



Mining of Ebola virus genome for the construction of multi-epitope vaccine to combat its infection

Uma Shankar , Neha Jain , Subodh Kumar Mishra , Md Fulbabu Sk , Parimal Kar & Amit Kumar

To cite this article: Uma Shankar , Neha Jain , Subodh Kumar Mishra , Md Fulbabu Sk , Parimal Kar & Amit Kumar (2021): Mining of Ebola virus genome for the construction of multi-epitope vaccine to combat its infection, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2021.1874529](https://doi.org/10.1080/07391102.2021.1874529)

To link to this article: <https://doi.org/10.1080/07391102.2021.1874529>



[View supplementary material](#)



Published online: 19 Jan 2021.



[Submit your article to this journal](#)



Article views: 3



[View related articles](#)



[View Crossmark data](#)



Mining of Ebola virus genome for the construction of multi-epitope vaccine to combat its infection

Uma Shankar, Neha Jain, Subodh Kumar Mishra, Md Fulbabu Sk Parimal Kar and Amit Kumar

Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology, Indore, Madhya Pradesh, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

Ebola virus is the primary causative agent of viral hemorrhagic fever that is an epidemic disease and responsible for the massive premature deaths in humans. Despite knowing the molecular mechanism of its pathogenesis, to date, no commercial or FDA approved multiepitope vaccine is available against Ebola infection. The current study focuses on designing a multi-epitope subunit vaccine for Ebola using a novel immunoinformatic approach. The best predicted antigenic epitopes of Cytotoxic-T cell (CTL), Helper-T cells (HTL), and B-cell epitopes (BCL) joined by various linkers were selected for the multi-epitope vaccine designing. For the enhanced immune response, two adjuvants were also added to the construct. Further analysis showed the vaccine to be immunogenic and non-allergenic, forming a stable and energetically favorable structure. The stability of the unbound vaccine construct and vaccine/TLR4 was elucidated via atomistic molecular dynamics simulations. The binding free energy analysis ($\Delta G_{\text{bind}} = -194.2 \pm 0.5 \text{ kcal/mol}$) via the molecular mechanics Poisson-Boltzmann docking scheme revealed a strong association and thus can initiate the maximal immune response. Next, for the optimal expression of the vaccine construct, its gene construct was cloned in the pET28a + vector system. In summary, the Ebola viral proteome was screened to identify the most potential HTLs, CTLs, and BCL epitopes. Along with various linkers and adjuvants, a multi-epitope vaccine is constructed that showed a high binding affinity with the immune receptor, TLR4. Thus, the current study provides a highly immunogenic multi-epitope subunit vaccine construct that may induce humoral and cellular immune responses against the Ebola infection.

ARTICLE HISTORY

Received 30 June 2020

Accepted 6 December 2020

KEYWORDS

Ebola; multi-epitope vaccine; hemorrhagic fever; immunoinformatics; molecular dynamics simulation; MM-PBSA; reverse vaccinology

1. Introduction

Ebola virus (EBOV), a member of the Filoviridae virus family, possess a negative-sense single-stranded RNA genome and causes a lethal Ebola virus disease (EOV) formerly known as Ebola hemorrhagic fever in humans and non-human primates (Bukreyev et al., 2014; Mahanty & Bray, 2004; Wang et al., 2006). The first outbreak of EBOV was reported in 1976 in South Sudan and the Democratic Republic of Congo with a high rate of mortality (88%) [<http://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease>]. After that, EBOV had been reported to spread throughout the world with increased pathogenicity and emerged as one of the most ominous pathogenic agents in the world (To et al., 2015) and has also been linked to its use in bioterrorism (Cenciarelli et al., 2015). A recent major outbreak of EBOV occurred in 2014–2015, with a total of 28,652 cases and 11,325 deaths worldwide (Ajisegiri et al., 2018). The second largest Ebola outbreak is in continuation since August 2018 in the Democratic Republic of Congo, with more than 2280 lives lost and 3450 confirmed cases. The highly contagious nature of infection and difficulty in accurate diagnosis makes the situation more ominous. Therefore, currently, along with the development of accurate diagnostic and effective

therapeutics, vaccine development against EBOV is at the top-most position in the priority list of the World Health Organization [<https://www.who.int/blueprint/priority-diseases/en/>].

The vaccine development against EBOV started immediately after its first outbreak but unfortunately could not achieve the expected success (Lupton et al., 1980). So far, the scientific community has mainly utilized viral glycoproteins (GP) for the development of a vaccine (Bounds et al., 2017; Dash et al., 2017; Khan et al., 2015). However, usage of such vaccines only activated the humoral immune response and failed to induce the cell-mediated immune response (Gross et al., 2018). The most successful attempt of vaccine development utilized the glycoprotein and nucleoproteins from three human-infecting EBOV viral strains, namely ZEBOV, SEBOV, and ICEBOV (Sullivan et al., 2000). However, this attempt also failed to activate both humoral and cellular immune responses (Sharma et al., 2017). Therefore, it became imperative for investigators to put their endeavors into the development of such a vaccine that could induce both the pathways of immune response and can effectively evacuate the pathogen from the body (Coller et al., 2017). A subunit vaccine with potentially immunogenic Cytotoxic T lymphocytes (CTLs), Helper T lymphocytes (HTLs), and B lymphocytes

epitopes and an innate immunity response activator would be a more successful approach to develop a vaccine with optimal efficacy against this highly infectious virus. The availability of whole genomes and the emergence of immunoinformatics have increased the pace of vaccine development. Reverse vaccinology, where the whole genome is critically analyzed by various tools for the identification of potential T and B cell epitopes, has become immensely popular in recent times and have shown great results in various human pathogens, including bacteria and viruses (Mora et al., 2003; Moxon et al., 2019). It has also been applied to the most recently occurring viruses like Nipah (Majee et al., 2020; Ojha et al., 2019) and SARS-CoV-2 (Kalita et al., 2020; Lizbeth et al., 2020; Samad et al., 2020).

The genome of the Ebola virus encodes for seven structural and non-structural proteins, namely Nucleoprotein (NP), Polymerase complex protein (VP35), Matrix protein (VP40), glycoprotein (GP), Minor nucleoprotein (VP30), membrane associated protein (VP24), and RNA-dependent RNA polymerase (L) (Feldmann et al., 1992). The VP30 is a major component of the virus nucleocapsid, and L protein plays a vital role in viral replication and transcription (Biedenkopf et al., 2013; Martinez et al., 2011). VP40 work as matrix protein and is solely responsible for the viral budding whereas, glycoprotein (GP) forms the two isoforms, i.e. soluble GP (sGP) and the structural spike GP that help in the infusion of EBOV to the host cells (Harty et al., 2000; Lee et al., 2017; Sanchez et al., 1993). In the present study, we analyzed the sequence and structure of all these seven proteins for their antigenicity, immunogenicity and predicted their respective B cell, HTL, and CTL epitopes by employing various immunoinformatics approaches. Adjuvants are chemical agents that boost the immune response of the host cell by binding to the Toll-like receptors (TLR) when added along with the vaccine (Reed et al., 2009). Previously, TLR4 has been reported for its binding with Ebola virus glycoproteins (GP), and thus, plays a vital role in immunopathogenesis by enhancing the pro-inflammatory cytokines and suppressor of cytokine signaling-1 (Okumura et al., 2010). A well-known agonist of TLR4, heparin-binding hemagglutinin was used as an adjuvant at the N terminal end of the subunit vaccine to maximize the host immune response (Zheng et al., 2017). Heparin-binding hemagglutinin is known to stimulate the dendritic cell maturation in TLR4 dependent manner and induces the expression of CD80, CD84, and CD40, both class-I and II MHC complexes, and pro-inflammatory cytokines (IL-6, IL-12, IL-1b, TNF- α , and CCR7) (Jung et al., 2011). The final construct was checked for its immunogenic and non-allergenic properties, and physiochemical properties of the constructed vaccine were predicted elucidate the half-life of the vaccine, instability index, theoretical pl, and grand average of hydropathicity. Tertiary structure analysis was performed to obtain the 3D structure of the constructed vaccine, and structural refinement was employed to check occupancy of maximum amino acids in the favored region of the Ramachandran plot. The molecular dynamic (MD) simulation confirmed the structural stability of the constructed vaccine. Moreover, docking studies revealed the tight interaction

between the constructed vaccine and TLR4 and MD simulation revealed the stability of the vaccine-TLR4 complex. Next, multi-epitope vaccine was used to construct the cDNA, codon optimized and inserted in a vector by using *in silico* cloning. Thus, present study provides a novel approach to develop the Ebola virus vaccine by utilizing both structural and non-structural proteins of the virus for preventing the EBOV infection in the human host.

2. Material and methods

2.1. Retrieval of antigenic Ebola proteome

The genome of EBOV synthesizes seven proteins. All seven protein sequences were retrieved from National Center for Biotechnology Information and were subjected to antigenic property analysis using Vaxigen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) server (Doytchinova & Flower, 2007).

2.2. Prediction of MHC-I specific cytotoxic T lymphocytes epitopes

Virus antigenic proteins were checked for cytotoxic T-cell epitopes by employing the NetCTL-1.2 server (<http://www.cbs.dtu.dk/services/NetCTL/>) (Larsen et al., 2007). NetCTL-1.2 predictions are based on an integrated prediction of TAP transporter efficiency associated with antigen processing, proteasomal C-terminal cleavage and MHC class-I binding that provides highly sensitive and top-scoring epitopes. All antigenic proteins of EBOV were submitted in FASTA format, and the data were collected for MHC type-I supertypes. The epitopes were filtered by using a threshold score of 0.75. The filtered epitopes were checked for their enhanced immunogenic properties using class-I Immunogenicity server available at The Immune Epitope Database (IEDB) [<http://tools.iedb.org/immunogenicity/>] (Calis et al., 2013).

2.3. Prediction of MHC-II specific helper T lymphocyte epitopes

Due to their involvement in the induction of both cellular and humoral immune response, HTL epitopes are essential for the development of effective immunotherapeutic vaccine development. The MHC-II prediction tool available at IEDB (<http://tools.iedb.org/mhcii/>) was employed for the prediction of 15-mer MHC-II binding epitopes. IEDB recommended parameter for the epitope prediction uses the consensus approach from various aligning methods like NN-align, SMM-align, CombLib, and Sturniolo. To get the high-affinity MHC-II binder, predicted epitopes were sorted by $IC_{50} \leq 50$ (nM) and the percentile rank.

Efficient immunity against a pathogen depends on the MHC-II based activation of both innate and adaptive immune response. During the viral infection, IFN- γ plays a crucial role in the activation of natural killer cells and macrophages. Therefore, the filtered HTL epitopes were evaluated for their ability to activate the IFN- γ using the IFNepitope server

available at <http://crdd.osdd.net/raghava/ifnepitope/scan.php> (Dhanda et al., 2013). IFNepitope is based on the support vector machine (SVM) algorithm and classifies the list of submitted epitopes into IFN- γ inducing or non-inducing MHC II binders.

2.4. Population coverage of the vaccine construct

The world population coverage of the vaccine construct was determined on the basis of frequency of occurrence for the respective HLA alleles. The population coverage tool of IEDB (<http://tools.iedb.org/population/>) that uses the allelefrequencies.net database was utilized to check the population coverage of the HLAs used in the analysis (Bui et al., 2006).

2.5. MHCs and epitope structural interaction analysis

The structure of all the available EBOV protein and HLA molecules were downloaded from the PDB database. The proteins that do not have known crystal structure were undergone structural modeling by employing the I-TASSER server and structural validation by the RAMPAGE server. I-TASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) provides the *ab-intio* prediction of protein structure (Yang et al., 2015). The most energetically favorable models were further refined using the GalaxyRefine server. Next, the selected epitopes were extracted from the respective PDB files and stored in PDB format using the PyMOL tool. To assess the binding propensity of the predicted epitope with the MHC I/II peptide-binding groove, the ClusPro docking server was exploited.

2.6. Prediction of B-cell epitopes

The BCPRED server (<http://ailab.ist.psu.edu/bcpred/>) was explored for the linear B cell epitope prediction (El-Manzalawy et al., 2008). It provides 20-mer length epitopes with a default specificity of 75%. All the EBOV protein sequences were submitted in FASTA format, and epitopes were predicted by a combination of both SVM and subsequence kernel approach.

2.7. Designing of multi-epitope subunit vaccine sequence

To construct the multi-epitope subunit vaccine, CTL, HTL, and B cell epitopes were combined using various linkers. Two previously known adjuvants, namely, Heparin-binding hemagglutinin and TpD were added to the vaccine construct for increased immunogenicity and the activation of TLR based cascade signaling.

2.8. Evaluation of allergenicity and antigenicity and physicochemical properties of subunit vaccine construct

The allergenicity of the constructed vaccine was predicted by employing three servers, namely AllerTop (<http://www.ddg-pharmfac.net/AllerTOP/>), AllergenFP (<http://www.ddg-pharmfac.net/AllergenFP/>) and AlgPred (<http://crdd.osdd.net/raghava/algpred/>) (Dimitrov et al., 2013, 2014; Saha & Raghava, 2006). AllerTop and AllergenFP use the k-nearest neighbor algorithm that is based on a training set of 2427 known allergens and non-allergens from different species. AlgPred predicts the allergen by various methods like SVM that uses the amino acid and dipeptide composition. The antigenicity of the peptide was predicted by VaxiJen v2.0 and ANTIGENPro (<http://scratch.proteomics.ics.uci.edu/>). (Cheng et al., 2005; Doytchinova & Flower, 2007)

The ProtParam Tool (<https://web.expasy.org/protparam/>) was used to predict various physicochemical properties of the designed vaccine such as molecular weight, theoretical pI, instability index, aliphatic index, and grand average hydrophobicity (GRAVY) value (Gasteiger et al., 2003).

2.9. Prediction of secondary and tertiary structure of the vaccine construct and its refinement

The secondary structure of the multi-epitope vaccine construct was predicted by PSIPRED server (<http://bioinf.cs.ucl.ac.uk/index.php?id=779>) (McGuffin et al., 2000). PSIPRED server uses two feed-forward neural networks and relies on output obtained by Position Specific Iterated-BLAST.

To model the complete tertiary structure of multi-epitope vaccine constructs, the I-TASSER server was utilized. I-TASSER relies on the confidence score (C-Score). Higher C-Score stands for higher accuracy in the generated model. On the basis of the C-score, the Ramachandran plot was constructed for the top five models. The model with the maximum residues in the favorable region of the Ramachandran plot and higher C-Score was used for further analysis.

The GalaxyRefine (<http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>) and ModRefiner (<https://zhanglab.ccmb.med.umich.edu/ModRefiner>) servers were used to refine the structure obtained from the I-TASSER (Heo et al., 2013; Xu & Zhang, 2011). The results from GalaxyRefine and ModRefiner were analyzed by PROCHECK (<http://servicesn.mbi.ucla.edu/PROCHECK/>) server (Laskowski et al., 1993). PROCHECK provides detailed information about the stereochemistry of a protein structure and uses five-set of programs for generating the Ramachandran plot.

2.10. Molecular docking of multi-epitope vaccine construct with TLR4

ClusPro (<https://cluspro.bu.edu/login.php>) was utilized for the docking analysis between the complex of multi-epitope vaccine construct and TLR4 molecule (Kozakov et al., 2017). The structure of TLR4 was obtained from the PDB database (PDB ID: 3FXI).

2.11. Stability of the multi-epitope vaccine constructs and vaccine-TLR4 complex

For analyzing the stability of the multi-epitope vaccine construct and the vaccine TLR4 construct molecular dynamics

(MD) simulation was performed by using the *pmemd.cuda* module of AMBER18 (Case et al., 2018) package. AMBER ff14SB force field was used to describe TLR4 and multi-epitope vaccine constructs as it accurately predicts the parameters for backbone and the side-chains of a protein (Maier et al., 2015). Missing hydrogens were added by using the *tLEaP* module of AMBER18. The multi-epitope vaccine construct and TLR-4/vaccine complex were solvated in a 10 Å octahedron periodic box with an explicit TIP3P water model (Price & Brooks, 2004). Chloride ions (Cl^-) were added to neutralize both the systems. The temperature of the system was maintained at 310K using a Langevin thermostat (Pastor et al., 1988). The SHAKE algorithm (Kräutler et al., 2001) was used to constraint the movement of bond lengths involving hydrogen atoms. For the consideration of long-range electrostatic forces, Particle Mesh Ewald (PME) (Darden et al., 1993) was incorporated to generate periodic boundary conditions. Firstly, each system was energy minimized by 500 steps of the steepest descent algorithm followed by another 500 steps of the conjugate gradient method. The temperature of each system was raised from 0K to 300K in multiple stages. Each system was equilibrated for 1 ns in the NVT ensemble. After that, a 100 ns NPT production run was conducted where a constant pressure of 1.0 bar was maintained using Berendsen Barosat (Berendsen et al., 1984). The Cartesian coordinates were stored every 10 ps, and a total of 10,000 snapshots for each production simulation were accumulated. A similar protocol was used in our previous studies (Jonniiya et al., 2019; Singh et al., 2020; Sk et al., 2020a).

All analyses, such as root-mean-squared deviation (RMSD), root-mean-squared fluctuation (RMSF), the radius of gyration (R_g), solvent accessible surface area (SASA), and hydrogen bonds between TLR4 and vaccine analysis, were performed by using the *Cpptraj* module (Roe & Cheatham, 2013). The final structure obtained from the production simulation was employed to construct the interaction profile between TLR4 and vaccine utilizing the Dimplot module of LigPlot++ (Wallace et al., 1995).

2.12. MM-PBSA and critical residue analysis of the multi-epitope vaccine construct—TLR4 complex

To analyze the stability of the multi-epitope vaccine—TLR4 complex, the binding free energy of the complex formation was calculated using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method (Gohlke et al., 2003; Gohlke & Case, 2004; Homeyer & Gohlke, 2012; Jonniya et al., 2020; Kollman et al., 2000; Sk et al., 2020b, 2020c, 2020d). Briefly, the MM-PBSA.py (Miller et al., 2012) script was utilized to calculate the free energy of the docked complex by using the following equations:

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{internal}} + \Delta G_{\text{solv}} - T\Delta S \quad (1)$$

$$\Delta E_{\text{internal}} = \Delta E_{\text{covalent}} + \Delta E_{\text{elec}} + \Delta E_{\text{vdW}} \quad (2)$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{pol}} + \Delta G_{\text{np}} \quad (3)$$

where $\Delta E_{\text{internal}}$, ΔG_{solv} , and $T\Delta S$ represent the total internal energy, solvation free energy, and conformational entropy, respectively. The internal energy is further composed of

$\Delta E_{\text{covalent}}$ (bond, dihedral, and angle), ΔE_{elec} (electrostatic) and ΔE_{vdW} (van der Waals). On the other hand, desolvation free energy is composed of polar (ΔG_{pol}) and non-polar (ΔG_{np}) free energies. A total of 1000 structural frames at a frequency of 4 ps from the last 40 ns trajectory was used for the enthalpy (ΔH) calculations. Due to the high computational cost, we avoided the estimation of conformational entropy (ΔS). To find the critical residues involved in the complex stabilization, per-residue decomposition was performed using the MM-GBSA method (Gohlke et al., 2003; Kar & Knecht, 2012a, 2012b, 2012c, 2012d).

2.13. cDNA construct, codon optimization and *in silico* cloning of designed vaccine

The Reverse translate tool available at EXPASY server and a Java Codon Adaptation Tool was employed to generate the cDNA construct of the designed multi-epitope vaccine for its optimal expression in *E. coli*. The obtained cDNA was further analyzed for the presence of rho-independent terminator sites and absence of *Xba*I, and *Bam*H I restriction sites in its ORF region. Next, SnapGene was used for *in silico* cloning of gene construct in the pET28a(+) vector expressing the designed multi-epitope vaccine.

3. Results and discussion

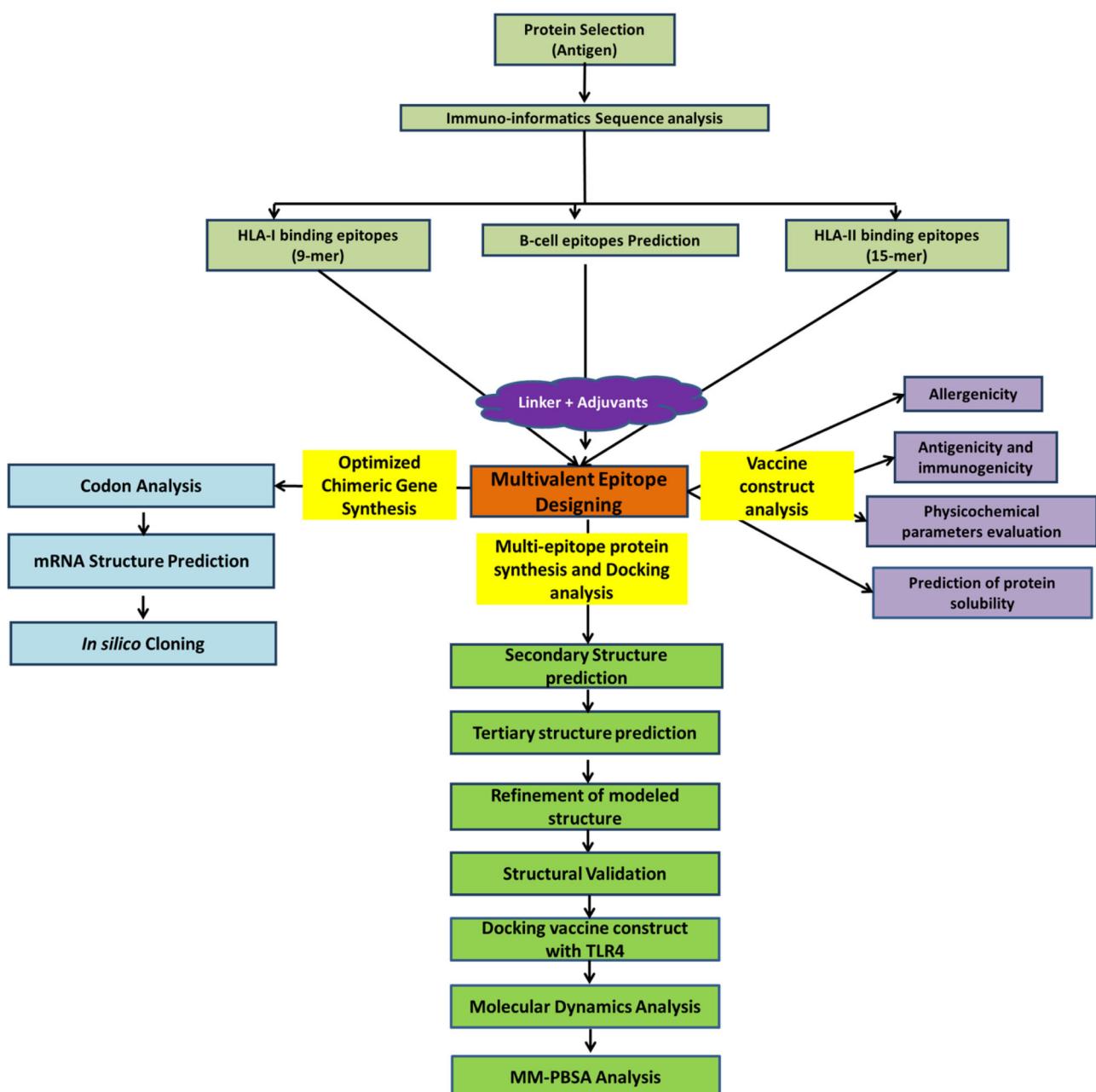
3.1. Retrieval of EBOV protein sequences

The genome of EBOV encodes for seven structural and two nonstructural proteins. Structural proteins include Nucleoprotein (NP), Polymerase complex protein (VP35), Matrix protein (VP40), Spike glycoprotein (sGP), Minor nucleoprotein (VP30), Membrane-associated protein (VP24), and RNA-dependent RNA polymerase (L). In contrast, non-structural proteins include Small secreted glycoprotein (ssGP), and secreted glycoprotein (GP). The NP, VP30, VP35, and L protein plays an active role in transcription and replication of virus genome inside the host cell while transmembrane glycoproteins, VP24, and VP40 are associated with the formation of virus membrane. The virus glycoproteins bind and fuse with the cell surface receptor and facilitate its entry into the host cell. The amino acid sequence of all these EBOV proteins was retrieved from the NCBI database and checked for their antigenicity on the Vaxigen server (Table 1).

In Vaxigen, a protein that possesses antigenic scores more than 0.4 stands for high antigenicity for viruses. In the analysis of antigenicity, all the nine EBOV proteins possess the antigenic score more than the threshold cutoff (0.4) that suggests the high antigenicity of all the viral proteins. Therefore, all these proteins were chosen for CTL, HTL, and B cell epitopes prediction by utilizing various available online tools. A schematic representation of the strategy followed is depicted in Figure 1.

Table 1. Prediction of antigenicity of Ebola virus protein from Vaxigen server.

#	Gene name	Protein name	Function	NCBI protein accession number	Vaxigen analysis (Threshold—0.4)
1	NP	Nucleoprotein	Encapsidation of genomic RNA	NP_066243.1	0.4448 (Probable antigen)
2	VP35	Polymerase complex protein	RNA-dependent RNA polymerase cofactor	NP_066244.1	0.5098 (Probable antigen)
3	VP40	Matrix protein	Coalesce nucleocapsids and cell membranes in virion assembly (budding)	NP_066245.1	0.5138 (Probable antigen)
4	GP	Spike glycoprotein	Receptor binding and fusion	NP_066246.1	0.4919 (Probable antigen)
5		Small secreted glycoprotein		NP_066247.1	0.4675 (Probable antigen)
6		Second secreted glycoprotein		NP_066248.1	0.5327 (Probable antigen)
7	VP30	Minor nucleoprotein	Encapsidation of genomic RNA	NP_066249.1	0.5203 (Probable antigen)
8	VP24	Membrane-associated protein	Interferon antagonism, transcriptional and replication control during infection	NP_066250.1	0.4787 (Probable antigen)
9	L	RNA-dependent RNA polymerase	Synthesis of viral RNAs; transcriptional RNA editing	NP_066251.1	0.4088 (Probable antigen)

**Figure 1.** Flowchart of the immuno-informatic strategy used for the EBOV vaccine construct analysis.

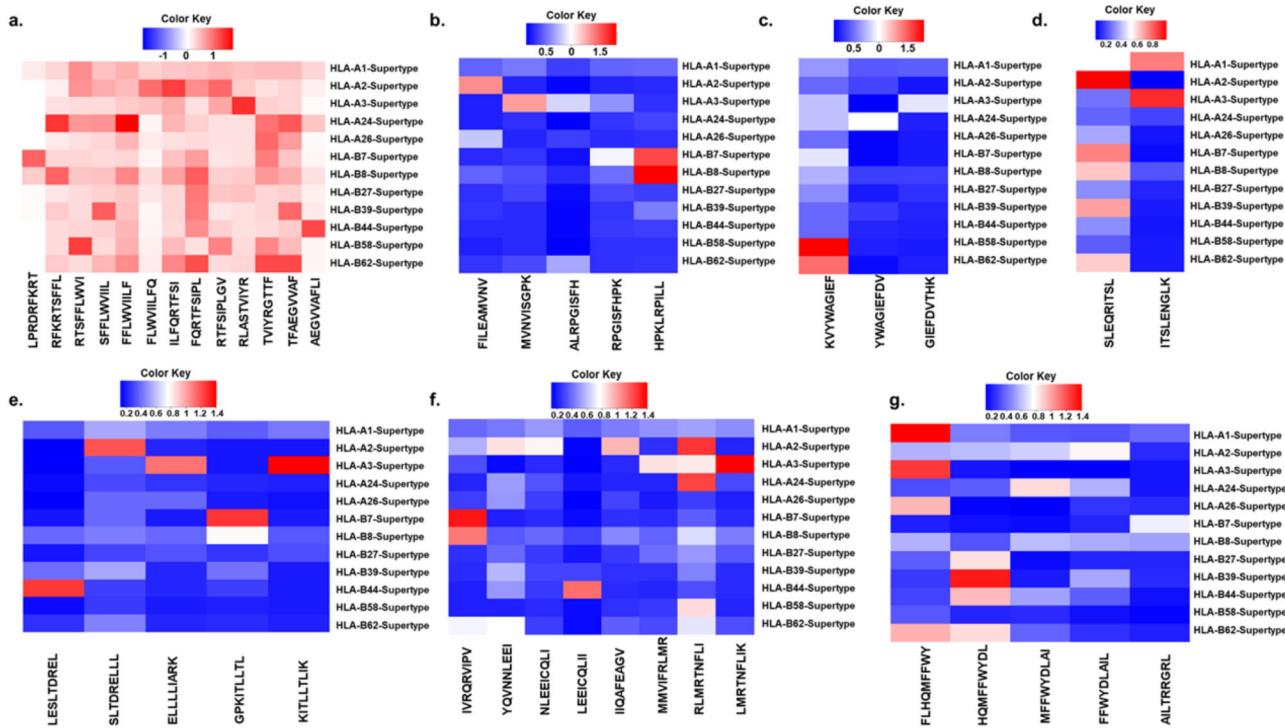


Figure 2. Heatmaps depicting the affinity of CTL epitopes with 12 HLA class-I molecules that covers >99% of the world population for the following Ebola proteins: (a) Glycoprotein (GP), (b) Matrix protein (VP40), (c) Membrane associated protein (VP24), (d) Polymerase complex protein (VP35), (e) Minor Nucleoprotein (VP30), (f) Nucleoprotein (NP), (g) RNA polymerase (L).

3.2. Prediction of MHC-I specific CTL epitopes

Cytotoxic T lymphocytes are the major components of the adaptive immune response. The CTLs are known to be activated by interacting with antigens that are processed and presented by the MHC class-I complex on the surface of virus-infected cells. The NetCTLs server was used for retrieving the CTL epitopes from EBOV proteins. This server uses artificial neural networks to predict 9-mer epitopes that harbor a high propensity to proteasomal cleavage at its C terminal end and high affinity towards MHC class-I molecules. This server also employs a weight matrix to determine the transportation efficiency of the predicted epitope by Transporter associated with antigen processing (TAP) protein. In the present study we used twelve different alleles of MHC-I namely: HLA-A*01:01, HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*24:02, HLA-B*07:02, HLA-B*08:01, HLA-B*27:02, HLA-B*27:05, HLA-B*35:01, HLA-B*35:03, HLA-B*44:02, HLA-B*44:03, HLA-B*58:01 HLA-B*58:02 due to their existence in >90% of world population. To obtain high epitope-HLA interaction specificity of more than 97%, the NetCTLs server's score was set as 0.75. Given that immunogenicity of vaccine is essential for maximal immune response, Class-I Immunogenicity server available at IEDB was employed to check the immunogenicity of the filtered epitopes obtained from NetCTLs analysis. The tool analysis 9-mer epitope and provides a positive or negative score. The higher positive score stands for the higher immunogenicity of the peptides, whereas a negative score depicts low immunogenic peptides. The heatmap depicting the combined score of HLA-I allele-CTL epitope

pair is displayed in Figure 2. The epitopes that possess the positive score were further evaluated for their specificity to the maximum number of MHC class-I supertypes. Overlapping sequences were merged and used for vaccine construction. This analysis provided one epitope for each VP24, VP35, and L protein, whereas two epitopes for each VP40, sGP, VP30, and NP proteins (Table S1, supplementary material).

3.3. Prediction of MHC-II specific helper T lymphocyte (HTL) epitopes

HTL epitopes are important for both the humoral and cellular immune responses. MHC-II specific HTL epitope (15-mers) were predicted by the IEDB tool. To cover the maximum population throughout the world, a set of twenty three most frequent MHC-II molecules namely: HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*05:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*01:01/DQB1*05:01, HLA-DQA1*01:02/DQB1*03:02, HLA-DQA1*04:01/DQB1*04:02, HLA-DQA1*05:01/DQB1*02:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*01:01, HLA-DRB1*03:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*12:01, HLA-DRB1*13:02, HLA-DRB1*15:01 that covers 99.99% of the world population were used. The epitopes were filtered on the basis of IC₅₀ and percentile rank. The IC₅₀ values ≤ 50 nM signify high-affinity epitopes, whereas the percentile rank is inversely

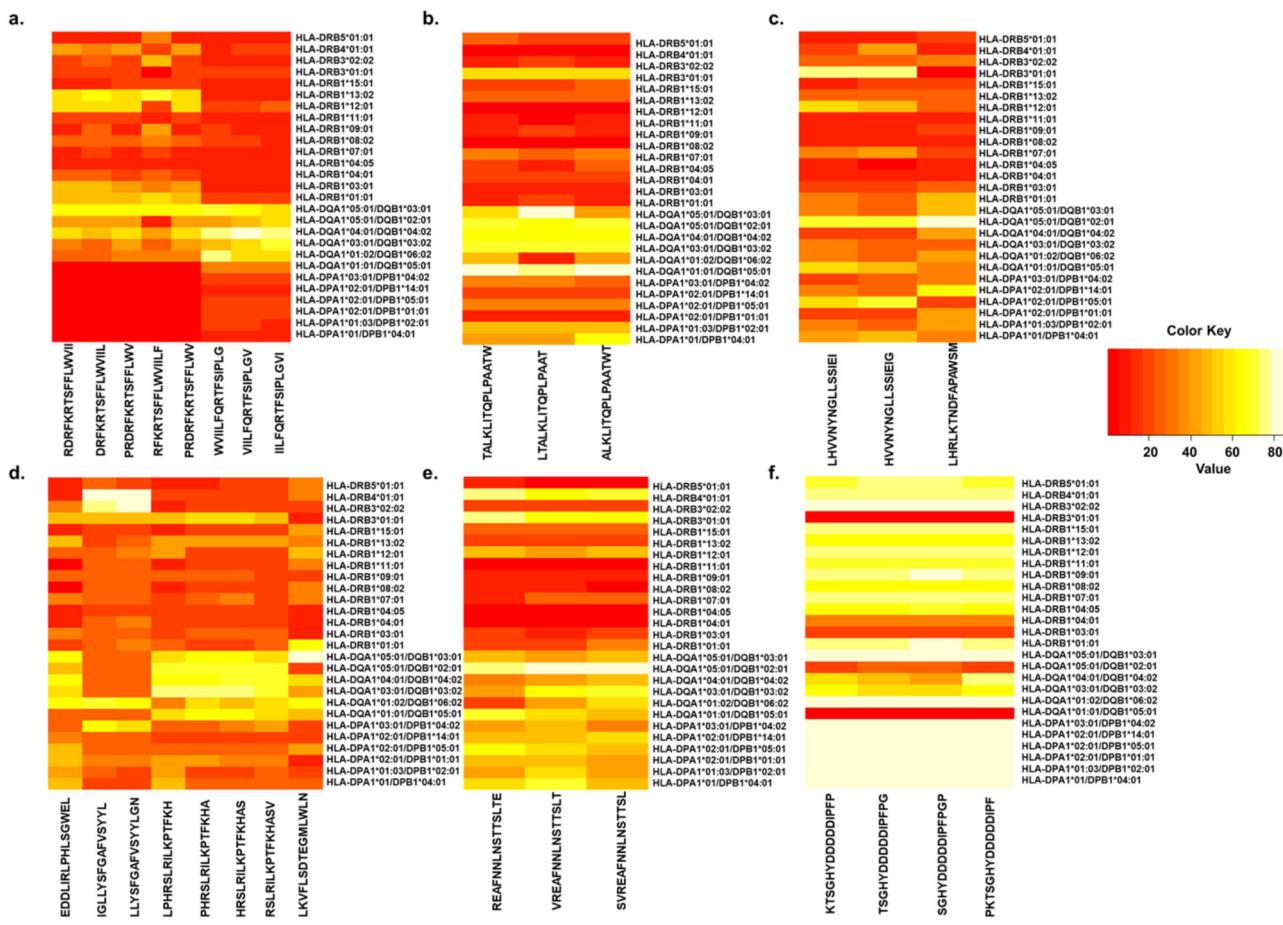


Figure 3. Heatmaps depicting the affinity of HTL epitopes with 27 HLA class-II molecules that covers >99% of the world population for the following Ebola proteins: (a) Glycoprotein (GP), (b) Matrix protein (VP40), (c) Membrane associated protein (VP24), (d) RNA polymerase (L), (e) Polymerase complex protein (VP35), (f) Nucleoprotein (NP).

proportional to the binding affinity. Thus, the epitopes having the minimum IC₅₀ and the lowest percentile rank for each MHC-II molecule were selected for further analysis. The heatmap depicting the percentile rank of the HLA-class-II alleles and HTL epitope pairs for EBOV proteins are depicted in Figure 3 (also see Table S2a-f, supplementary material). In this analysis, we found one best HLA-epitope pair for NP, VP40, VP24, and VP35, two for sGP, and four best HLA-epitope sets for L protein (Table S3, supplementary material).

The type II interferon-gamma (IFN- γ) is a regulator of immunity and acts as a connecting link between innate and adaptive immune responses. IFN- γ synthesized by the naive helper T cells is required for its differentiation into the Th1 population, immunoglobulin class switching, and viral load clearance in the CNS region of humans. IFN- γ expression is found to be increased when HTL interacts with the MHC-II molecule presenting virus antigen. Therefore for the identification of MHC-II epitopes that are able to induce the IFN- γ synthesis can be a better subunit vaccine target. The IFNepitope tool was used to confirm the ability of the selected MHC-II epitopes to induce the production of IFN- γ . Out of the ten selected MHC-II binders, three epitopes showed a negative score and, therefore excluded from further analysis (Table S3, supplementary material). Thus, on the basis of IC₅₀, percentile rank, and ability to induce the IFN- γ

production, six HTL epitopes were selected and added to the vaccine construct (Table S3, supplementary material).

3.4. Molecular interaction of the HTL and CTLs epitopes with the HLA complexes

To check the interaction between selected HTL/CTL epitopes and HLA molecules, docking analysis was performed by employing the ClusPro protein docking server. All the epitopes showed the energetically favorable interaction with the MHC-I and MHC-II complex structures and bind to the peptide-binding groove of the HLA (Figures S1-S8, supplementary material). The molecular interaction of the epitopes with the HLAs further affirms the high affinity of the epitopes for the helper and cytotoxic T cells.

3.5. Linear B cell epitope prediction

B cell epitopes act as an antigenic determinant of a pathogenic infection inside the body and thus plays a major role in vaccination. Therefore, to identify the highest affinity linear B cell epitopes from EBOV proteins, BCPRED tool was employed. BCPRED is SVM based classifier and uses five different kernel algorithms that provide five-fold cross-

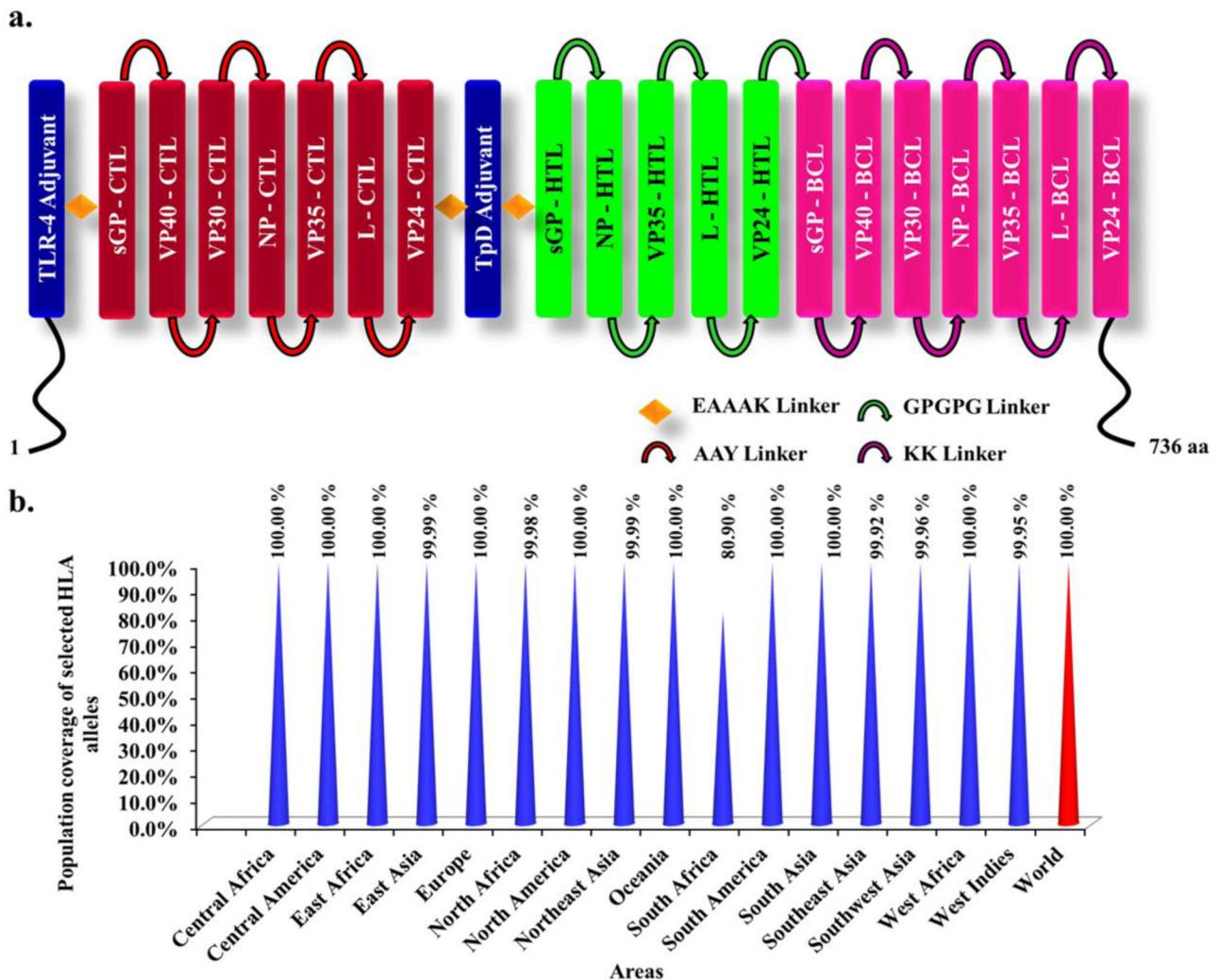


Figure 4. Multiepitope vaccine constructs with its population coverage. (a) Linear representation of the construct multiepitope vaccine of EBOV. The length of the multiepitope vaccine is 736 amino acid residues, and N terminal have TLR4 adjuvant which linked with multiepitope sequences with the help of EAAAK linker. CTL (red) epitopes followed by HTL (green) and BCL (pink) joined by EAAAK linker. HTL, helper T-lymphocyte; CTL, cytotoxic T-lymphocyte, and BCL; B cell lymphocytes. (b) Plot depicting the percentage (%) coverage of MHC-I and MHC-II alleles used in the analysis in the world and its various regions.

validation. It provides the similarity score of the submitted epitope with the already known B cell epitopes. The higher the similarity, the higher is the propensity of the submitted epitope to interact with the B-cells. Thus, the best epitope for each EBOV protein having the highest score was taken into consideration for further analysis (Table S4, supplementary material).

3.6. Construction of multi-epitope subunit vaccine

In the present study, we merged 11 CTL epitopes, 5 HTL epitopes, and 7 B-cell epitopes to construct a multi-epitope subunit vaccine by utilizing the AAY, GPGPG, and KK linkers, respectively (Table S5, supplementary material, and Figure 4a). To enhance the immunogenicity of the vaccine construct, two adjuvants were also added to the final construct using EAAAK linker as it is reported to separate different domains in a protein structure efficiently.

TLRs are the best-known cell surface receptor to recognize various pathogenic patterns present on the pathogens. TLRs

upon interaction with antigen activates the signaling cascades that are responsible for specific adaptive immune response inside the host by increasing the expression of Interferon-beta. Due to their crucial role during infection, TLRs have been one of the centers of attraction for the construction of vaccine adjuvant strategies against bacteria and viruses. (Hedayat et al., 2011) TLR4, TLR5, and TLR9 have been previously targeted by various agonist and vaccine complex that leads to the increased immune response during the infection of hepatitis B virus, human immunodeficiency virus, and Epstein Barr virus (Hedayat et al., 2011). EBOV infection decreases the expression of TLR4 that leads to diminished Interferon beta-mediated immune response (Saghazadeh & Rezaei, 2017). Therefore, the addition of TLR4 agonists with a vaccine would increase the Interferon beta-mediated immune response. Heparin-binding hemagglutinin (hbhA) is one such peptide of *Mycobacterium tuberculosis* that has been reported to agonist the TLR4 activity when added with a cancer vaccine (Jung et al., 2011). Therefore its sequence was retrieved from UniProt and added at the N terminal region of the vaccine construct.

Table 2. Allergenicity, antigenicity, and physiochemical properties of the construct vaccine.

Properties of the vaccine construct	Tool used	Result
Allergenicity Prediction	AllerTop Tool AllergenFP Tool AlgPred Tool	Probable non-allergen Probable non-allergen Non allergen Score = -1.1252241 [Threshold = -0.4]
Antigenicity Prediction	VaxiJen Server AntigenPro Server	0.4852 (Probable ANTIGEN) Predicted Probability of Antigenicity: 0.822274
Physiochemical Properties by ProtParam	Molecular weight Instability index Estimated half-life	79,989.89 35.80 [Stable] 30 h (mammalian reticulocytes, in vitro). >20 h (yeast, in vivo). >10 h (<i>Escherichia coli</i> , in vivo).
Solubility	Aliphatic index Grand average of hydropathicity (GRAVY) Solpro Tool	85.94 -0.290 Soluble with probability 0.923650

Helper T cells help in B cells activation and maturation by immunoglobulin class switching and thus play a central role in the humoral immune response. Therefore, for the enhanced adaptive humoral immune response, another adjuvant TpD was also added in between the CTL and HTL epitopes. TpD is a derivative of tetanus and diphtheria toxoid that interact with the majority of MHC class-II molecules and is able to initiate the humoral response (Fraser et al., 2014). The CTL, HTL, and B cell epitopes, along with hhbA, TpD adjuvant, and all linkers, forms a 736 amino acids long multi-epitope vaccine construct (Figure 4a).

3.7. Population coverage of the vaccine

Vaccines work by presenting the epitopes to CTLs, HTLs, and B cell by representing itself through class-I and II MHC molecules. MHC molecules are widely diversified in nature, but the distribution of their alleles are uneven. Therefore, to check the distribution of MHCs throughout the world population, the Population coverage tool available at the IEDB server was utilized. The analysis provided 91.27% and 99.99% conserveness of selected HLA class-I and class-II alleles, respectively (Figures S9 and S10, [supplementary material](#)). Taking HLA class-I and class-II together, the selected HLA alleles showed 100% population coverage (Figure 4b)

3.8. Multi-epitope subunit vaccine constructs allergenicity, immunogenicity and physiochemical analysis

An effective and safe vaccine must not be allergic in nature and must possess maximum immunogenicity. Allergens are the antigens or peptides that can evoke IgE leading to hypersensitivity reactions. The severe effects of these hypersensitivity reactions can lead to anaphylaxis that can be lethal (Broadfield et al., 2002). Therefore, to check the allergenicity of the vaccine construct AllerTop, AllergenFP, and AlgPred tools were used. The AllerTop and AllergenFP are based on different algorithms and predicted the subunit vaccine construct as non-allergen. AlgPred gave a score of -1.125 that also signifies the non-allergenicity of the multi-epitope construct (Table 2).

To check the immunogenicity of the vaccine, we employed Vaxijen and ANTIGENPro servers. Based on physiochemical properties of the submitted peptides, the Vaxigen server has an in-built threshold of 0.4 for the virus model. Thus, any peptide above this threshold is antigenic for viruses. Likewise, ANTIGENpro predicts the peptide antigenicity by utilizing the microarray data. Predicted antigenicity score for a peptide >0.5 in ANTIGENpro is considered a good antigen. Vaxigen server provided the antigenic score of 0.485, whereas ANTIGENPro gave the score of 0.822 that signifies a better immunogenicity of the constructed vaccine (Table 2).

To determine the physiochemical properties of the vaccine construct, ProtParam server was used. The molecular weight of the vaccine construct was predicted to be 79.9 kDa that was following the ideal molecular weight of the vaccine (>50 kDa). Recent research have shown that, higher is the molecular weight of the bioconjugate, the higher is the exposure rate in the lymphatic region. Thus, higher will be the chances of coming in contact with the naïve T and B cells in the lymph nodes (Liu & Irvine, 2015). The pI of the vaccine was obtained as 9.6 that depicts the basic nature of the vaccine construct. The stability of the construct is one of the major factors for its efficiency. The instability index <40 depicts the stability of protein, whereas the instability index >40 depicts the instability of the protein. The instability index of the construct was found as 35.80 that shows the stable nature of the subunit vaccine. The half-life of a protein is the time required for half amount of protein to be denatured inside the body. The half-life of the EBOV vaccine construct was estimated as >30 h for mammalian reticulocyte *in vitro*, >20 h for yeast and >10 h for *E. coli* *in vivo* (Table 2).

3.9. Structure prediction and validation of the multi epitope subunit vaccine construct

To simulate the stability of vaccine alone and vaccine-immuno-molecule complex, its secondary and tertiary structures were predicted using PSIPRED and I-TASSER server. PSIPRED is based on neural networks and predicts the propensity of the amino acids to lie in the alpha-helical, beta-sheet, or coil regions. The PSIPRED analysis of the multi-epitope vaccine construct showed the formation of alpha-helix and coils with few beta coils in the middle (Figure 11, [supplementary material](#)). I-TASSER predicts the complete 3D structure of the given

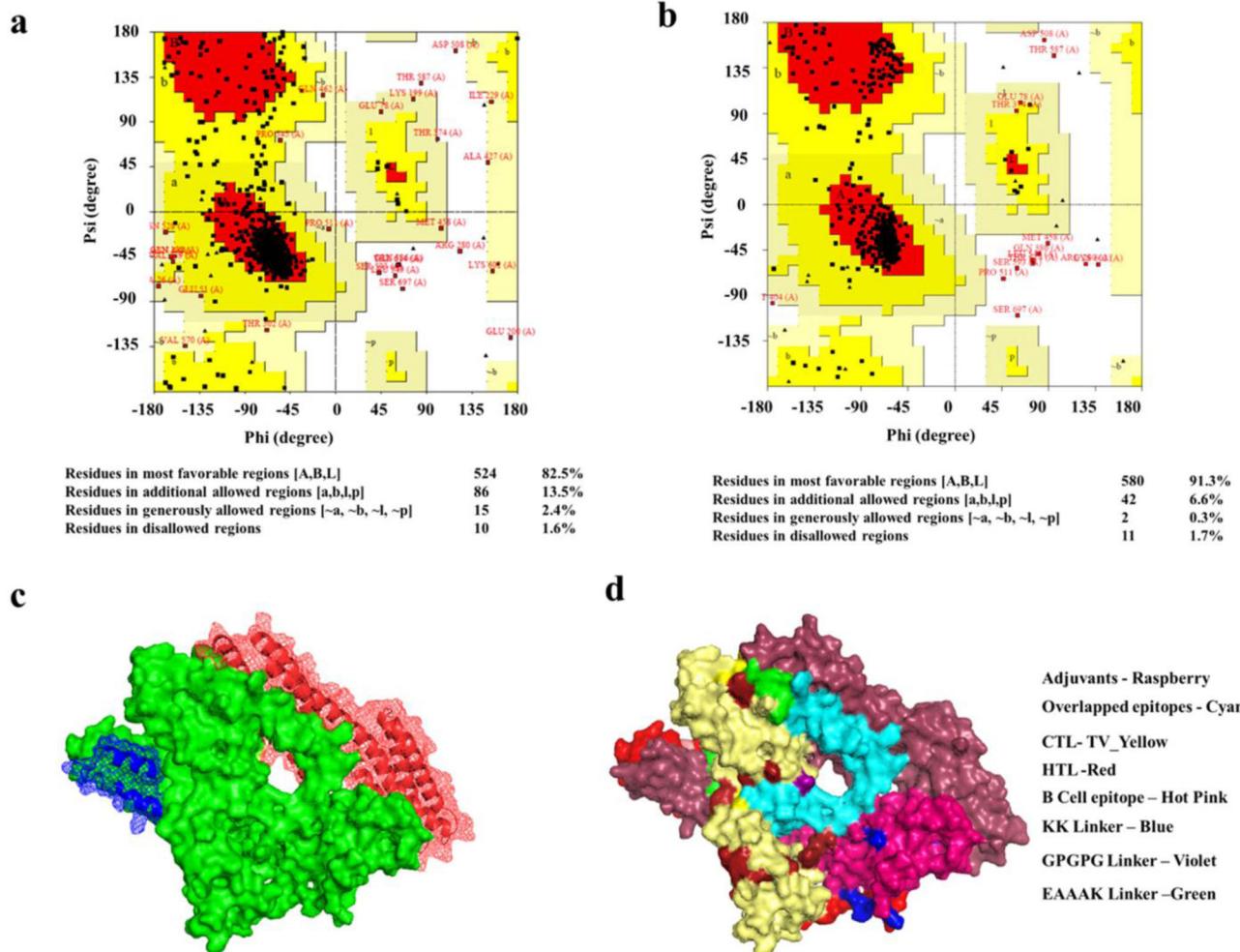


Figure 5. Tertiary structure and validation of multi-epitope vaccine construct (a) Ramachandran plot before refinement and (b) after refinement showing allowed and disallowed region of amino acids. (c) The tertiary structure of the vaccine constructs depicting TLR-4 adjuvant, Heparin-binding hemagglutinin (hbHA) (red), and universal TpD adjuvant (blue). (d) The tertiary structure of the vaccine constructs depicting various linkers, CTL, HTL, and B cell epitopes.

peptide sequence on the basis of the multi-template based threading method and provides several models with a rank of C-Score. C-score is based on the query sequence similarity with the threaded templates. It ranges between -5 and 2 and is used for estimating the quality of the generated models. Higher C-score represents a model with higher confidence. The first model was constructed with a C-score of -0.48, thus depicting a model with a good confidence score (Table S6, supplementary material). The predicted structure of the obtained model was further validated by constructing a Ramachandran plot by using the PROCHECK server. Ramachandran plot deduced that 82.5% residues of predicted structure lied in the most favorable region (Figure 5a). The model construct was further refined by using GalaxyRefine and ModRefiner. The GalaxyRefine employs repeated structural perturbations and MD simulations for refining the tertiary structure of a protein. First, it rebuilds the side chain by assigning the highest possible rotational isomers, and then rebuilding of side chains is performed by an aggressive and mild relaxation method. GalaxyRefine generates five structural models, where the first model is a result of structural perturbations applied only to side chains, whereas the rest is generated by more aggressive perturbations in the loop and other

secondary structural elements. Whereas, Modrefiner algorithm is based on refinement at the atomic level and can refine the submitted structure on the basis of a reference model or by *ab initio* means. The Ramachandran plot was constructed for all the refined structures and the best model having 91.3% residues in the most favorable region along with 6.6% in the additionally allowed region and only 0.3% in the generously allowed region generated by GalaxyRefine with GDT-HA of 0.95, RMSD 0.425, MolProbity 2.443 and clash score of 1.4 was selected as the best vaccine construct and used for further analysis (Figure 5b and Table S7, supplementary material).

3.10. Multi-epitope subunit vaccine constructs interaction with TLR4 and its simulation analysis

A TLR4 adjuvant was added at the N terminal region of the vaccine construct. Therefore to check the interaction of sub-unit vaccine with TLR4 protein, protein-protein docking was performed by employing the ClusPro server. TLR4 was submitted as a receptor, while the vaccine constructs as a ligand. ClusPro performs the docking calculation in three steps that involve a rigid body-based docking from the billions of available conformation; RMSD based clustering of the lowest

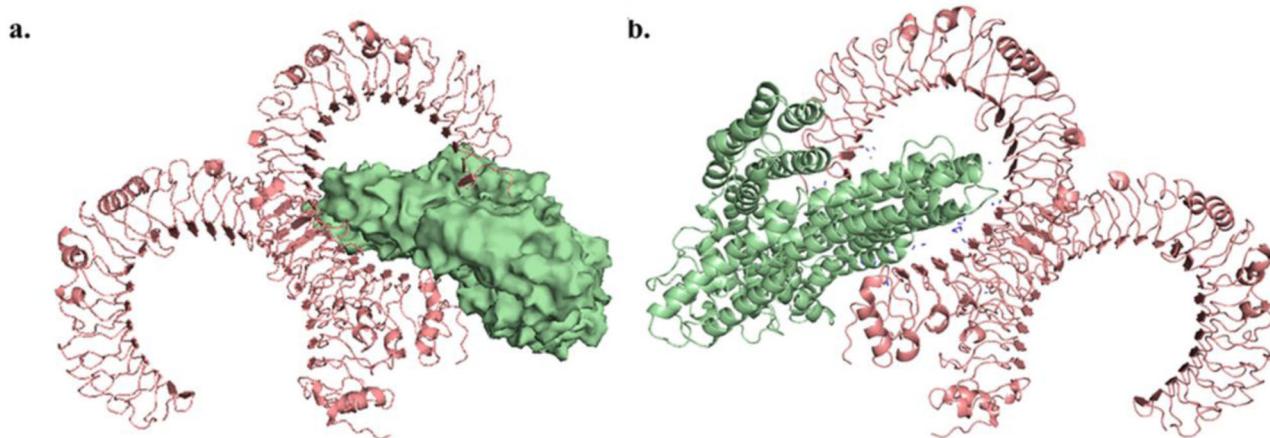


Figure 6. Docked complex and Molecular dynamic simulation analysis of Multi-epitope vaccine with TLR-4 complex. (a) The best conformer for the complex of multi-epitope vaccine constructs (pale green) with human TLR-4 complex (Salmon color). (b) Hydrogen bonds (blue) formed between vaccine construct and TLR-4 complex. Docking analysis depicted the interaction of the vaccine construct with one subunit of the human TLR-4 complex.

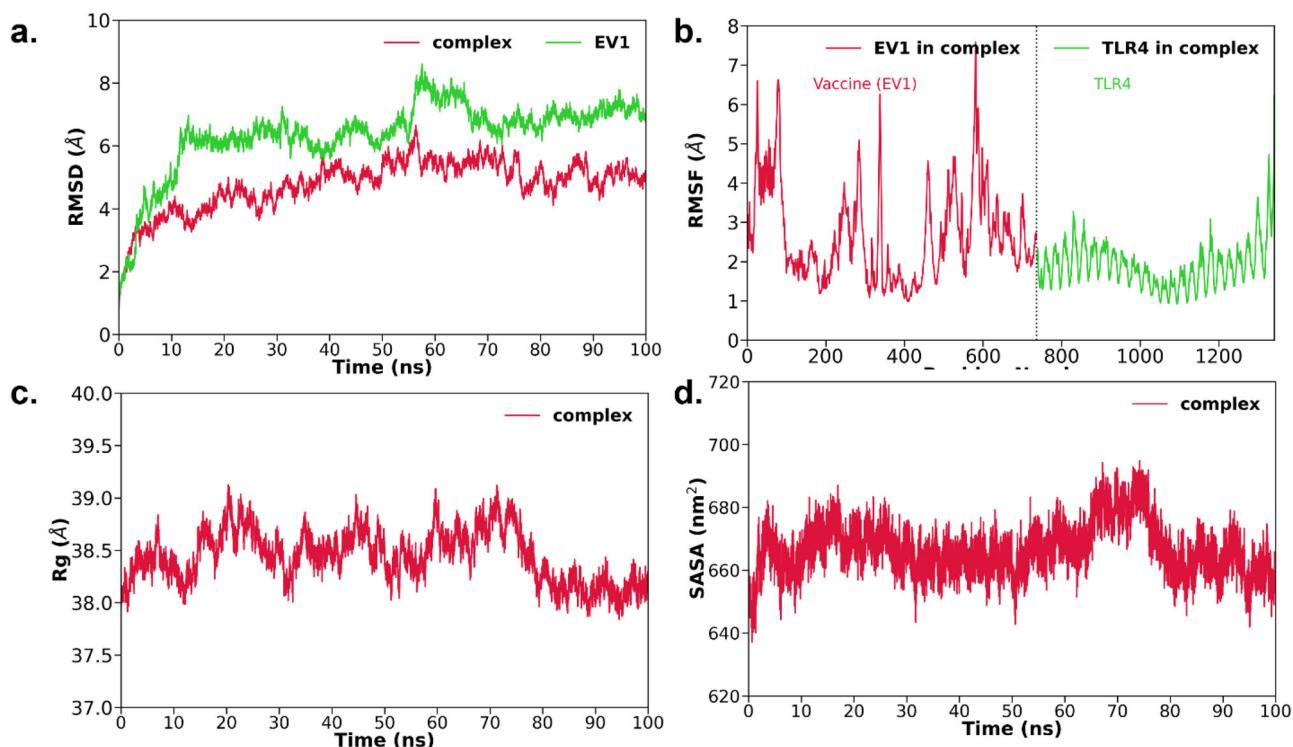


Figure 7. Molecular dynamics simulation analysis of Ebola multi-epitope vaccine construct, and it's complex with TLR4. (a) The time evolution of RMSDs of backbone atoms for the multi-epitope vaccine construct (green) and vaccine-TLR4 complex (red). (b) RMSF of C_z atoms for the vaccine construct (red) and TLR4 (green) in the complex. (c) Radius of gyration (R_g) of the complex during the 100 ns simulation of the vaccine-TLR4 complex. (d) Time evolution of the solvent accessible surface area (SASA).

energy model, and energy minimization of the refined structure. ClusPro generated 30 docked models between the receptor and ligand-based on the weighted score for the energy of the receptor-ligand complex formed. The model with the lowest energy was selected for further analysis (Figure 6a and 6b).

3.11. Stability of the multi-epitope subunit vaccine construct complexed with TLR4 receptor

In order to evaluate the structural stability of the constructed multi-epitope vaccine and its complex with TLR4, a 100 ns MD simulation was performed and trajectory driven root-

mean-squared deviation (RMSD), root-mean-squared fluctuations (RMSFs), the radius of gyration (R_g), and solvent accessible surface area (SASA) analyses were performed (see Figure 7 and Table S8, supplementary material).

Trajectory analysis of the vaccine construct as well as the TLR4/vaccine complex showed the RMSD value of the backbone atoms reached equilibrium at ~55 ns and remained stable till the end of the simulation (Figure 7a). The unchanged RMSD value over the course of the last 45 ns simulation depicted the formation of optimal electrostatic interactions and stable conformation of the complex. The average RMSD

Table 3. Energetic components of the binding free energy for vaccine and receptor complex in kcal/mol.

Components	ΔE_{vdw}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{non-polar}$	$\Delta G_{Bind}^{\text{sim}}$
ET complex	-182.7 (0.3)	-5112.4 (2.8)	5123.9 (2.7)	-23.0 (0.0)	-194.2 (0.5)

Standard errors of the mean (SEM) are provided in parentheses.

Table 4. Critical residues involved in complex stabilization and highest energy contributing residues of vaccine and receptor.

Vaccine construct	Energy (kcal/mol)	Receptor (TLR4)	Energy (kcal/mol)
Arg418	-6.0	Arg289	-6.1
Glu117	-6.0	Arg264	-5.5
Tyr422	-5.0	Glu27	-5.0
Lys478	-4.6	Arg598	-4.7
Lys199	-2.8	Val30	-3.0
Tyr399	-2.8	Val32	-2.6
Val130	-2.7	Pro28	-2.4
Leu419	-2.3	Lys362	-2.4
Leu345	-2.3	Phe573	-2.0
Val140	-2.2	Asn265	-1.7
Phe482	-2.2	Gln39	-1.7
Ala146	-2.2	Val338	-1.6
Arg416	-2.1	Val33	-1.6
Glu135	-2.0	Pro34	-1.5
Arg415	-2.0	Val602	-1.5
Ile353	-2.0	His456	-1.5
Cys349	-2.0	Asp60	-1.5
Val147	-1.7	Phe500	-1.5

Only residues with more than -1.5 kcal/mol free energy contributions are listed.

value of the complex ($4.9 \pm 0.6 \text{ \AA}$) is lower than the unbounded vaccine construct ($\sim 6.4 \pm 1.1 \text{ \AA}$). Further, the average RMSD value ($\sim 5.1 \pm 0.8 \text{ \AA}$) of the vaccine decreased when it was bound to TLR4 compared to the unbound vaccine ($\sim 6.4 \pm 1.1 \text{ \AA}$). This suggests that the vaccine construct gets stabilized after the complex formation with TRL4.

Next, we computed the root-mean-squared fluctuations (RMSFs) of C_{α} -atoms for the complex, and shown in Figure 7(b). It is evident from Figure 7(b) that the residual fluctuation in the vaccine construct is relatively higher compared to TLR4. It is further apparent from the RMSF plot that the interface area of both vaccine and TLR4 shows relatively low and stable RMSF values. The radius of gyration (R_g) reveals the compactness and order of the system. The high compactness of a system indicates its higher stability. The R_g analysis of the vaccine—TLR4 complex trajectory showed the stable nature of the complex with a mean R_g of $38.4 \pm 0.2 \text{ \AA}$, reflecting compactness of the complex system throughout 100 ns simulation (Figure 7c).

The solvent-accessible surface area (SASA) analysis of the trajectory is another way to monitor the complex's stability. Lower peaks in the SASA profile depicts the protein complex's contraction while the higher peaks reveal the expansion. Thus, the lesser the deviation, the more stable is the complex. The SASA profile analysis of the vaccine—TLR4 complex showed an average value of $667.2 \pm 8.1 \text{ nm}^2$ depicting the stable conformation of the complex (Figure 7d). Thus, MD simulation analysis strengthened the docking interaction analysis and showed the vaccine construct interacts with the immune receptors like TLR4 and can generate a significant immune response against the Ebola virus.

3.12. Free energy calculations and energy decompositions using MM-PBSA reveals the stability and high affinity of vaccine construct to the immune receptor

To analyze the energetic basis of the binding of the multi-epitope vaccine construct and TLR4, we calculated the binding free energy via MM-PBSA (see Table 3). The calculated change in Gibbs free energy (ΔG) or binding affinity (ΔG_{Bind}) was observed to be $-194.2 \pm 0.5 \text{ kcal/mol}$. It is evident from Table 3 that the intermolecular electrostatic (ΔE_{elec}) and van der Waals (ΔE_{vdw}) interactions, as well as non-polar solvation free energy ($\Delta G_{non-polar}$), favor the association of the vaccine construct to TLR4. On the other hand, polar solvation free energy (ΔG_{polar}) opposes the vaccine/TLR4 complex formation. The formation of the complex is mainly contributed by the electrostatic interaction energy ($\Delta E_{elec} = -5112.8 \pm 2.8 \text{ kcal/mol}$) followed by the van der Waals interactions ($\Delta E_{vdw} = -182.7 \pm 0.3 \text{ kcal/mol}$). The highest contribution of ΔE_{elec} in the complex binding energy strengthens the hydrogen bond interactions between the vaccine construct and the immune response. The total polar part ($\Delta E_{elec} + \Delta E_{polar} = 11.5 \pm 3.9 \text{ kcal/mol}$) of free energy disfavor the complex formation and the total non-polar component ($\Delta E_{vdw} + \Delta E_{non-polar} = -205.7 \pm 0.3 \text{ kcal/mol}$) favored the vaccine binding. Overall, the complex formation is mainly driven by the van der Waal interactions.

To find the critical residues involved in stabilizing the vaccine-construct-TLR4 complex system, the total binding free energy was subjected to pairwise residual decomposition using the MM-GBSA scheme. Residues with the absolute free energy of $\geq 1.5 \text{ kcal/mol}$ are listed in Table 4. The per-residue decomposition study showed a total of 18 residues of the vaccine construct, namely Arg418, Glu117, Tyr422, Lys478, Lys199, Tyr399, Val130, Leu419, Leu345, Val140, Phe482, Ala146, Arg416, Glu135, Arg415, Ile353, Cys349, and Val147 that participate rigorously in binding interactions with the vaccine construct and the absolute contribution of each residue to the binding energy is $\geq 1.5 \text{ kcal/mol}$ (see Table 4). Most of these residues are from the binding side region of TLR4, which favors direct contact with the vaccine. On the other hand, the following residues, namely Arg289, Arg264, Glu27, Arg598, Val30, Val32, Pro28, Lys362, Phe573, Asn265, Gln39, Val338, Val33, Pro34, Val602, His456, Asp60, and Phe500 of TLR4 formed favorable interactions with the vaccine construct (Table 4).

3.13. Hydrogen bonds and hydrophobic interactions

H-bonds (h-bonds) play a critical role in the molecular recognition and stability of the protein complex conformation (Hubbard & Haider, 2010). To complement the per-residue decomposition analysis, h-bond analyses were performed for

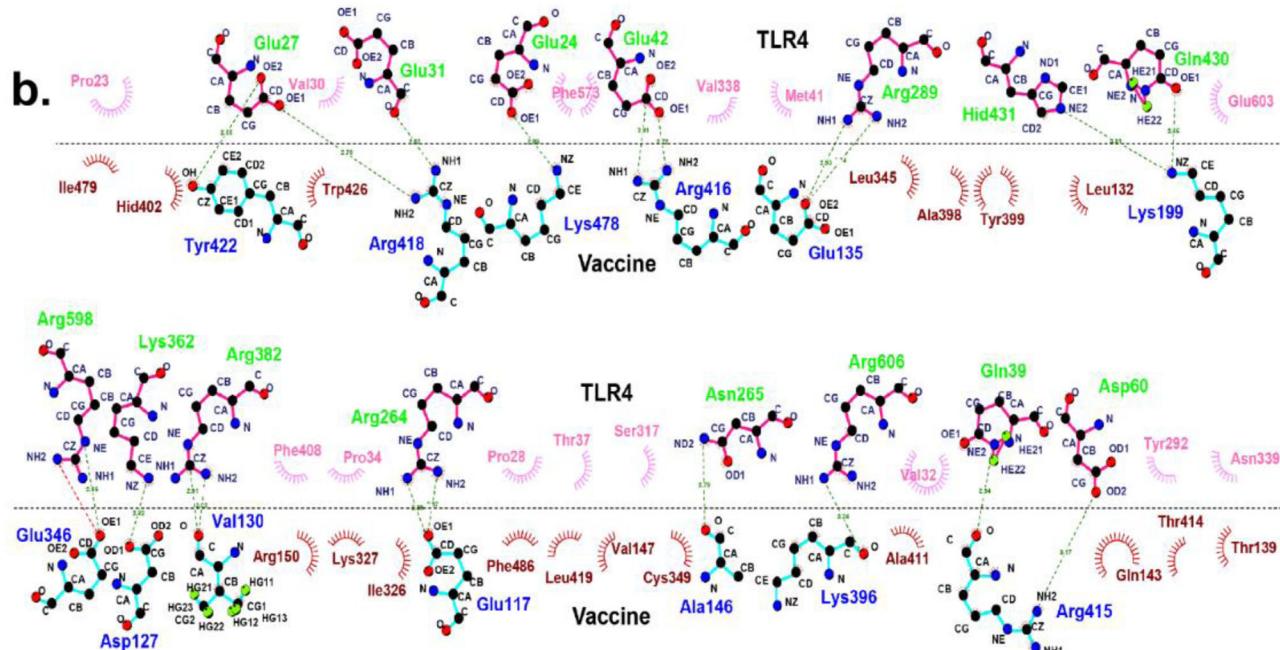
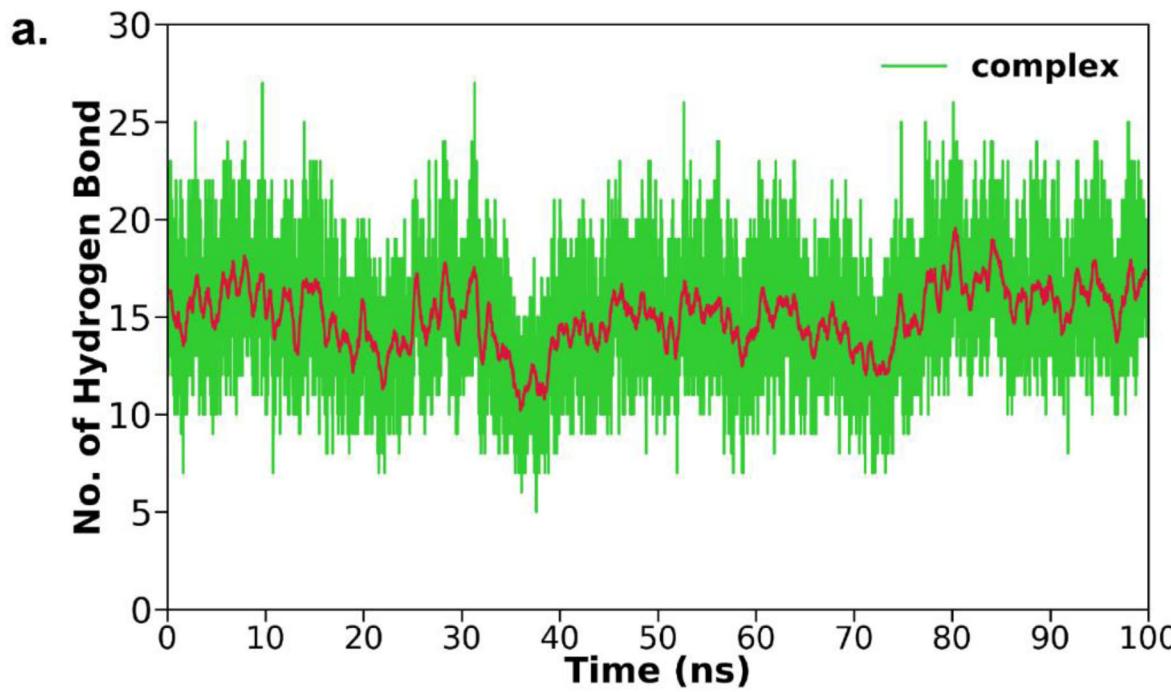


Figure 8. (a) Time evolution of the total number of hydrogen bonds between TLR4 and the vaccine construct (in green) and running average of hydrogen bonds with 50 windows (in red). (b) The interaction profile of the final structure of the TLR4-vaccine construct. The upper part is related to TLR4, and the lower part is for the vaccine construct. The hydrogen bonds are depicted as dotted lines (lime green), while the hydrophobic interactions are represented in red color for the vaccine construct and pink for TLR4. The plot was generated using LigPlot+.

the complex simulation trajectory, and the corresponding percentage occupancy is recorded in Table S9, supplementary material. H-bond occupancy from the simulations is defined as the percentage of h-bonds throughout the simulated trajectory by setting the acceptor-donor distance of $<3.5\text{ \AA}$ and angle cut off $\geq 120^\circ$. The H-bond analysis provides valuable information regarding the strongest h-bond formed between TLR4 and vaccine. The h-bond interaction analysis of the trajectory revealed the formation of ~ 15 h-bonds between the vaccine construct and TLR4 complex (Figure 8a). These h-bonds are relatively stable

throughout the simulation. The donor and acceptor pairs for the hydrogen bond formation between the vaccine construct and receptor with more than 20% occupancy are depicted in Figure 8(b). The key residues forming h-bond include Asn265, Glu27, Arg264, Gln39, Arg289, Arg382, Arg499, Cyx29, and Arg598 from TLR4 and Ala146, Arg418, Glu117, Tyr422, Glu135, Val130, and Glu346 from the vaccine construct. It should further be noted here that Glu117 and Glu135 from vaccine form four and three h-bonds with TLR4, respectively (see in Table S9, supplementary material).

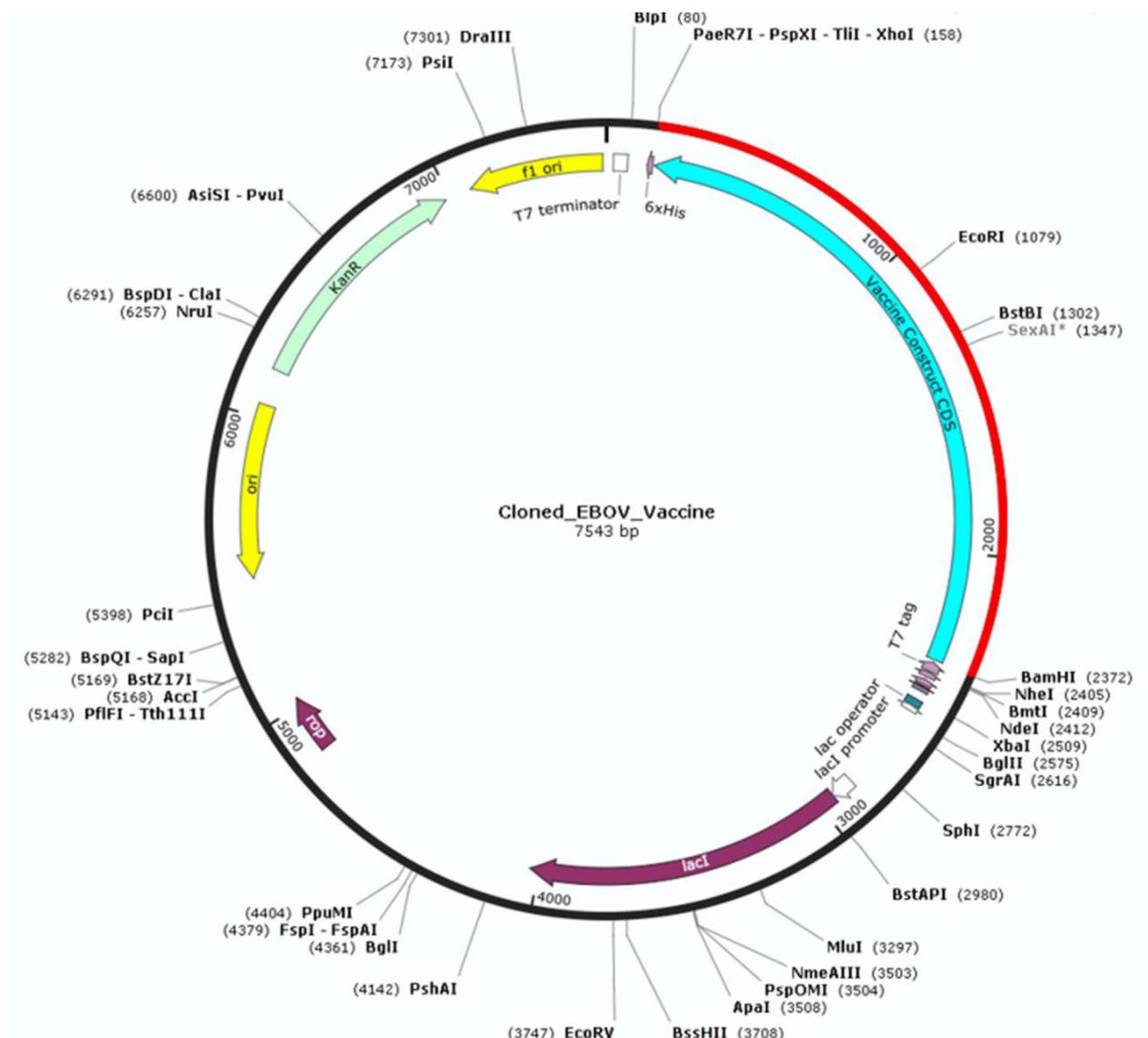


Figure 9. *In silico* cloning of multi-epitope vaccine construct: *In silico* cloning of construct vaccine in pET21+ vector, construct vaccine depicted in blue color and present in between the sites of *Xba*I and *Bam*H I of the vector backbone.

In addition, we complement the above findings by providing information on h-bond and hydrophobic interactions from the final conformation of production simulation by plotting a 2D Ligplot diagram and shown in Figure 8(b). Hydrogen bonds are shown with green dotted lines, and the hydrophobic residues are shown in red and pink semi-circle spikes. The major hydrophobic interacting residues in TLR4 comprises of Pro23, Val30, Phe573, Val338, Met41, Glu603, Phe408, Pro34, Pro28, Thr37, Ser317, Val32, Tyr292, Asn339, and in vaccine includes Ile479, His402, Trp426, Leu345, Ala398, Tyr399, Leu132, Arg150, Leu327, Ile326, Phe486, Leu419, Val147, Cys349, Ala411, Gln143, Thr414, Thr139. This large number of interactions account for the high stability and good binding affinity of the vaccine to TLR4.

3.14. Codon optimization and *in silico* cloning

Generally, during the translation process, the usage of the codon for the specific amino acid varies among different

organisms. The codon usage can be represented by the codon adaptation index (CAI). The range of CAI varies from 0 to 1 where 0 corresponds to no codon usage while 1 corresponds to the highly adapted codons in the given organism. To generate a gene construct that can stably translate into the multi-epitope subunit vaccine, its cDNA was constructed using the reverse translation tool. The resultant cDNA construct was 22.8 Kb long and possess a CAI of 0.58 with a GC content of 59.19% as calculated by Java codon adaptation tool (JCAT) for *Escherichia coli* (strain K12). The codon usage varies among the host that is used for cloning and expression. Unadapted codons may lead to lower expression. As well, the presence of rho-independent transcription termination sites and commonly used restriction enzyme cleavage sites inside the codon sequence may result in the transcription failure of the vaccine construct, and thus such sites are required to be avoided. Therefore to adopt the codon's usage for *Escherichia coli* and avoid rho-independent transcription terminator sites, codon adaptation and optimization

were performed using JCAT. This adaptation and optimization lead to an increase in the CAI from 0.58 for unadapted to 0.98 for the adapted sequence. The GC content further depicted a formation of a stable DNA construct and the good propensity of protein expression in the *E. coli* system (Figure S12a–b).

The vaccine DNA construct was then reversed, and the sites for *Xhol* and *BamH1* were added at the 5' and 3' end, respectively. The final gene construct was cloned in between the *Xhol* and *BamH1* restriction sites of the pET28a(+) vector, and a 6 mer histidine tag was added at the C-terminal end for further purification (Figure 9).

4. Conclusion

Ebola causes a highly severe hemorrhagic fever with a high mortality rate. To control this epidemic viral infection, the availability of an effective vaccine is of utmost importance. There are various attempts made towards its vaccine development, but no successful vaccine with higher efficacy and efficiency is available to date. The advancement in the field of immunoinformatics analysis has overcome the disadvantage of traditional vaccine development. This study provides a brief immunoinformatics approach to find a novel vaccine candidate against Ebola virus infection. Here we explored the proteome of the Ebola virus to search for the best CTL, HTL, and B cell epitopes to construct a multiepitope subunit vaccine. Sequence analysis of the vaccine construct showed the predicted epitopes are antigenic and non-allergic in nature. MD simulation depicted the stability of the tertiary structure of the vaccine construct and its complex with TLR4 complex and *in silico* cloning analysis helped in the generation of an effective vector that can be used for the multi-epitope vaccine expression. In summary, an extensive immunoinformatics approach was exploited to develop an effective and efficacious multi-epitope vaccine against the deadly Ebola virus.

Acknowledgements

US would like to acknowledge the Ministry of Education, Govt. of India, New Delhi, NJ to the Council of Scientific & Industrial Research (CSIR), Govt. of India, New Delhi and MFS to Department of Science and Technology (DST/INSPIRE_FELLOWSHIP/2017/IF170145), Govt. of India, New Delhi for their individual fellowships.

Disclosure statement

The authors have declared no competing interest.

ORCID

Md Fulbabu Sk  <http://orcid.org/0000-0003-4341-8370>
Parimal Kar  <http://orcid.org/0000-0001-8451-9739>

References

- Ajisegiri, W. S., Chughtai, A. A., & MacIntyre, C. R. (2018). A risk analysis approach to prioritizing epidemics: Ebola virus disease in West Africa as a case study. *Risk Analysis*, 38(3), 429–441. <https://doi.org/10.1111/risa.12876>
- Berendsen, H. J., Postma, J. v., van Gunsteren, W. F., DiNola, A., & Haak, J. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8), 3684–3690. <https://doi.org/10.1063/1.448118>
- Biedenkopf, N., Hartlieb, B., Hoenen, T., & Becker, S. (2013). Phosphorylation of Ebola virus VP30 influences the composition of the viral nucleocapsid complex: Impact on viral transcription and replication. *The Journal of Biological Chemistry*, 288(16), 11165–11174. <https://doi.org/10.1074/jbc.M113.461285>
- Bounds, C. E., Terry, F. E., Moise, L., Hannaman, D., Martin, W. D., De Groot, A. S., Suschak, J. J., Dupuy, L. C., & Schmaljohn, C. S. (2017). An immunoinformatics-derived DNA vaccine encoding human class II T cell epitopes of Ebola virus, Sudan virus, and Venezuelan equine encephalitis virus is immunogenic in HLA transgenic mice. *Human Vaccines & Immunotherapeutics*, 13(12), 2824–2836. <https://doi.org/10.1080/21645515.2017.1329788>
- Broadfield, E., McKeever, T. M., Scrivener, S., Venn, A., Lewis, S. A., & Britton, J. (2002). Increase in the prevalence of allergen skin sensitization in successive birth cohorts. *The Journal of Allergy and Clinical Immunology*, 109(6), 969–974. <https://doi.org/10.1067/mai.2002.124772>
- Bui, H. H., Sidney, J., Dinh, K., Southwood, S., Newman, M. J., & Sette, A. (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*, 7, 153. <https://doi.org/10.1186/1471-2105-7-153>
- Bukreyev, A. A., Chandran, K., Dolnik, O., Dye, J. M., Ebihara, H., Leroy, E. M., Mühlberger, E., Netesov, S. V., Patterson, J. L., Paweska, J. T., Saphire, E. O., Smither, S. J., Takada, A., Towner, J. S., Volchkov, V. E., Warren, T. K., & Kuhn, J. H. (2014). Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012–June 2013. *Archives of Virology*, 159(4), 821–830. <https://doi.org/10.1007/s00705-013-1846-9>
- Calis, J. J. A., Maybeno, M., Greenbaum, J. A., Weiskopf, D., De Silva, A. D., Sette, A., Keşmir, C., & Peters, B. (2013). Properties of MHC Class I presented peptides that enhance immunogenicity. *PLoS Computational Biology*, 9(10), e1003266. <https://doi.org/10.1371/journal.pcbi.1003266>
- Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, I. T. E., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Ghoreishi, D., Gilson, M. K., Gohlke, H., Goetz, A. W., Greene, D., Harris, R., Homeyer, N., Huang, Y., Izadi, S., Kovalenko, A., Kurtzman, T., ... Kollman, D. M. Y. a. P. A. (2018). AMBER 2018. University of California, San Francisco.
- Cenciarelli, O., Gabbarini, V., Pietropaoli, S., Malizia, A., Tamburini, A., Ludovici, G. M., Carestia, M., Giovanni, D. D., Sassolini, A., Palombi, L., Bellecci, C., & Gaudio, P. (2015). Viral bioterrorism: Learning the lesson of Ebola virus in West Africa 2013–2015. *Virus Research*, 210, 318–326. <https://doi.org/10.1016/j.virusres.2015.09.002>
- Cheng, J., Randall, A. Z., Sweredoski, M. J., & Baldi, P. (2005). SCRATCH: A protein structure and structural feature prediction server. *Nucleic Acids Res*, 33(Web Server issue), W72–6. <https://doi.org/10.1093/nar/gki396>
- Coller, B. G., Blue, J., Das, R., Dubey, S., Finelli, L., Gupta, S., Helmond, F., Grant-Klein, R. J., Liu, K., Simon, J., Troth, S., VanRheenen, S., Waterbury, J., Wivel, A., Wolf, J., Heppner, D. G., Kemp, T., Nichols, R., & Monath, T. P. (2017). Clinical development of a recombinant Ebola vaccine in the midst of an unprecedented epidemic. *Vaccine*, 35(35 Pt A), 4465–4469. <https://doi.org/10.1016/j.vaccine.2017.05.097>
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N-log (N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089–10092. <https://doi.org/10.1063/1.464397>
- Dash, R., Das, R., Junaid, M., Akash, M. F. C., Islam, A., & Hosen, S. Z. (2017). In silico-based vaccine design against Ebola virus glycoprotein. *Advances and Applications in Bioinformatics and Chemistry*, 10, 11–28. <https://doi.org/10.2147/AABC.S115859>
- Dhanda, S. K., Vir, P., & Raghava, G. P. (2013). Designing of interferon-gamma inducing MHC class-II binders. *Biology Direct*, 8(1), 30. <https://doi.org/10.1186/1745-6150-8-30>

- Dimitrov, I., Flower, D. R., & Doytchinova, I. (2013). AllerTOP-a server for in silico prediction of allergens. *BMC Bioinformatics*, 14(Suppl 6), S4. <https://doi.org/10.1186/1471-2105-14-S6-S4>
- Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2014). AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics (Oxford, England)*, 30(6), 846–851. <https://doi.org/10.1093/bioinformatics/btt619>
- Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8(1), 4. <https://doi.org/10.1186/1471-2105-8-4>
- El-Manzalawy, Y., Dobbs, D., & Honavar, V. (2008). Predicting linear B-cell epitopes using string kernels. *Journal of Molecular Recognition*, 21(4), 243–255. <https://doi.org/10.1002/jmr.893>
- Feldmann, H., Muhlberger, E., Randolph, A., Will, C., Kiley, M. P., Sanchez, A., & Klunk, H. D. (1992). Marburg virus, a filovirus: Messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Research*, 24(1), 1–19. [https://doi.org/10.1016/0168-1702\(92\)90027-7](https://doi.org/10.1016/0168-1702(92)90027-7)
- Fraser, C. C., Altreuter, D. H., Ilyinskii, P., Pittet, L., LaMothe, R. A., Keegan, M., Johnston, L., & Kishimoto, T. K. (2014). Generation of a universal CD4 memory T cell recall peptide effective in humans, mice and non-human primates. *Vaccine*, 32(24), 2896–2903. <https://doi.org/10.1016/j.vaccine.2014.02.024>
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31(13), 3784–3788. <https://doi.org/10.1093/nar/gkg563>
- Gohlke, H., & Case, D. A. (2004). Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry*, 25(2), 238–250. <https://doi.org/10.1002/jcc.10379>
- Gohlke, H., Kiel, C., & Case, D. A. (2003). Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes. *Journal of Molecular Biology*, 330(4), 891–913. [https://doi.org/10.1016/S0022-2836\(03\)00610-7](https://doi.org/10.1016/S0022-2836(03)00610-7)
- Gross, L., Lhomme, E., Pasin, C., Richert, L., & Thiebaut, R. (2018). Ebola vaccine development: Systematic review of pre-clinical and clinical studies, and meta-analysis of determinants of antibody response variability after vaccination. *International Journal of Infectious Diseases*, 74, 83–96. <https://doi.org/10.1016/j.ijid.2018.06.022>
- Harty, R. N., Brown, M. E., Wang, G., Huibregtse, J., & Hayes, F. P. (2000). A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: Implications for filovirus budding. *Proceedings of the National Academy of Sciences of the United States of America*, 97(25), 13871–13876. <https://doi.org/10.1073/pnas.250277297>
- Hedayat, M., Netea, M. G., & Rezaei, N. (2011). Targeting of Toll-like receptors: A decade of progress in combating infectious diseases. *The Lancet*, 377(9771), 702–712. [https://doi.org/10.1016/S1473-3099\(11\)70099-8](https://doi.org/10.1016/S1473-3099(11)70099-8)
- Heo, L., Park, H., & Seok, C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Research*, 41(Web Server issue), W384–388. <https://doi.org/10.1093/nar/gkt458>
- Homeyer, N., & Gohlke, H. (2012). Free energy calculations by the molecular mechanics Poisson-Boltzmann Surface Area Method. *Molecular Informatics*, 31(2), 114–122. <https://doi.org/10.1002/minf.201100135>
- Hubbard, R. E., & Haider, M. K. (2010). Hydrogen bonds in proteins: Role and strength. *eLS*.
- Jonniya, N. A., Sk, M. F., & Kar, P. (2019). Investigating phosphorylation-induced conformational changes in WNK1 kinase by molecular dynamics simulations. *ACS Omega*, 4(17), 17404–17416. <https://doi.org/10.1021/acsomega.9b02187>
- Jonniya, N. A., Sk, M. F., & Kar, P. (2020). A comparative study of structural and conformational properties of WNK kinase isoforms bound to an inhibitor: Insights from molecular dynamic simulations. *Journal of Biomolecular Structure and Dynamics*, 1–16. <https://doi.org/10.1080/07391102.07392020.01827035>
- Jung, I. D., Jeong, S. K., Lee, C. M., Noh, K. T., Heo, D. R., Shin, Y. K., Yun, C. H., Koh, W. J., Akira, S., Whang, J., Kim, H. J., Park, W. S., Shin, S. J., & Park, Y. M. (2011). Enhanced efficacy of therapeutic cancer vaccines produced by co-treatment with Mycobacterium tuberculosis heparin-binding hemagglutinin, a novel TLR4 agonist. *Cancer Research*, 71(8), 2858–2870. <https://doi.org/10.1158/0008-5472.CAN-10-3487>
- Kalita, P., Padhi, A. K., Zhang, K. Y. J., & Tripathi, T. (2020). Design of a peptide-based subunit vaccine against novel coronavirus SARS-CoV-2. *Microbial Pathogenesis*, 145, 104236. <https://doi.org/10.1016/j.micpath.2020.104236>
- Kar, P., & Knecht, V. (2012a). Energetics of mutation-induced changes in potency of lersivirine against HIV-1 reverse transcriptase. *Journal of Physical Chemistry B*, 116(22), 6269–6278. <https://doi.org/10.1021/jp300818c>
- Kar, P., & Knecht, V. (2012b). Energetic basis for drug resistance of HIV-1 protease mutants against amprenavir. *Journal of Computer-Aided Molecular Design*, 26(2), 215–232. <https://doi.org/10.1007/s10822-012-9550-5>
- Kar, P., & Knecht, V. (2012c). Origin of decrease in potency of darunavir and two related antiviral inhibitors against HIV-2 compared to HIV-1 protease. *The Journal of Physical Chemistry B*, 116(8), 2605–2614. <https://doi.org/10.1021/jp211768n>
- Kar, P., & Knecht, V. (2012d). Mutation-induced loop opening and energetics for binding of tamiflu to influenza N8 neuraminidase. *The Journal of Physical Chemistry B*, 116(21), 6137–6149. <https://doi.org/10.1021/jp3022612>
- Khan, M. A., Hossain, M. U., Rakib-Uz-Zaman, S. M., & Morshed, M. N. (2015). Epitope-based peptide vaccine design and target site depiction against Ebola viruses: An immunoinformatics study. *Scandinavian Journal of Immunology*, 82(1), 25–34. <https://doi.org/10.1111/sji.12302>
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., & Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research*, 33(12), 889–897. <https://doi.org/10.1021/ar000033j>
- Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols*, 12(2), 255–278. <https://doi.org/10.1038/nprot.2016.169>
- Kräutler, V., van Gunsteren, W. F., & Hünenberger, P. H. (2001). A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry*, 22(5), 501–508. [https://doi.org/10.1002/1096-987X\(20010415\)22:5<501::AID-JCC1021>3.0.CO;2-V](https://doi.org/10.1002/1096-987X(20010415)22:5<501::AID-JCC1021>3.0.CO;2-V)
- Larsen, M. V., Lundsgaard, C., Lamberth, K., Buus, S., Lund, O., & Nielsen, M. (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics*, 8, 424. <https://doi.org/10.1186/1471-2105-8-424>
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. <https://doi.org/10.1107/S0021889892009944>
- Lee, J., Nyenhuis, D. A., Nelson, E. A., Cafiso, D. S., White, J. M., & Tamm, L. K. (2017). Structure of the Ebola virus envelope protein MPER/TM domain and its interaction with the fusion loop explains their fusion activity. *Proceedings of the National Academy of Sciences of the United States of America*, 114(38), E7987–E7996. <https://doi.org/10.1073/pnas.1708052114>
- Liu, H., & Irvine, D. J. (2015). Guiding principles in the design of molecular bioconjugates for vaccine applications. *Bioconjugate Chemistry*, 26(5), 791–801. <https://doi.org/10.1021/acs.bioconjchem.5b00103>
- Lizbeth, R.-S. G., Jazmín, G.-M., José, C.-B., & Marlet, M.-A. (2020). Immunoinformatics study to search epitopes of spike glycoprotein from SARS-CoV-2 as potential vaccine. *Journal of Biomolecular Structure and Dynamics*, 1–15.
- Lupton, H., Lambert, R., Bumgardner, D., Moe, J., & Eddy, G. (1980). Inactivated vaccine for Ebola virus efficacious in guinea pig model. *The Lancet*, 316(8207), 1294–1295. [https://doi.org/10.1016/S0140-6736\(80\)92352-1](https://doi.org/10.1016/S0140-6736(80)92352-1)
- Mahanty, S., & Bray, M. (2004). Pathogenesis of filoviral haemorrhagic fevers. *The Lancet*, 363(9414), 487–498. [https://doi.org/10.1016/S1473-3099\(04\)01103-X](https://doi.org/10.1016/S1473-3099(04)01103-X)

- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation*, 11(8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>
- Majee, P., Jain, N., & Kumar, A. (2020). Designing of a multi-epitope vaccine candidate against Nipah virus by in silico approach: A putative prophylactic solution for the deadly virus. *Journal of Biomolecular Structure and Dynamics*, 1–20.
- Martinez, M. J., Volchkova, V. A., Raoul, H., Alazard-Dany, N., Reynard, O., & Volchkov, V. E. (2011). Role of VP30 phosphorylation in the Ebola virus replication cycle. *The Journal of Infectious Diseases*, 204(Suppl 3), S934–S940. <https://doi.org/10.1093/infdis/jir320>
- McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics (Oxford, England)*, 16(4), 404–405. <https://doi.org/10.1093/bioinformatics/16.4.404>
- Miller, B. R., III, McGee, T. D., Jr., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation*, 8(9), 3314–3321. <https://doi.org/10.1021/ct300418h>
- Mora, M., Veggi, D., Santini, L., Pizza, M., & Rappuoli, R. (2003). Reverse vaccinology. *Drug Discovery Today*, 8(10), 459–464. [https://doi.org/10.1016/S1359-6446\(03\)02689-8](https://doi.org/10.1016/S1359-6446(03)02689-8)
- Moxon, R., Reche, P. A., & Rappuoli, R. (2019). Editorial: Reverse vaccinology. *Frontiers in Immunology*, 10, 2776. <https://doi.org/10.3389/fimmu.2019.02776>
- Ojha, R., Pareek, A., Pandey, R. K., Prusty, D., & Prajapati, V. K. (2019). Strategic development of a next-generation multi-epitope vaccine to prevent nipah virus zoonotic infection. *ACS Omega*, 4(8), 13069–13079. <https://doi.org/10.1021/acsomega.9b00944>
- Okumura, A., Pitha, P. M., Yoshimura, A., & Harty, R. N. (2010). Interaction between Ebola virus glycoprotein and host toll-like receptor 4 leads to induction of proinflammatory cytokines and SOCS1. *Journal of Virology*, 84(1), 27–33. <https://doi.org/10.1128/JVI.01462-09>
- Pastor, R. W., Brooks, B. R., & Szabo, A. (1988). An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Molecular Physics*, 65(6), 1409–1419. <https://doi.org/10.1080/00268978800101881>
- Price, D. J., & Brooks, C. L. III (2004). A modified TIP3P water potential for simulation with Ewald summation. *Journal of Chemical Physics*, 121(20), 10096–10103. <https://doi.org/10.1063/1.1808117>
- Reed, S. G., Bertholet, S., Coler, R. N., & Friede, M. (2009). New horizons in adjuvants for vaccine development. *Trends in Immunology*, 30(1), 23–32. <https://doi.org/10.1016/j.it.2008.09.006>
- Roe, D. R., & Cheatham, T. E. (2013). PTraj and CPPtraj: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, 9(7), 3084–3095. <https://doi.org/10.1021/ct400341p>
- Saghazadeh, A., & Rezaei, N. (2017). Implications of Toll-like receptors in Ebola infection. *Expert Opinion on Therapeutic Targets*, 21(4), 415–425. <https://doi.org/10.1080/14728222.2017.1299128>
- Saha, S., & Raghava, G. P. S. (2006). AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Research*, 34(Web Server issue), W202–W209. <https://doi.org/10.1093/nar/gkl343>
- Samad, A., Ahammad, F., Nain, Z., Alam, R., Imon, R. R., Hasan, M., & Rahman, M. S. (2020). Designing a multi-epitope vaccine against SARS-CoV-2: An immunoinformatics approach. *Journal of Biomolecular Structure and Dynamics*, 1–17.
- Sanchez, A., Kiley, M. P., Holloway, B. P., & Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: Organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Research*, 29(3), 215–240. [https://doi.org/10.1016/0168-1702\(93\)90063-S](https://doi.org/10.1016/0168-1702(93)90063-S)
- Sharma, R., Jangid, K., & Anuradha. (2017). Ebola vaccine: How far are we? *Journal of Clinical and Diagnostic Research*, 11(5), DE01–DE04. <https://doi.org/10.7860/JCDR/2017/22184.9863>
- Singh, S., Sk, M. F., Sonawane, A., Kar, P., & Sadhukhan, S. (2020). Plant-derived natural polyphenols as potential antiviral drugs against SARS-CoV-2 via RNA-dependent RNA polymerase (RdRp) inhibition: An in-silico analysis. *Journal of Biomolecular Structure and Dynamics*. <https://doi.org/10.1080/07391102.07392020.01796810>
- Sk, M. F., Jonniya, N. A., & Kar, P. (2020a). Exploring the energetic basis of binding of currently used drugs against HIV-1 subtype CRF01_AE protease via molecular dynamics simulations. *Journal of Biomolecular Structure and Dynamics*, 1–18. <https://doi.org/10.1080/07391102.01794965>
- Sk, M. F., Jonniya, N. A., Roy, R., Poddar, S., & Kar, P. (2020b). Computational investigation of structural dynamics of SARS-CoV-2 methyltransferase-stimulatory factor heterodimer nsp16/nsp10 bound to the cofactor SAM. *Frontiers in Molecular Biosciences*, 7, 353. <https://doi.org/10.3389/fmoleb.2020.590165>
- Sk, M. F., Roy, R., & Kar, P. (2020c). Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. *Journal of Biomolecular Structure and Dynamics*, 1–16. <https://doi.org/10.1080/07391102.07392020.01724196>
- Sk, M. F., Roy, R., Jonniya, N. A., Poddar, S., & Kar, P. (2020d). Elucidating biophysical basis of binding of inhibitors to SARS-CoV-2 main protease by using molecular dynamics simulations and free energy calculations. *Journal of Biomolecular Structure and Dynamics*, 1–13. <https://doi.org/10.1080/07391102.07392020.01768149>
- Sullivan, N. J., Sanchez, A., Rollin, P. E., Yang, Z. Y., & Nabel, G. J. (2000). Development of a preventive vaccine for Ebola virus infection in primates. *Nature*, 408(6812), 605–609. <https://doi.org/10.1038/35046108>
- To, K. K., Chan, J. F., Tsang, A. K., Cheng, V. C., & Yuen, K. Y. (2015). Ebola virus disease: A highly fatal infectious disease reemerging in West Africa. *Microbes and Infection*, 17(2), 84–97. <https://doi.org/10.1016/j.micinf.2014.11.007>
- Wallace, A. C., Laskowski, R. A., & Thornton, J. M. (1995). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering*, 8(2), 127–134. <https://doi.org/10.1093/protein/8.2.127>
- Wang, D., Raja, N. U., Trubey, C. M., Juompan, L. Y., Luo, M., Woraratanadarm, J., Deitz, S. B., Yu, H., Swain, B. M., Moore, K. M., Pratt, W. D., Hart, M. K., & Dong, J. Y. (2006). Development of a cAdVax-based bivalent Ebola virus vaccine that induces immune responses against both the Sudan and Zaire species of Ebola virus. *Journal of Virology*, 80(6), 2738–2746. <https://doi.org/10.1128/JVI.80.6.2738-2746.2006>
- Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, 101(10), 2525–2534. <https://doi.org/10.1016/j.bpj.2011.10.024>
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The iTASSER Suite: Protein structure and function prediction. *Nature Methods*, 12(1), 7–8. <https://doi.org/10.1038/nmeth.3213>
- Zheng, Q., Li, Z., Zhou, S., Zhang, Q., Zhou, L., Fu, X., Yang, L., Ma, Y., & Hao, X. (2017). Heparin-binding hemagglutinin of Mycobacterium tuberculosis is an inhibitor of autophagy. *Frontiers in Cellular and Infection Microbiology*, 7, 33–33. <https://doi.org/10.3389/fcimb.2017.00033>