of the docked complex (Supplementary Fig. S11)

Table 5

B. Table showing statistics of docked vaccine and TLR2 cluster. 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).

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

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. 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 supplementary figure S17. The docked structure along with amino acid interaction is shown in Fig. 9. A detailed overview and the interaction with amino acid sequences are given in supplementary figure S18 and supplementary material SM4, respectively. The secondary structure of the docked complex was predicted (Supplementary Fig. S16), and further Ramachandran plot and Z-score analysis were carried out for structural validation of the docked complex (Supplementary Fig. S15).

Table 5

C. Table showing statistics of docked vaccine and TLR1 cluster. 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).

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

Energy minimization and Molecular dynamics simulation of vaccine and vaccine-TLR4 complex

Molecular Dynamics Simulation is used for studying the thermodynamic properties and time dependent phenomena [71, 72]. In this study Molecular Dynamics Simulation was conducted in order 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 molecular dynamics simulation 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 (Supplementary Fig. 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, in order to study its various thermodynamic properties and the effect of temperature and pressure on the system. The first equilibration phase was conducted under the NVT ensemble in order to check if the system remained stable at the desired temperature of 300K. The temperature of the system quickly reached 300K and was maintained for the equilibration period with minimal fluctuations, indicating the stability of the system at 300K (Supplementary Fig. S19). The system was then subjected to a second equilibration phase under the NPT ensemble in order 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 (Supplementary Fig. 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 (Supplementary Fig. 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 in order 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 300K and was maintained for most of the equilibration phase, indicating its stability (Supplementary Fig. S20). Similarly, the NPT equilibration of the vaccine-TLR4 complex suggested that the system remains stable at an atmospheric pressure of 1 bar which is evident from the very little fluctuations as observed in the plot (Supplementary Fig. 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 30 ns in order to monitor the overall stability of the vaccine and vaccine-TLR4 complex. The RMSD plot of the vaccine obtained after the simulation period of 30 ns showed successive fluctuation up to 10 ns, and then it reached a steady state for the rest of the simulation (Fig. 10A). The RMSD plot of the vaccine-TLR4 complex indicated that the complex endured fluctuations until 12 ns and then reached a steady state which remained stable for the rest of the simulation (Fig. 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 vaccine-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 in order 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 (Fig. 10B). In particular, Pro372, Leu383, Phe385, 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 (Fig. 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 in order 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 (Supplementary Fig. S19 & S20).

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% [65]. The CAI (Codon Adaptation Index) value was calculated to be 1, meaning the multi-epitope vaccine construct ventures high level expression in *E. coli* K-12 strain [66]. The cDNA of the vaccine was then inserted into pET-28a (+) vector for restriction cloning (Fig. 11).

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 Fig. 12. The graphs showed significantly higher secondary

and tertiary responses as compared to the primary response. Elevated levels of immunoglobulin 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 behaviour, suggesting possible class switching and memory development (Fig. 12A & B) (Supplementary Fig. S21). A relatively high response was observed in populations of TH (helper) and TC (cytotoxic) cells with pre activation of TCs during vaccination (Fig. 12D & C) (Supplementary Fig. S21). Higher macrophage activity was demonstrated during exposure, while dendritic and NK cell activity was found to be consistent (Supplementary Fig. S21). High levels of IFN- γ and IL-2 were also evident which says that a good immune response has been generated (Fig. 12E). In order to check the efficacy of the vaccine, a live replicating virus was injected at around day 366. The results from the antigen graph (Fig. 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. In order 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 (Supplementary Fig. S22).

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 [73]. There is no specific treatment available for the disease till date but a symptomatic treatment is done using mannitol to reduce the brain edema [16]. 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 [19]. A β -propio lactone (BPL) inactive tissue culture based vaccine has also been developed [74]. Unfortunately, for both types of vaccines developed so far, clinical trials have been performed only in mice but not in humans [10]. 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 unfavourable epitopes unlike the case of using complete antigen in recombinant vaccine technology [20, 75, 76].

The advances in the field of immunoinformatics, along with the knowledge of host immune response is being successfully used for the development of epitope based vaccines against various pathogens [77, 78]. Designing of epitope driven vaccine is a modern idea, which is successfully applied in several immunological studies for the development of vaccines [79]. Similarly, the present study is centred 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 effects is lowered because it is less likely to contain unwanted components [29,30,70,80–83]. 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 [29]. The cellular and humoral responses generated by the multi-epitope vaccines are highly specific with increased cytokines production [84–86]. A multi-epitope vaccine developed with such precautions can thus become a powerful tool in the battle against viral infections [87].

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*, i.e., good candidate antigens might be overlooked [32]. *In silico* methods whereas, screens for all the probable candidate antigens, as predicted by various tools and algorithms which might

otherwise remain undetected [32]. 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 [88]. The CTL, HTL and IFN- γ epitopes selected for the study were screened against various immunological filters (Table 1 & Table 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 non-allergenic. 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 [41]. Moreover, the vaccine has exhibited a fair percentage of solubility in over-expressed conditions. The GRAVY index of the vaccine was 0.083 indicating the vaccine's polar nature and its effective interaction with water [89]. 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) [90]. The half-life of the vaccine is 30 hours in mammalian reticulocytes, > 20 hours in yeast and > 10 hours in *Escherichia 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 in order to assess the allergenicity and physicochemical properties of their designed vaccine candidate against *Toxoplasma gondii* [91]. 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 [91]. The physicochemical properties predicted in our study were comparable to those predicted by Foroutan *et al.*, in their recently published work [91]. In fact, the instability index and aliphatic index of our vaccine candidate was found to be better when compared to the values reported by [91]. The Ramachandran Plot, ERRAT score, Verify3D score and Z score analysis validated the overall quality of the vaccine construct (Fig. 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 [92].

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 [93–95]. CHPV regulates TLR4, which leads to the secretion of pro-inflammatory cytokines and Nitric oxide (NO) in monocytes-macrophage cell line of mouse [96]. In humans, TLR4 is expressed in immune cells such as monocytes, macrophages, granulocytes and immature dendritic cells [97]. Cholera toxin B (CTB) has been proven to act as a potential viral adjuvant [98–100]. Activation of TLR4 by CTB is presumably due to the direct interaction between TLR4 and CTB [101]. 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- κ B activation in TLR4 receptor cells by direct binding, which has been demonstrated using ELISA-based assays [101]. Further, TLR2 has been associated with the recognition of viral envelope glycoprotein [93]. The core TLR2 signalling pathway utilizes myeloid differentiation factor 88 (MyD88)

as the primary adaptor, and results in NF- κ B and mitogen-activated protein kinase (MAPK) activation as well as secretion of a core panel of cytokines [93]. It has also been reported that TLR2 acts as heterodimer with TLR1 for the generation of innate immune response and has been shown to recognise viral proteins [59, 93–95]. TLR1/TLR2 dimer generates intracellular signalling via IRAK4 mediated activation of IRAK1/2, which results in the activation of NF- κ B, p38 and JNK proteins in the cytoplasm, involved in triggering innate immune response [102].

Figure13. Proposed mechanism of working (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 up regulate 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.

Hence, the interaction pattern of multi-epitope vaccine against TLR4/TLR2/TLR1 was analyzed using Molecular Docking Analysis (Figs. 7, 8 & 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 co-receptor protein and the vaccine. Thus, docking studies indicate that both TLR4 and MD-2 are responsible for a stable interaction of the vaccine. Molecular Dynamics Simulation (MDS) for both the vaccine and vaccine-TLR4 complex was performed in order 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 300K 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 (Fig. 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 intra molecular interactions between the vaccine and 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 (+) (Fig. 11). Several research groups have recently applied similar strategy to design a multi-epitope vaccine against *Klebsiella pneumoniae* [88], *Kaposi Sarcoma* [20], *Pseudomonas aeruginosa* [103], Epstein Barr virus [104], Malaria [105], Hendra virus [106] and Nipah virus [107]. Similar approaches have also been used for developing vaccine against cancer antigens [22, 108–110] The proposed mechanism of action was also predicted for the final vaccine model (Fig. 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.

Conclusion

CHPV is a fast growing pathogen that causes acute encephalitis. Applying computational methods in vaccine design will greatly improve the vaccine discovery process and can achieve that goal in less time and at lower cost. 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 non-allergenic. Molecular docking, molecular dynamics simulation, and *in silico* immune simulation were used to check the final vaccine's stabilization and immune responses. The results obtained can help in the design of an effective vaccine against CHPV infection.

Abbreviations

CHPV- Chandipura vesiculovirus

CTL- Cytotoxic T lymphocyte

HTL- Helper T lymphocyte

IFN- Interferon

JTT- Jones Taylor-Thorton

VIPR- Virus Pathogen Database and Analysis Resource

NJ- Neighbor Joining

IEDB- Immune Epitope Database

ACC- Autocross Variance

QSAR- Quantitative Structure Activity Relationship

HLA- Human Leukocyte Antigen

TLR- Toll Like Receptors

MHC- Major Histocompatibility Complex

MDS- Molecular Dynamics Simulation

OPLS-AA- Optimized Potential for Liquid Simulation- All Atom

RMSD- Root mean square deviation

RMSF- Root mean square fluctuation

Declarations

Availability of data and materials

Correspondence: Correspondence and requests for materials should be addressed to Anurag P. Srivastava, anuiitkgp@gmail.com

Competing interests

There are no competing financial interests.

Funding

Not Applicable

Authors' contributions

D.D., designed and performed experiments and helped in the writing, S. B., designed and performed epitope selection prediction experiment, and helped in the writing, T.K., performed the docking experiments and helped in the writing, U.N., performed molecular dynamics simulation experiment and helped in the writing, F.C., performed immune simulation experiment and helped in writing. A. Paul., helped in phylogenetic analysis, A. Pandey and A. P. S. devised and supervised experiments, analyzed data, and helped in writing the manuscript.

Acknowledgements

The authors thank Dr. Joseph V.G., Chancellor Garden City University for his constant support to carry out this research work. FC acknowledges partial support from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826121.

References

1. Bhatt PN, Rodrigues FM (1967) Chandipura: a new Arbovirus isolated in India from patients with febrile illness. *Indian J Med Res* 55
2. Rao BL, Basu A, Wairagkar NS et al (2004) A large outbreak of acute encephalitis with high fatality rate in children in Andhra Pradesh, India, in 2003, associated with Chandipura virus. *Lancet* 364:869–874
3. Chadha MS, Arankalle VA, Jadi RS et al (2005) An outbreak of Chandipura virus encephalitis in the eastern districts of Gujarat state, India. *Am J Trop Med Hyg* 73:566–570
4. Gurav YK, Tandale BV, Jadi RS et al (2010) Chandipura virus encephalitis outbreak among children in Nagpur division, Maharashtra, 2007. *Indian J Med Res* 132:395
5. Sapkal GN, Sawant PM, Mourya DT (2018) Suppl-2, M2: Chandipura Viral Encephalitis: A Brief Review. *Open Virol J* 12:44
6. Peiris JS, Dittus WP, Ratnayake CB (1993) Seroepidemiology of dengue and other arboviruses in a natural population of toque macaques (*Macaca sinica*) at Polonnaruwa, Sri Lanka. *J Med Primatol* 22:240–245
7. Fontenille D, Traore-Lamizana M, Trouillet J et al (1994) First isolations of arboviruses from phlebotomine sand flies in West Africa. *Am J Trop Med Hyg* 50:570–574
8. Traore-Lamizana M, Fontenille D, Diallo M et al (2001) 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* 38:480–492
9. Bisen PS, Raghuvanshi R (2013) *Emerging epidemics: Management and control*. John Wiley & Sons
10. Sudeep AB, Gurav YK, Bondre VP (2016) Changing clinical scenario in Chandipura virus infection. *Indian J Med Res* 143:712
11. Cherian SS, Gunjekar RS, Banerjee A et al (2012) Whole genomes of Chandipura virus isolates and comparative analysis with other rhabdoviruses. *PLoS One* 7:e30315
12. Dhanda V, Rodrigues FM, Ghosh SN (1970) Isolation of Chandipura virus from sandflies in Aurangabad. *Indian J Med Res* 58:179–180

13. Geevarghese G, Arankalle VA, Jadi R et al (2005) Detection of chandipura virus from sand flies in the genus *Sergentomyia* (Diptera: Phlebotomidae) at Karimnagar District, Andhra Pradesh, India. *J Med Entomol* 42:495–496
14. Marriott AC (2005) Complete genome sequences of Chandipura and Isfahan vesiculoviruses. *Arch Virol* 150:671–680
15. Banerjee AK (1987) Transcription and replication of rhabdoviruses. *Microbiol Rev* 51:66
16. Menghani S, Chikhale R, Raval A et al (2012) Chandipura virus: an emerging tropical pathogen. *Acta Trop* 124:1–14
17. Basak S, Mondal A, Polley S et al (2007) Reviewing Chandipura: a vesiculovirus in human epidemics. *Biosci Rep* 27:275–298
18. Neumann G, Whitt MA, Kawaoka Y (2002) A decade after the generation of a negative-sense RNA virus from cloned cDNA—what have we learned? *J Gen Virol* 83:2635–2662
19. Venkateswarlu CH, Arankalle VA (2009) Recombinant glycoprotein based vaccine for Chandipura virus infection. *Vaccine* 27:2845–2850
20. Chauhan V, Rungta T, Goyal K, Singh MP (2019) Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Sci Rep* 9:1–15
21. Faisal A-R, Imtiaz SH, Zerín T et al (2017) Computer aided epitope design as a peptide vaccine component against Lassa virus. *Bioinformatics* 13:417
22. Atapour A, Mokarram P, MostafaviPour Z et al (2019) Designing a fusion protein vaccine against HCV: an in silico approach. *Int J Pept Res Ther* 25:861–872
23. Dorosti H, Eslami M, Negahdaripour M et al (2019) Vaccinomics approach for developing multi-epitope peptide pneumococcal vaccine. *J Biomol Struct Dyn* 37:3524–3535
24. Hajighahramani N, Eslami M, Negahdaripour M et al (2019) Computational design of a chimeric epitope-based vaccine to protect against *Staphylococcus aureus* infections. *Mol Cell Probes* 46:101414
25. Brennick CA, George MM, Corwin WL et al (2017) Neoepitopes as cancer immunotherapy targets: key challenges and opportunities. *Immunotherapy* 9:361–371
26. Palatnik-de-Sousa CB, Soares I, da S, Rosa DS (2018) Epitope discovery and Synthetic Vaccine design. *Front Immunol* 9:826
27. Patronov A, Doytchinova I (2013) T-cell epitope vaccine design by immunoinformatics. *Open Biol* 3:120139
28. Brusich V, Petrovsky N (2005) Immunoinformatics and its relevance to understanding human immune disease. *Expert Rev Clin Immunol* 1:145–157
29. Lu C, Meng S, Jin Y et al (2017) A novel multi-epitope vaccine from MMSA-1 and DKK 1 for multiple myeloma immunotherapy. *Br J Haematol* 178:413–426
30. He R, Yang X, Liu C et al (2018) Efficient control of chronic LCMV infection by a CD4 T cell epitope-based heterologous prime-boost vaccination in a murine model. *Cell Mol Immunol* 15:815–826
31. Khan A, Junaid M, Kaushik AC et al (2018) Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PLoS One* 13:e0196484
32. Davies MN, Flower DR (2007) Harnessing bioinformatics to discover new vaccines. *Drug Discov Today* 12:389–395
33. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113
34. Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549

35. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8:275–282
36. Källberg M, Wang H, Wang S et al (2012) Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 7:1511–1522
37. Heo L, Park H, Seok C (2013) GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res* 41:W384–W388
38. Lovell SC, Davis IW, Arendall WB III et al (2003) Structure validation by Ca geometry: ϕ , ψ and C β deviation. *Proteins: Struct Funct Bioinf* 50:437–450
39. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
40. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519
41. Walker JM (2005) *The proteomics protocols handbook*. Springer
42. Larsen MV, Lundegaard C, Lamberth K et al (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* 8:424
43. Moutaftsi M, Peters B, Pasquetto V et al (2006) A consensus epitope prediction approach identifies the breadth of murine T CD8⁺-cell responses to vaccinia virus. *Nat Biotechnol* 24:817–819
44. Jensen KK, Andreatta M, Marcatili P et al (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154:394–406
45. Ponomarenko J, Bui H-H, Li W et al (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* 9:514
46. Dhanda SK, Vir P, Raghava GPS (2013) Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct* 8:30
47. Bui H-H, Sidney J, Dinh K et al (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics* 7:1–5
48. Stratmann T (2015) Cholera toxin subunit B as adjuvant—–an accelerator in protective immunity and a break in autoimmunity. *Vaccines* 3:579–596
49. Yang J, Anishchenko I, Park H et al (2020) Improved protein structure prediction using predicted interresidue orientations. *Proceedings of the National Academy of Sciences* 117:1496–1503
50. Eisenberg D, Lüthy R, Bowie JU (1997) [20] VERIFY3D: assessment of protein models with three-dimensional profiles. In: *Methods in enzymology*. Elsevier, pp 396–404
51. Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 8:4
52. Magnan CN, Zeller M, Kayala MA et al (2010) High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics* 26:2936–2943
53. Dimitrov I, Bangov I, Flower DR, Doytchinova I (2014) AllerTOP v. 2—a server for in silico prediction of allergens. *J Mol Model* 20:2278
54. Dimitrov I, Naneva L, Doytchinova I, Bangov I (2014) AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* 30:846–851
55. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580
56. Nielsen H (2017) Predicting secretory proteins with SignalP. In: *Protein function prediction*. Springer, pp 59–73

57. Lamiable A, Thévenet P, Rey J et al (2016) PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res* 44:W449–W454
58. Van Zundert GCP, Rodrigues J, Trellet M et al (2016) The HADDOCK2. 2 web server: user-friendly integrative modeling of biomolecular complexes. *J Mol Biol* 428:720–725
59. Patel MC, Shirey KA, Pletneva LM et al (2014) Novel drugs targeting Toll-like receptors for antiviral therapy. *Future Virol* 9:811–829
60. Chen J, Ng MM-L, Chu JJH (2015) Activation of TLR2 and TLR6 by dengue NS1 protein and its implications in the immunopathogenesis of dengue virus infection. *PLoS Pathog* 11:e1005053
61. de Vries SJ, Bonvin AM (2011) CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLoS One* 6:e17695
62. Laskowski RA, Jabłońska J, Pravda L et al (2018) PDBsum: Structural summaries of PDB entries. *Protein Sci* 27:129–134
63. Abraham MJ, Murtola T, Schulz R et al (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1:19–25
64. Turner PJ (2005) XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research. Oregon Graduate Institute of Science and Technology, Beaverton
65. Grote A, Hiller K, Scheer M et al (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 33:W526–W531
66. Sharp PM, Li W-H (1987) The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15:1281–1295
67. Rapin N, Lund O, Bernaschi M, Castiglione F (2010) Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One* 5:e9862
68. Castiglione F, Mantile F, De Berardinis P, Prisco A (2012) 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
69. Baquero E, Albertini AA, Raux H et al (2015) Structure of the low pH conformation of Chandipura virus G reveals important features in the evolution of the vesiculovirus glycoprotein. *PLoS Pathog* 11:e1004756
70. Saadi M, Karkhah A, Nouri HR (2017) Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infect Genet Evol* 51:227–234
71. McQuarrie D (1976) *Statistical Mechanics*, Chap. 21
72. Adcock SA, McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* 106:1589–1615
73. Ghosh S, Dutta K, Basu A (2013) Chandipura virus induces neuronal death through Fas-mediated extrinsic apoptotic pathway. *J Virol* 87:12398–12406
74. Jadi RS, Sudeep AB, Barde PV et al (2011) Development of an inactivated candidate vaccine against Chandipura virus (Rhabdoviridae: Vesiculovirus). *Vaccine* 29:4613–4617
75. Sbai H, Mehta A, DeGroot AS (2001) Use of T cell epitopes for vaccine development. *Current drug targets-Infectious disorders* 1:303–313
76. Sette A, Fikes J (2003) Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. *Curr Opin Immunol* 15:461–470
77. Gershoni JM, Roitburd-Berman A, Siman-Tov DD et al (2007) Epitope mapping: the first step in developing epitope-based vaccines. *BioDrugs* 21:145–156

78. Iurescia S, Fioretti D, Fazio VM, Rinaldi M (2012) Epitope-driven DNA vaccine design employing immunoinformatics against B-cell lymphoma: a biotech's challenge. *Biotechnol Adv* 30:372–383
79. Gaafar B, Ali SA, Abd-elrahman KA, Almofti YA (2019) Immunoinformatics Approach for Multiepitope Vaccine Prediction from H, M, F, and N Proteins of Peste des Petits Ruminants Virus. *Journal of Immunology Research* 2019
80. Lu I-N, Farinelle S, Sausy A, Muller CP (2017) 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* 14:511–520
81. Jiang P, Cai Y, Chen J et al (2017) Evaluation of tandem Chlamydia trachomatis MOMP multi-epitopes vaccine in BALB/c mice model. *Vaccine* 35:3096–3103
82. Lennerz V, Gross S, Gallerani E et al (2014) Immunologic response to the survivin-derived multi-epitope vaccine EMD640744 in patients with advanced solid tumors. *Cancer Immunol Immunother* 63:381–394
83. Zhu S, Feng Y, Rao P et al (2014) Hepatitis B virus surface antigen as delivery vector can enhance Chlamydia trachomatis MOMP multi-epitope immune response in mice. *Appl Microbiol Biotechnol* 98:4107–4117
84. Lv X, Yang J, Song H et al (2014) Therapeutic efficacy of the multi-epitope vaccine CTB-UE against Helicobacter pylori infection in a Mongolian gerbil model and its microRNA-155-associated immuno-protective mechanism. *Vaccine* 32:5343–5352
85. Lin X, Chen S, Xue X et al (2016) Chimerically fused antigen rich of overlapped epitopes from latent membrane protein 2 (LMP2) of Epstein–Barr virus as a potential vaccine and diagnostic agent. *Cell Mol Immunol* 13:492–501
86. Li Y, Zheng K, Tan Y et al (2019) A recombinant multi-epitope peptide vaccine based on MOMP and CPSIT_p6 protein protects against Chlamydia psittaci lung infection. *Appl Microbiol Biotechnol* 103:941–952
87. Zhang L (2018) Multi-epitope vaccines: a promising strategy against tumors and viral infections. *Cell Mol Immunol* 15:182–184
88. Dar HA, Zaheer T, Shehroz M et al (2019) Immunoinformatics-Aided Design and Evaluation of a Potential Multi-Epitope Vaccine against Klebsiella Pneumoniae. *Vaccines* 7:88
89. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157:105–132
90. Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* 88:1895–1898
91. Foroutan M, Ghaffarifar F, Sharifi Z, Dalimi A (2020) 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* 69:101413
92. Bazhan SI, Antonets DV, Karpenko LI et al (2019) In silico designed ebola virus T-cell multi-epitope DNA vaccine constructions are immunogenic in mice. *Vaccines* 7:34
93. Boehme KW, Compton T (2004) Innate sensing of viruses by toll-like receptors. *J Virol* 78:7867–7873
94. Xagorari A, Chlichlia K (2008) Toll-like receptors and viruses: induction of innate antiviral immune responses. *Open Microbiol J* 2:49
95. Carty M, Bowie AG (2010) Recent insights into the role of Toll-like receptors in viral infection. *Clinical Experimental Immunology* 161:397–406
96. Anukumar B, Shahir P (2012) Chandipura Virus infection in mice: the role of toll like receptor 4 in pathogenesis. *BMC Infect Dis* 12:125
97. Vaure C, Liu Y (2014) A comparative review of toll-like receptor 4 expression and functionality in different animal species. *Front Immunol* 5:316
98. Tamura S-I, Funato H, Nagamine T et al (1989) Effectiveness of cholera toxin B subunit as an adjuvant for nasal influenza vaccination despite pre-existing immunity to CTB. *Vaccine* 7:503–505

99. Kim HJ, Kim J-K, Seo SB et al (2007) Intranasal vaccination with peptides and cholera toxin subunit B as adjuvant to enhance mucosal and systemic immunity to respiratory syncytial virus. *Arch Pharm Res* 30:366–371
100. Hou J, Liu Y, Hsi J et al (2014) Cholera toxin B subunit acts as a potent systemic adjuvant for HIV-1 DNA vaccination intramuscularly in mice. *Hum Vaccin Immunother* 10:1274–1283
101. Phongsisay V, Iizasa E, Hara H, Yoshida H (2015) Evidence for TLR4 and FcγR–CARD9 activation by cholera toxin B subunit and its direct bindings to TREM2 and LMIR5 receptors. *Mol Immunol* 66:463–471
102. Saeed U, Mazoor S, Jalal N, Piracha ZZ (2015) Contemplating the importance of Toll-like receptors I and II regarding human viral pathogenesis. *Jundishapur journal of microbiology* 8
103. Solanki V, Tiwari M, Tiwari V (2019) Prioritization of potential vaccine targets using comparative proteomics and designing of the chimeric multi-epitope vaccine against *Pseudomonas aeruginosa*. *Sci Rep* 9:1–19
104. Ali A, Khan A, Kaushik AC et al (2019) Immunoinformatic and systems biology approaches to predict and validate peptide vaccines against Epstein–Barr virus (EBV). *Sci Rep* 9:1–12
105. Pandey RK, Bhatt TK, Prajapati VK (2018) Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. *Sci Rep* 8:1–11
106. Kamthania M, Srivastava S, Desai M et al (2019) Immunoinformatics Approach to Design T-cell Epitope-Based Vaccine Against Hendra Virus. *Int J Pept Res Ther* 25:1627–1637
107. Ojha R, Pareek A, Pandey RK et al (2019) Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus zoonotic infection. *ACS omega* 4:13069–13079
108. Nezafat N, Ghasemi Y, Javadi G et al (2014) A novel multi-epitope peptide vaccine against cancer: an in silico approach. *J Theor Biol* 349:121–134
109. Parvizpour S, Razmara J, Pourseif MM, Omid Y (2019) In silico design of a triple-negative breast cancer vaccine by targeting cancer testis antigens. *BiolImpacts: BI* 9:45
110. Mishra S, Sinha S (2009) Immunoinformatics and modeling perspective of T cell epitope-based cancer immunotherapy: a holistic picture. *J Biomol Struct Dyn* 27:293–305

Figures

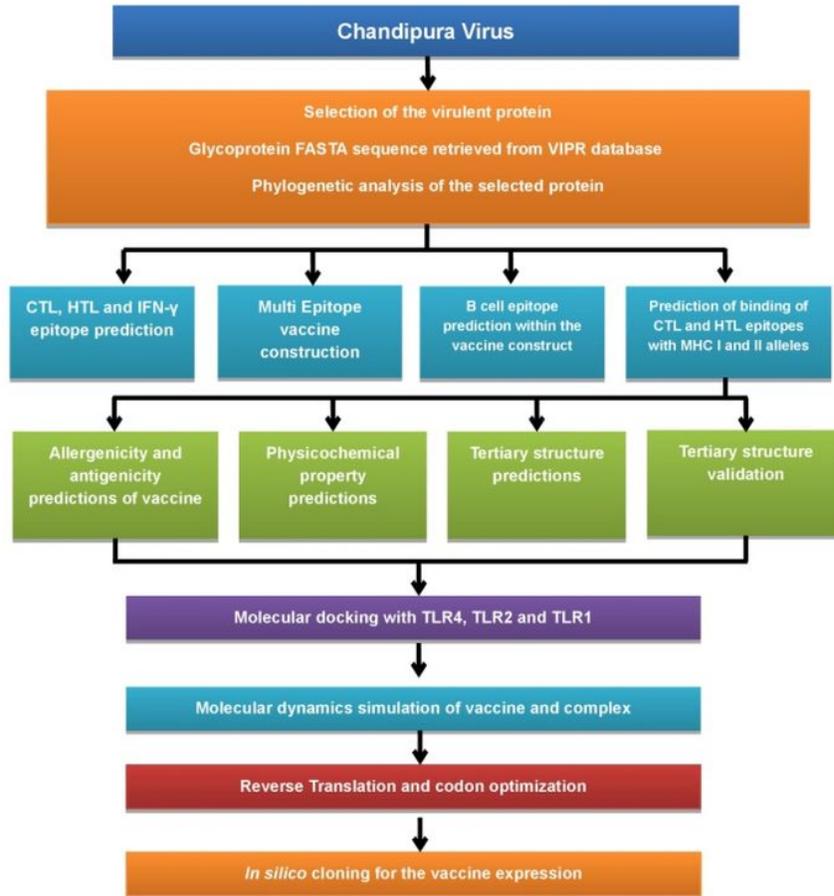


Figure 1

A systematic representation of all the steps followed in the designed study.

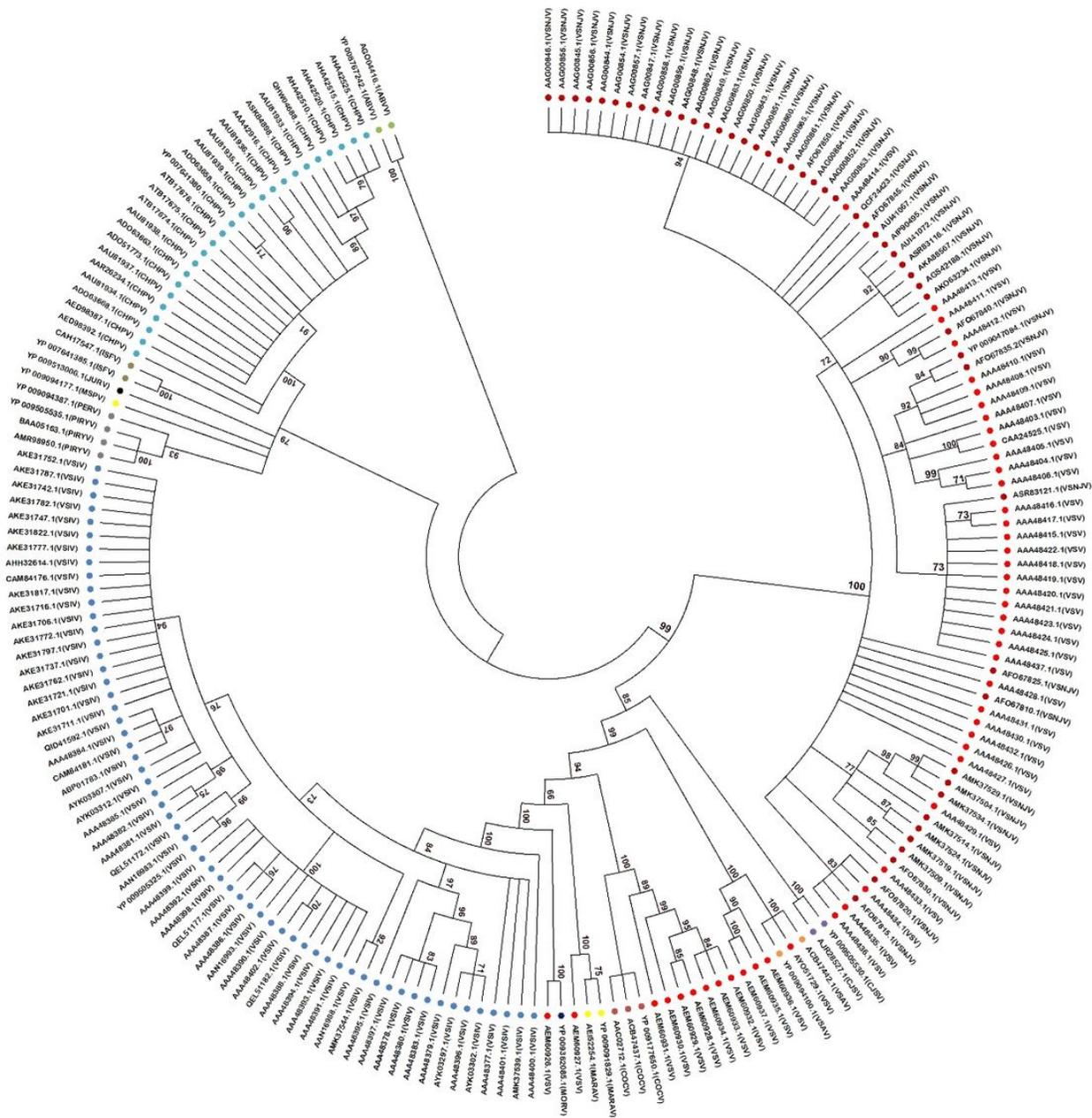


Figure 2

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.

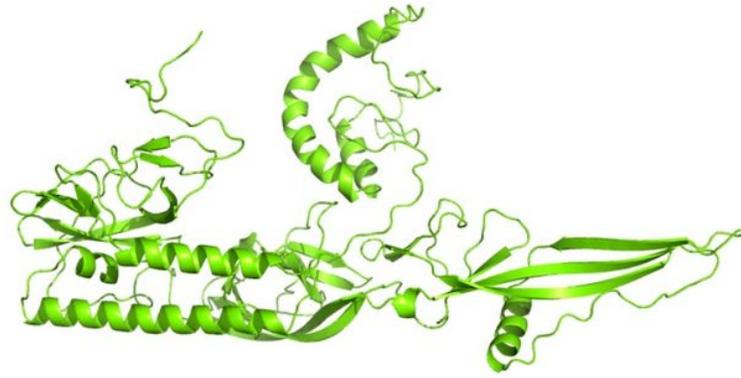


Figure 3

Modeled structure of CHPV glycoprotein with sequence length of 530 amino acids as predicted by RaptorX.

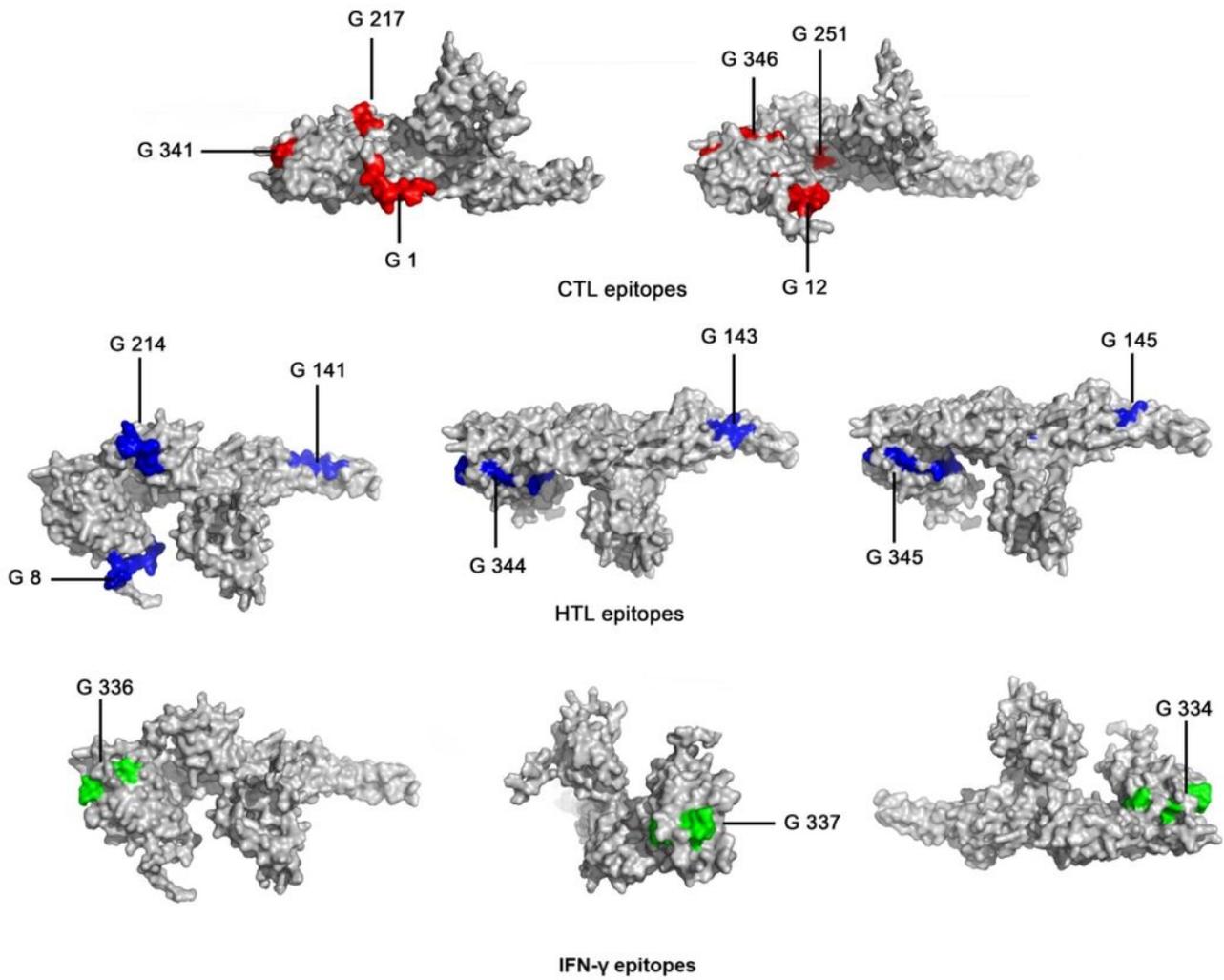


Figure 4

Glycoprotein 3D models showing the surface position of the epitopes included in the multi-epitope vaccine construct. CTL epitopes marked by red colour, HTL epitopes are marked by blue colour and IFN- γ epitopes are marked by green colour.

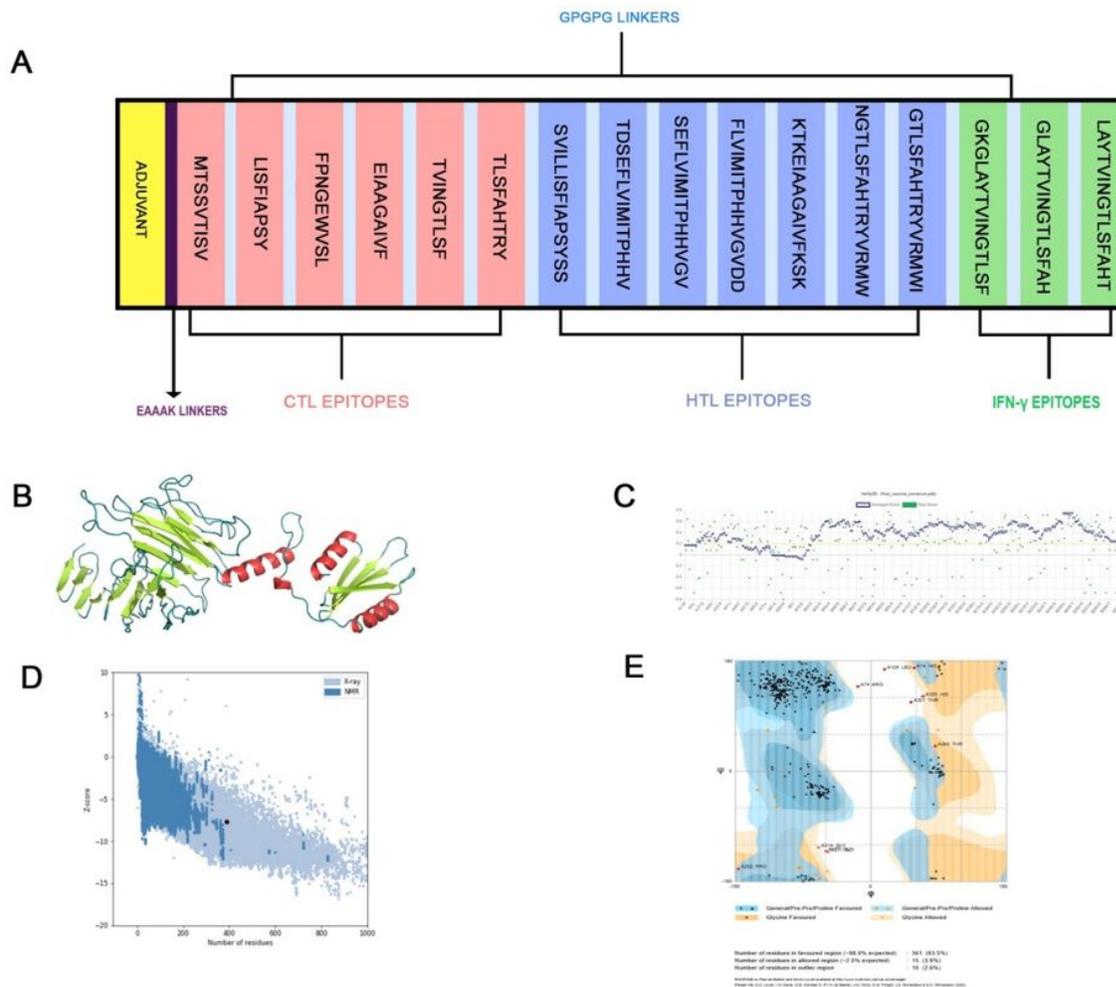


Figure 5

(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 colour represents the EAAAK linker and pale blue colour represents GPGPG linkers. (B) Tertiary structure of the final vaccine construct. Helices, sheets and loops are represented by red, light green and dark teal colour, respectively. (C) Verify3D 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.

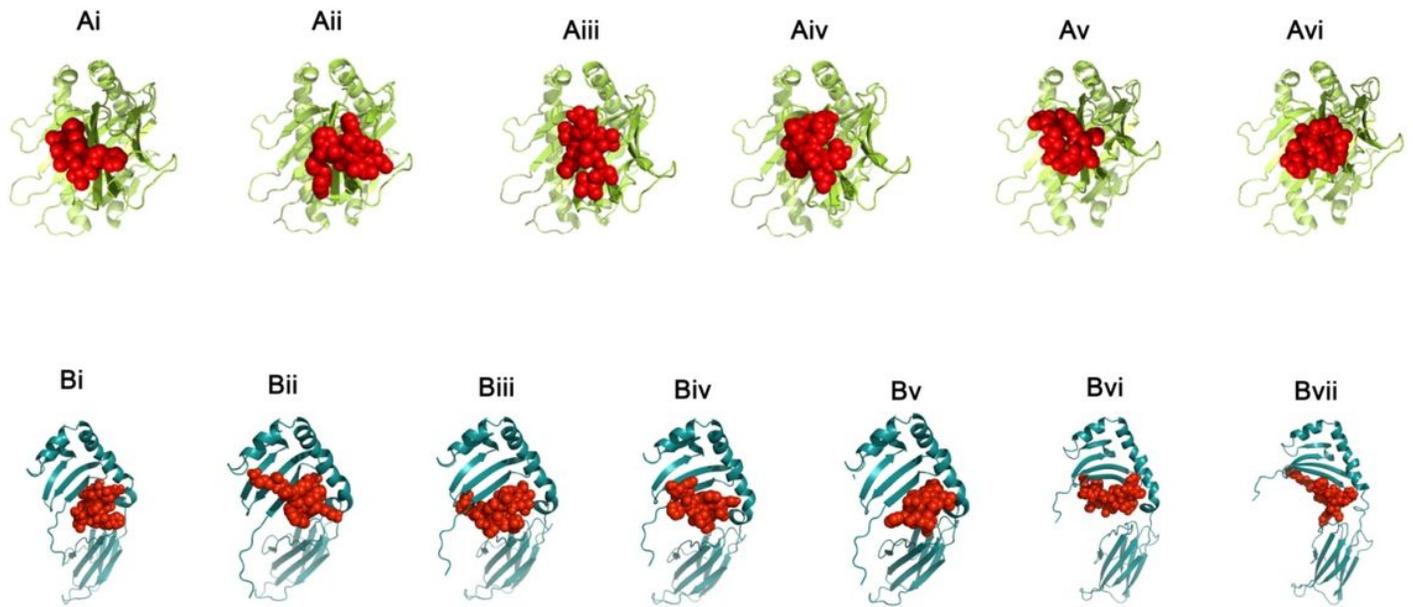


Figure 6

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.

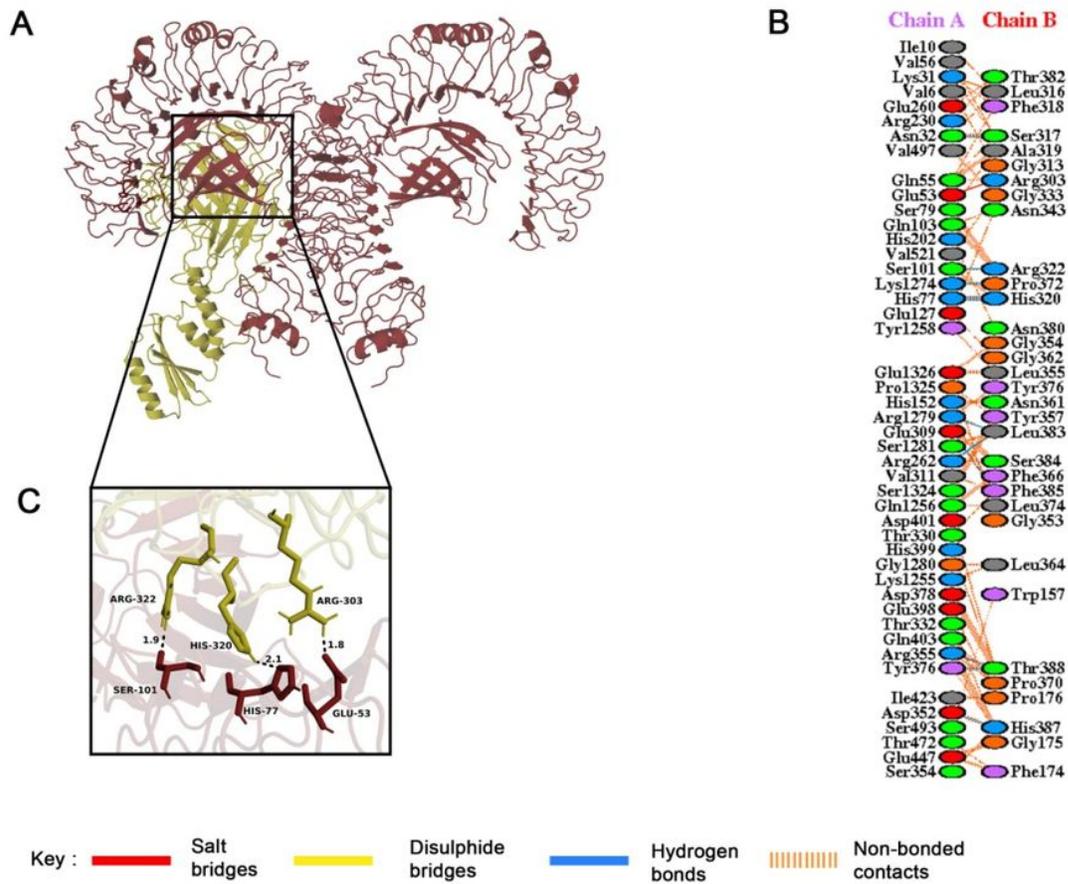


Figure 7

A) Molecular docking of vaccine-TLR4/MD2 complex. Vaccine construct and TLR4 receptor are shown in yellow and red colour, 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.

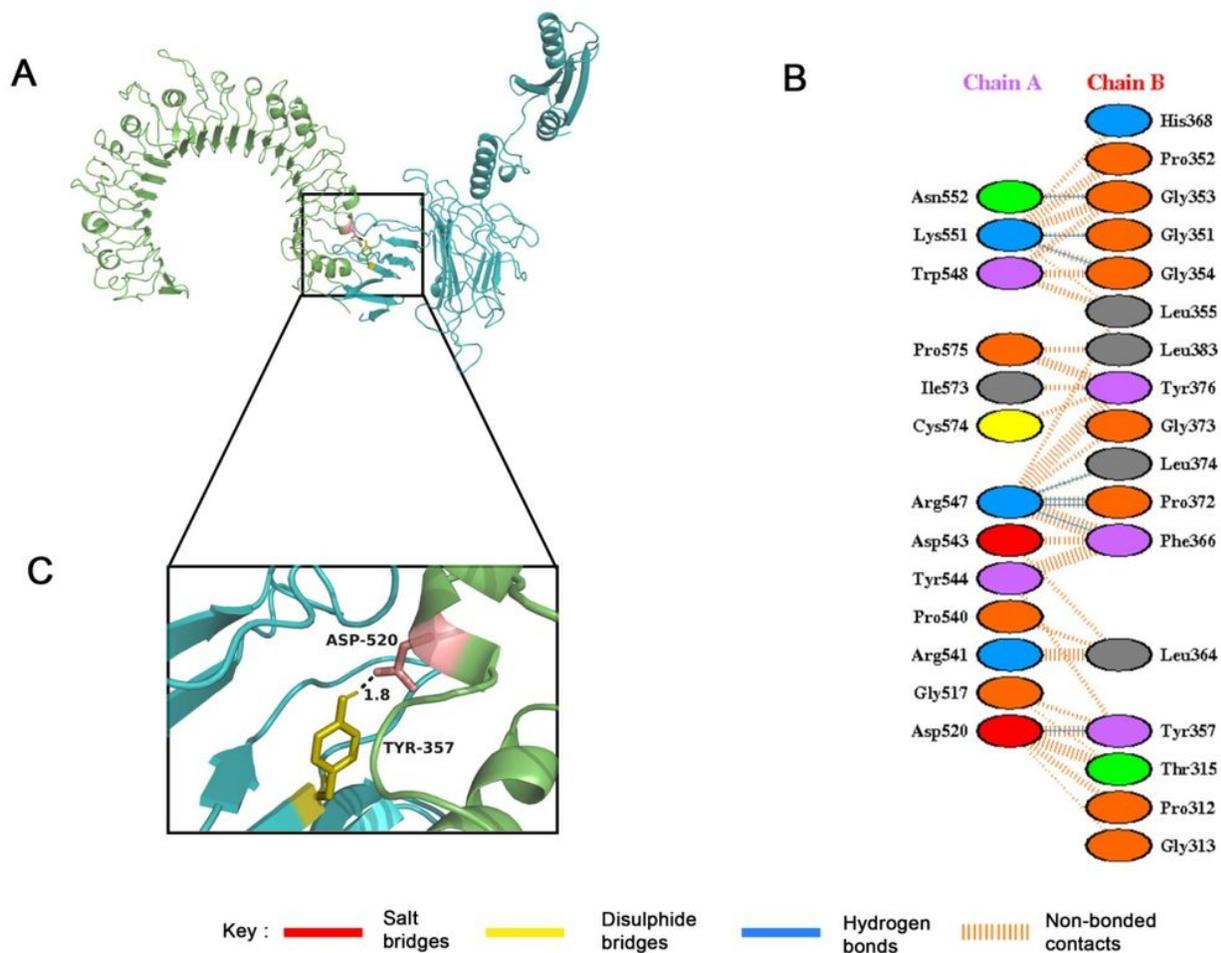


Figure 8

(A) Molecular docking of vaccine-TLR2 complex. Vaccine construct and TLR2 receptor are shown in blue and green colour, 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.

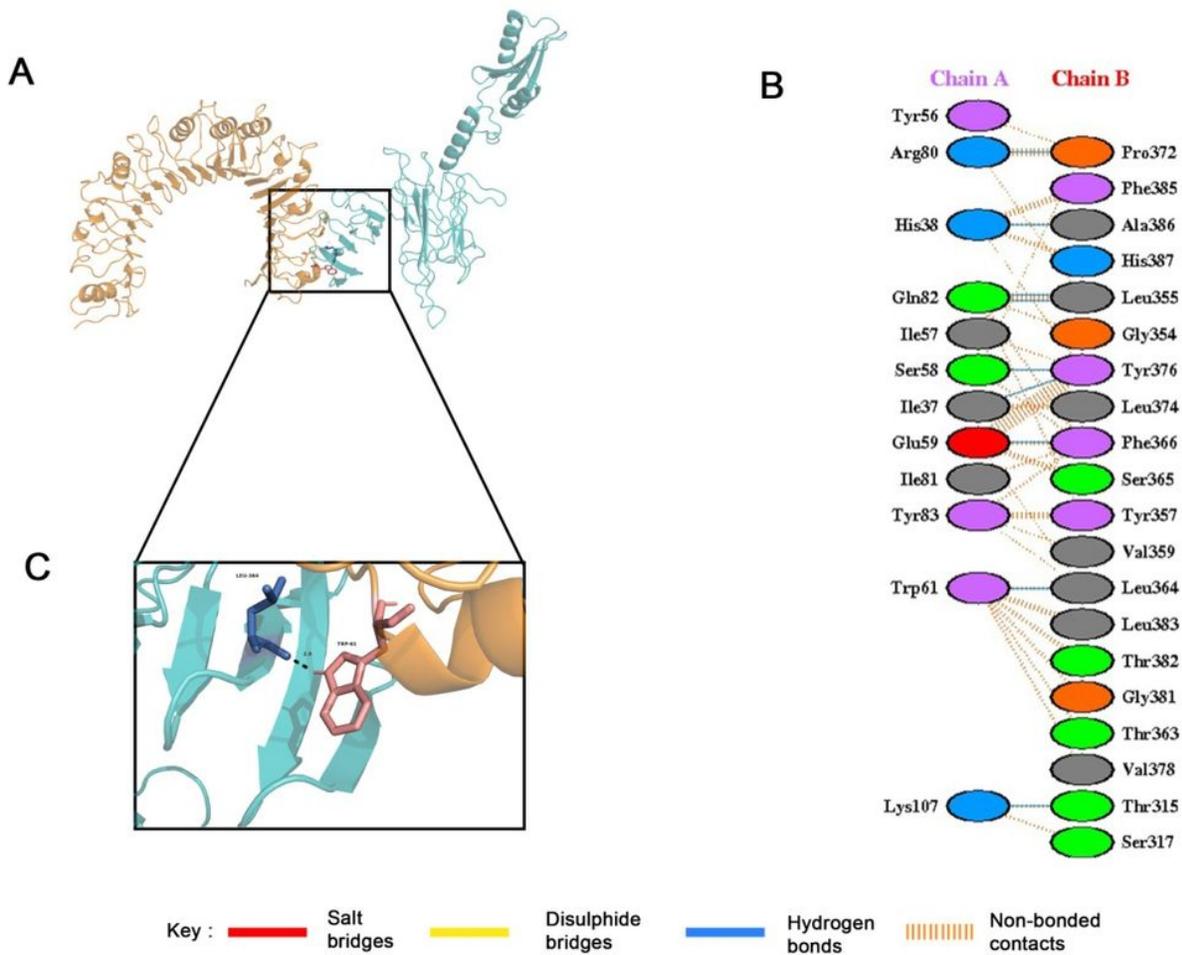


Figure 9

(A) Molecular docking of vaccine-TLR1 complex. Vaccine construct and TLR1 receptor are shown in blue and orange colour, 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.

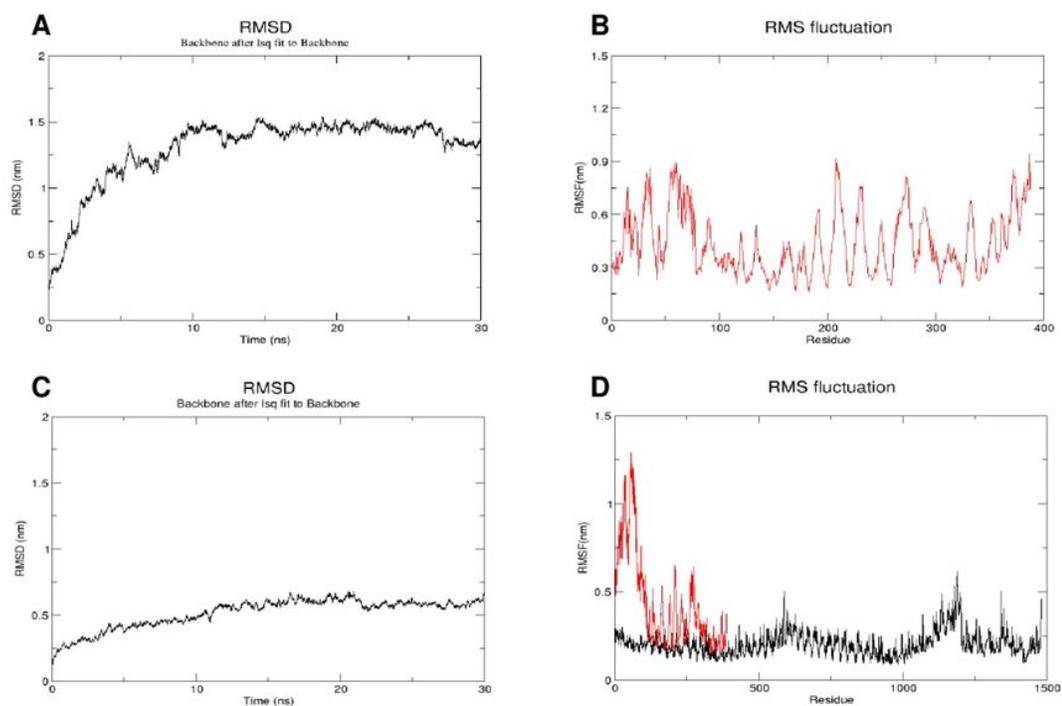


Figure 10

(A) RMSD plot of the vaccine backbone indicating stability. (B) RMSF plot of the vaccine (shown in red colour) with regions showing flexibility. (C) RMSD plot of the vaccine-TLR4/MD2 complex backbone with both the chains of TLR4 along with the co-receptor MD2, indicating stability. (D) RMSF plot of vaccine-TLR4/MD2 complex with both the chains of TLR4 along with the co-receptor MD2, indicating stability and flexibility. Vaccine is shown in red colour and TLR4 shown in black colour.

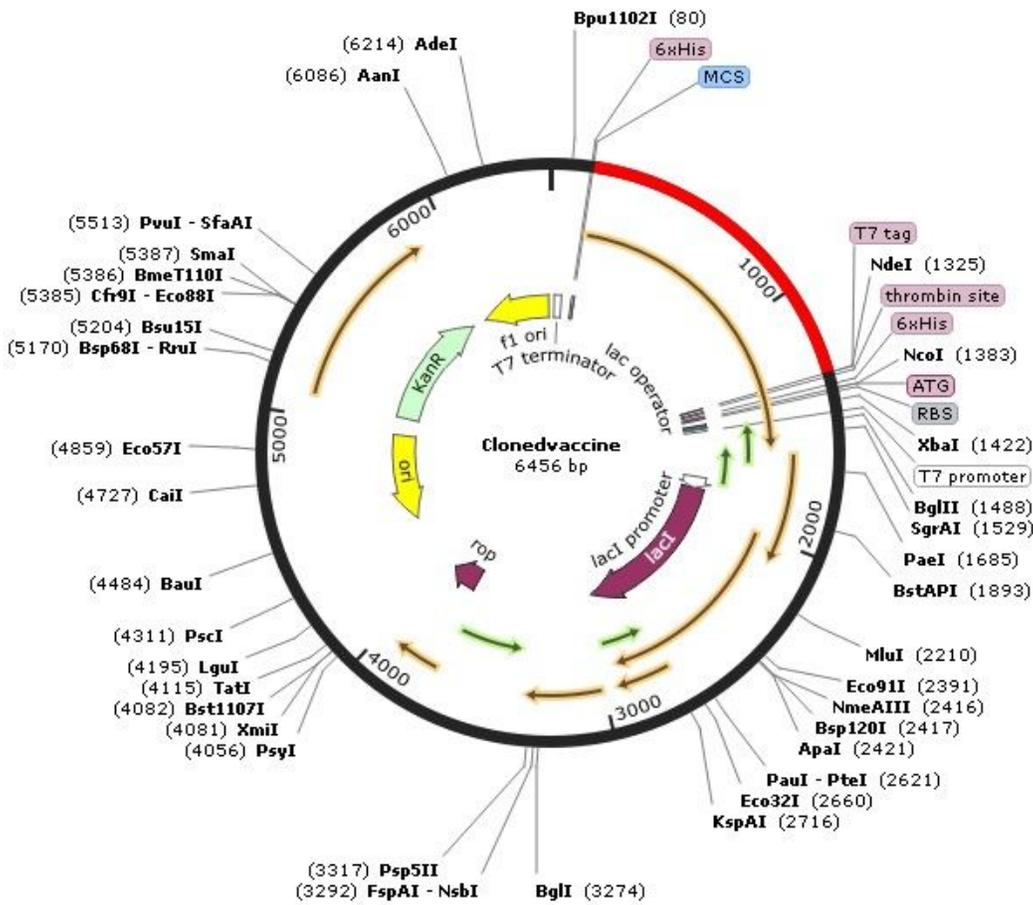


Figure 11

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

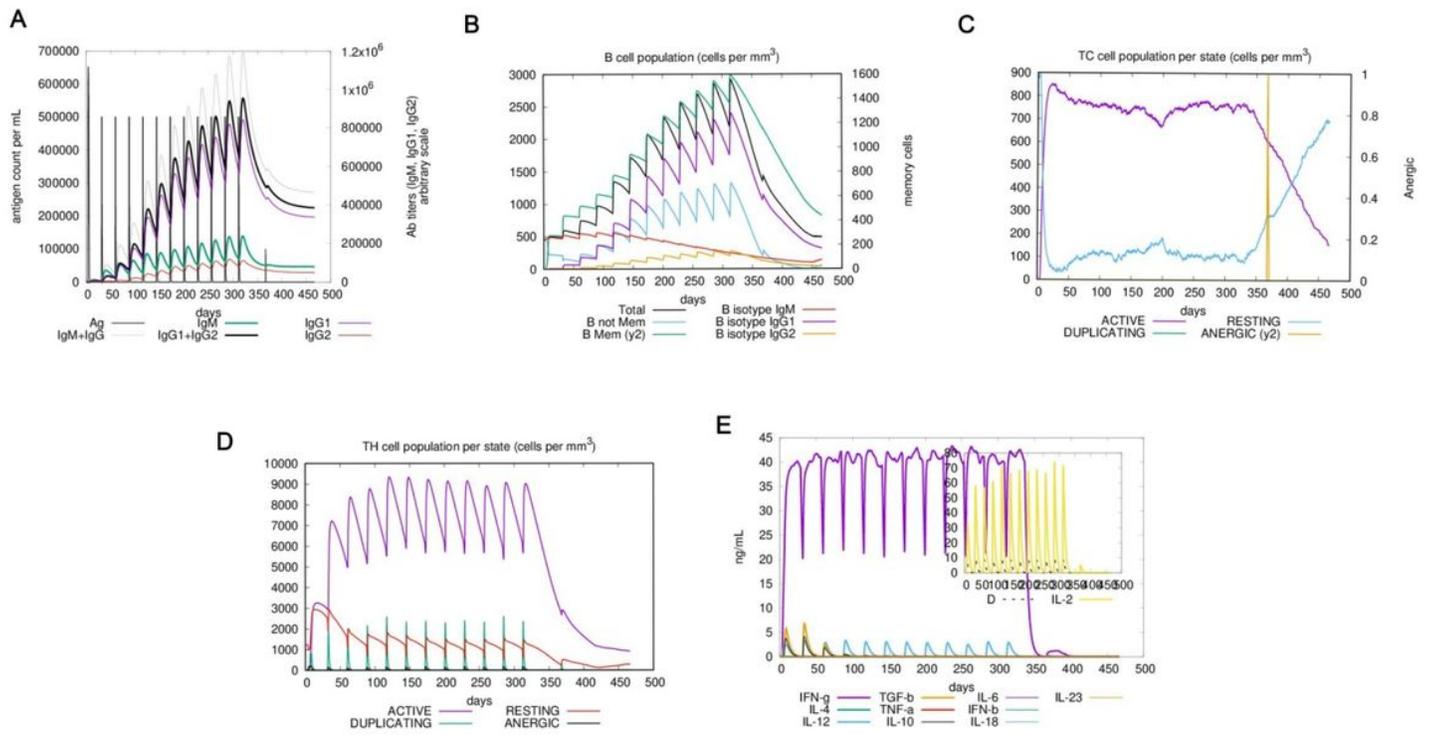


Figure 12

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.

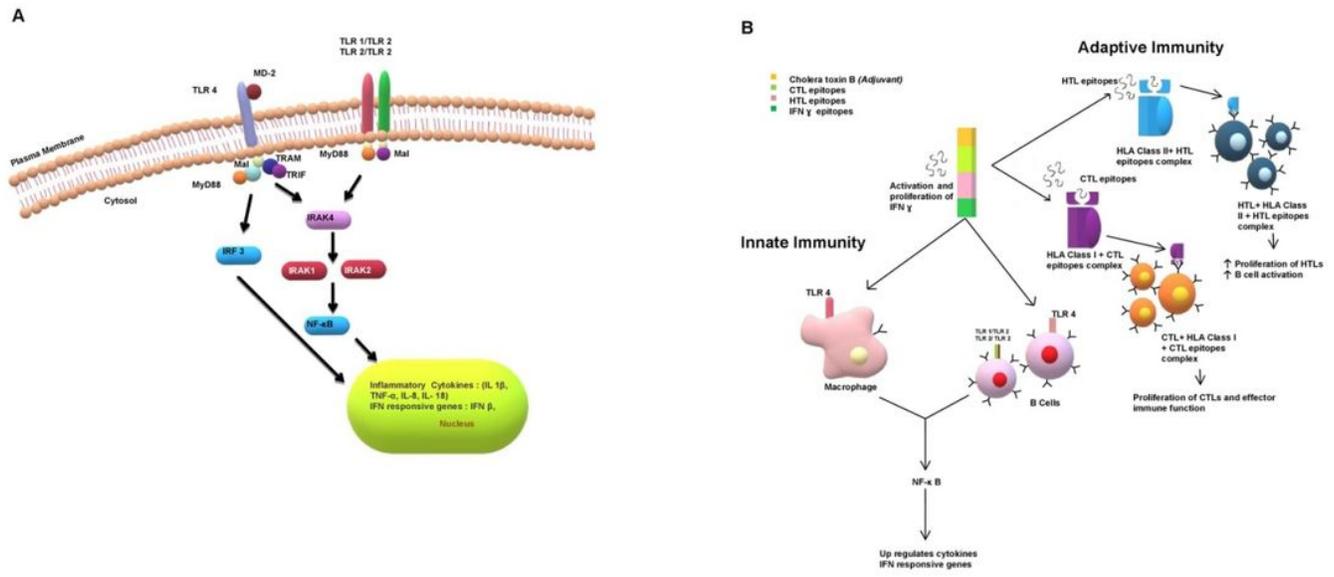


Figure 13

Proposed mechanism of working (A) TLR signal transduction pathway: TLR1/TLR2 heterodimer or TLR 2/2 homodimer utilizes MyD88 and MAL as primary adaptors to activate NF-κB that triggers inflammatory cytokine secretion. TLR4 uses four primary adaptors namely MyD88, MAL, TRIF and TRAM for NF-κB 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 up regulate 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.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [DebetaCHPVAdditionalFile1.docx](#)