Heliyon 9 (2023) e18059

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Immunoinformatic and reverse vaccinology-based designing of potent multi-epitope vaccine against Marburgvirus targeting the glycoprotein

Hassan Yousaf [a], Anam Naz [a,*], Naila Zaman [b], Mubashir Hassan [a,c], Ayesha Obaid [d], Faryal Mehwish Awan [d], Syed Sikander Azam [b]

[a] Institute of Molecular Biology and Biotechnology (IMBB), The University of Lahore (UOL), Lahore, Pakistan
[b] Computational Biology Lab, National Center for Bioinformatics, Quaid-i-Azam University, Islamabad, Pakistan
[c] The Steve and Cindy Rasmussen Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH, 43205, USA
[d] Department of Medical Lab Technology, The University of Haripur, Haripur, Pakistan

ARTICLE INFO

Keywords:
Lake Victoria Marburgvirus
Immunoinformatics
Glycoprotein
Molecular docking
TIM-1
Molecular dynamics simulation

ABSTRACT

Marburg virus (MARV) has been a major concern since its first outbreak in 1967. Although the deadly BSL-4 pathogen has been reported in few individuals with sporadic outbreaks following 1967, its rarity commensurate the degree of disease severity. The virus has been known to cause extreme hemorrhagic fever presenting flu-like symptoms (as implicated in COVID-19) with a 90% case fatality rate (CFR). After a number of plausible evidences, it has been observed that the virus usually originates from African fruit bat, *Rousettus aegyptiacus*, who themselves do not indicate any signs of illness. Thus, efforts have been made in the recent years for a universal treatment of the infection, but till date, no such vaccine or therapeutics could circumvent the viral pathogenicity. In an attempt to formulate a vaccine design computationally, we have explored the entire proteome of the virus and found a strong correlation of its glycoprotein (GP) in receptor binding and subsequent role in infection progression. The present study, explores the MARV glycoprotein GP1 and GP2 domains for quality epitopes to elicit an extended immune response design potential vaccine construct using appropriate linkers and adjuvants. Finally, the chimeric vaccine wass evaluated for its binding affinity towards the receptors via molecular docking and molecular dynamics simulation studies. The rare, yet deadly zoonotic infection with mild outbreaks in recent years has flustered an alarming future with various challenges in terms of viral diseases. Thus, our study has aimed to provide novel insights to design potential vaccines by using the predictive framework.

1. Introduction

The massive increase in human population globally has propounded the concept of urbanization and subsequent exposure to a myriad of diseases due to climate and ecological alterations [1]. In past decades, the emergence and rise in spread of RNA type viruses have posed serious threats to the human life, catering to a number of viral pathogens that remain a mysterious entity to researchers due to their surprisingly evolving genomes and immune evasion [2,3]. Unfortunately, even after years of clinical trials, Marburg viruses of

* Corresponding author.
E-mail addresses: anam.naz@imbb.uol.edu.pk, anam.naz88@live.com (A. Naz).

https://doi.org/10.1016/j.heliyon.2023.e18059
Received 27 September 2022; Received in revised form 30 June 2023; Accepted 5 July 2023
Available online 17 July 2023

the family filoviridae have no licensed vaccines or effective therapeutics for the Marburg virus disease (MVD) [4].

The initial rise of Marburg virus (MARV) outbreak is traced back to August 1967, in city of Marburg, Frankfurt and Belgrade. Initially 32 cases were reported, of which 7 (secondary cases) progressed to fatal disease outcomes with a resulting 21% case fatality rate (CFR). The source was found to be imported from Uganda; African green (Vervet) monkeys (*Chlorocebus aethiops*). However, research published in the Lancet in 1967 incorrectly claimed that the unexplained sickness was caused by rickettsia or chlamydia and was widely credited as the first description of Marburg virus disease (MVD) [5,6]. Though, it was not until 1976, that Marburgvirus identified and classified along with Ebolavirus (EBOV) as members of filoviridae; after a sporadic outbreak in Johannes-burg (1975) which was followed by a single fatality (young Australian who traveled from Zimbabwe) out of 3 [7]. Subsequently, further MARV outbreaks occurred sporadically in Nairobi with 2 reported cases in 1980 (CFR: 50%) and one case in 1987 [8,9]. Later, the cases were reported from Koltsovo (1988 & 1990) [10,11], Durba (1998–2000) [12], Uíge (2004–2005) [13,14], Kam-wenge (2007) [15], Colorado (2008) [16], Leiden (2008) [17] and recent outbreaks in Uganda (2012, 2014 & 2017) and Guinea (2021) (https://www.cdc.gov/vhf/marburg/outbreaks/chronology.html). Structurally, the Marburg virion are pleomorphic particles with an appearance of rod- or ring-like, crook- or six-shaped, or branched structures [5,18]. The virus comprises of a negative-sense and non-segmented RNA genome which range from 19,111 to 19,114 nucleotides and encodes 7 mono-cistronic genes in a linear order: nucleoprotein (NP), VP35, VP40, glycoprotein (GP), VP30, VP24, and the polymerase (L) [13]. Although with no clear and comprehensive knowledge on these protein products of the virion in previous years, it is has been found that VP40 and glycoprotein (GP) are the are the virus's two primary immunogens. The only surface protein that promotes target attachment is GP, which is a transmembrane protein (type I), introduced into the viral envelope as homo-trimeric spikes [19,20]. Whereas, VP40 (matrix protein) acts in combination with the glycoprotein for viral budding [5].

Initial attempts of vaccine development included the whole inactivated virions which resulted in contradictory results and hence unsuccessful trials [10]. Additionally, there are some notable studies which revealed the postexposure protection against Marburg hemorrhagic fever in non-human primates (NHPs) using recombinant vesicular stomatitis virus vector vaccines (VSV). In one of these studies, a total of 5 rhesus monkeys that were administered with rVSV MARV vectors (post-exposure) survived a high-dose lethal challenge of MARV suggesting the potential of VSV vaccines as not only preventive but also post exposure treatments [21]. Notably, for VSV vaccines, 6 additional studies evidenced 100% efficacy in NHPs along with the Mire Plos One 2014 study that reported 100% survivability in study animals 400 days post-vaccination [22–27]. However, contrary to the promising outcomes of the VSV vaccines, studies by Geisbert et al. and Woolsey et al. reported less efficacy of the vaccines. The Woolsey et al. study administered non-human primates with a high (1000 PFU) or low (50 PFU) dose of MARV Angola, which was followed by vaccinating the subjects with either a rVSVΔG-MARV-Angola-GP or rVSVN4CT1-MARV-AngolaGP. The results indicated that only one animal from high viral dose administered group (received rVSVΔG-MARV-Angola-GP inoculation) generated the specific IgM and IgG antibodies appearing on day 10 [28]. Moreover, Geisbert et al. study reported less efficacious results with 2/6 NHPs being survived in one group (vaccine administered 48 h after exposure) as compared to 2nd group which resulted in survival of 5/6 NHPs (vaccine administered 24 h after exposure) [29].

Despite the scarcity of data on human clinical trial vaccinations for the Marburg virus, only three such trials have been described in

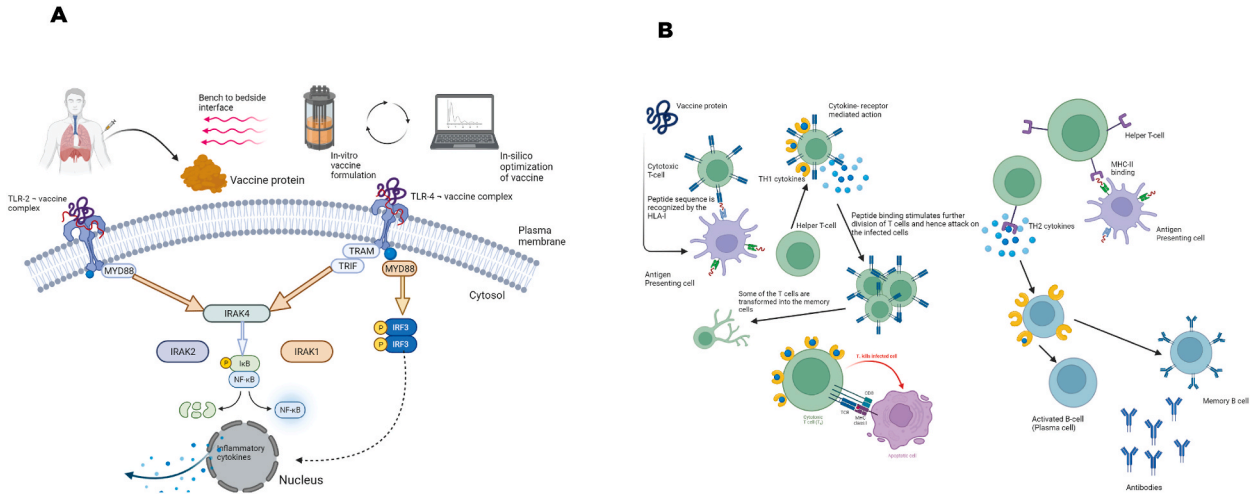


Fig. 1. Mechanism of predicted immune response by the formulated vaccine: **(A)** TLR signaling pathways are mediated by vaccine protein; the TLR-2 homodimer relies on MYD88 activation. (Myeloid Differentiation Primary Response 88), which is an innate signal transduction adaptor protein and works in combination with the 3 active adaptors stimulated by TLR-4, namely: TRAM, TRIF and MYD88 which regulates the IRAK4 (Interleukin 1 Receptor Associated Kinase 4) which in turn stimulates the nuclear factor-kappa-B along with activity of other IRAK molecules for the transcriptional activation of cytokine genes as a response. **(B)** Antigen presentation cells (APCs) will first process the vaccine, and the processed fragments will bind with the MHC-I molecule to drive CTLs (cytotoxic T lymphocytes) to divide and destroy infected cells. Interactions with MHC-II molecules result in the formation of helper T-cells and the stimulation of B cells, which leads to the production of memory B cells. In the final stage of this immunological cycle, antibodies produced by the activated B cell (plasma cell) will take part in infection clearance.

the literature, and only two of those indicated Marburg-specific immunogenicity, both of which were DNA plasmid vaccines (Kibuuka et al. and Sarwar et al.) Only 31% of the individuals (9/29) had an antibody response to the vaccine, according to Kibuuka et al. [30] (trial is registered at ClinicalTrials.gov, number NCT00997607). Similarly, only 23% (7/30) of those vaccinated for both Marburg and Ebola viruses produced an antibody response in the Sarwar et al. [31] experiment (NCT00605514). Four patients in these investigations had negative reactions to the vaccine, resulting in the vaccine being stopped. However, 52% (15/29) of participants in the Marburg vaccination group and 43% (13/30) in the group receiving both vaccines had a T-cell response. Despite this, the European Medicines Agency (EMA) awarded Zabdeno (Ad26.ZEBOV) and Mvabea (MVA-BN-Filo) marketing authorizations against Ebola virus disease in May 2020. The vaccine could be used to prevent MVD, but its effectiveness has yet to be proven in clinical studies.

Thus, the long-standing unsuccessful attempts to come up with a universal treatment against this deadly BSL-4 agent, computational vaccinology is another hopeful futuristic interface that combines the experimental immunology and the advanced computer science algorithms to predict the antigenic targets using the immune-informatics tools. A fundamental unit that generates a cellular or humoral immune response is called an antigenic epitope. A multi-epitope vaccine, made up of a succession of or overlapping peptides, is thus an excellent way to prevent and treat cancers and viral infections. Multiepitope vaccines feature unique design approaches when compared to traditional and single epitope vaccines: i) they contain multiple epitopes from different viral antigens that broaden the spectrum of targeted agents; ii) they contain multiple MHC-restricted epitopes that are recognized by T-cell receptors (TCRs) of multiple T-cell clones; iii) the presence of CTL, Th, and B-cell epitopes elicits strong cellular and humoral immunity; iv) the presence of adjuvants enhances the immunogenicity and long-lasting response [32,33].

To that end, the current work used a variety of immune-informatics methods and databases to predict both $CD4^+$ and $CD8^+$ (in combination and alone) epitopes, as well as consensus B cell epitopes that had the most population coverage globally and across specific regions. The selected peptides were then analyzed by concatenating various linkers to provide flexibility and long-lasting

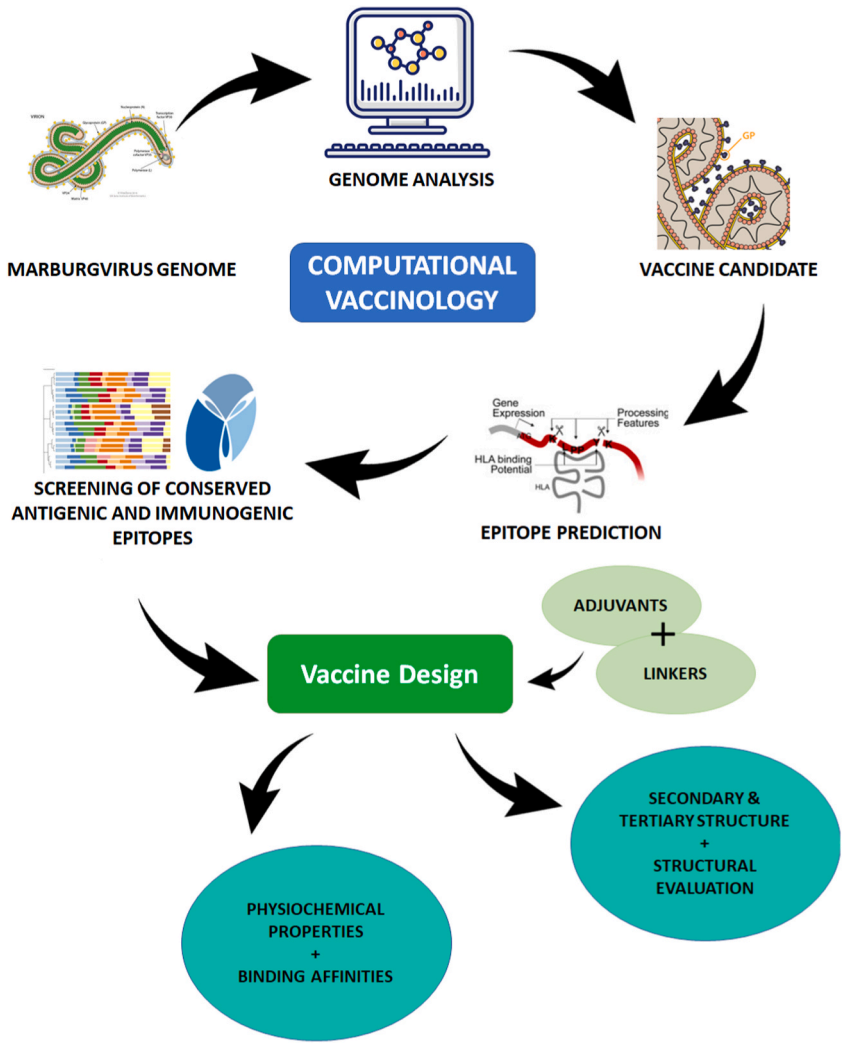


Fig. 2. Flow-diagram of computational vaccinology adapted to design vaccine constructs.

immune response. Furthermore, the physicochemical analysis of the constructed vaccine was performed and the 3D structure was predicted. The chimeric vaccine was then analyzed for its interactions with the HLA-A, HLA-B, HLA-DRB and TLR-2, TLR-4 receptors by molecular docking and the receptor interactions were analyzed by molecular dynamics simulations. The anticipated response of vaccine proposed in this study has been shown in Fig. 1A and B.

2. Materials and methods

2.1. Retrieval of MARV proteome and preliminary analysis

There are a total of 5 phylogenetic branches of MARV with strain Musoke-80 being commonly used in most pre-clinical work. Thus, Musoke-80 reference genome (NC_001608.3) was retrieved from NCBI database [34]. The whole genome of this reference strain was then analyzed using Viral Zone; an online resource of Swiss Institute of Bioinformatics (https://viralzone.expasy.org/) [35]. The entire proteome of musoke-80 was then retrieved via UniProt (https://www.uniprot.org/) in FASTA format. As a measure of an immune response, all the 7 structural proteins were analyzed for antigenicity on *VaxiJen* v2.0 server (http://www.ddg-pharmfac.net/vaxijen/) [36–38]. This server uses auto and cross-covariance (ACC) transformation method to maintain 70–89% prediction accuracy and a threshold of 0.4 was chosen. However, transmembrane proteins were prioritized in the antigenicity prediction analysis and hence the protein with highest antigenic score was selected for further analysis (Fig. 2).

2.2. Prediction of HLA-I interacting epitopes

The retrieved sequence of the MARV envelope glycoprotein was fed into the IEDB TepiTool (http://tools.iedb.org/tepitool/) [39] in order to scrutinize the HLA class-I binding epitopes. We selected a panel of 27 most frequent A and B alleles (A*01:01, A*02:03, A*03:01, A*11:01, A*24:02, A*23:01, A*26:01, A*32:01, A*30:01, A*30:02, A*31:01, A*33:01, A*02:06, A*68:01, A*68:02, A*02:01, B*58:01, B*53:01, B*44:03, B*35:01, B*07:02, B*08:01, B*15:01, B*40:01, B*44:02, B*51:01, B*57:01). We further applied the default settings of low number of peptides which included the removal of duplicate peptides and presented only 9mer peptides (prioritizing low residue peptides is primarily due to the limited ligand fitting capacity of MHC-I molecules which contains a binding groove; closed at both ends and thus only allowing 8–11 residues in length) Finally, we selected peptides using the IEDB recommended percentile rank method (peptides were selected with predicted consensus percentile rank $\leq$1). The percentile rank is calculated by comparing the epitope's allele binding affinity to a wide pool of peptide fragments of similar sizes from the SWISS-PROT database. Percentile rank is calculated using a standard uniform scale and can be used to compare different predictors. As a result, a lower percentile score indicates a higher binding affinity. Prioritizing HLA binding epitopes is intended to ensure an effective and targeted immune response in infected macrophages.

2.3. Prediction of HLA-II interacting epitopes

Similar to HLA-I epitope prediction, we used the TepiTool for HLA-II peptides. For this purpose, we initially deployed the 7-allele method, which relies on the seven most common DR alleles and utilizes the median of consensus percentile ranks of the most common of these alleles, namely, HLA-DRB3*01:01, HLA-DRB5*01:01, HLA-DRB1*03:01, HLA-DRB4*01:01, HLA-DRB1*07:01, HLA-DRB3*02:02, HLA-DRB1*15:01. The peptide length was chosen to be 15mer by default. Peptides with a median consensus percentile $\leq$20 were considered as the good binders. However, we also performed the 2nd method of choosing the pre-selected panel of most frequent DR (DRB4*01:01, DRB3*02:02, DRB1*01:01, DRB1*04:01, DRB5*01:01, DRB1*03:01, DRB1*04:05, DRB1*07:01, DRB1*08:02, DRB1*09:01, DRB1*11:01, DRB4*01:01, DRB1*12:01, DRB1*13:02, DRB3*01:01, DRB1*15:01), DP (DPA1*03:01/DPB1*04:02, DPA1*02:01/DPB1*05:01, DPA1*02:01/DPB1*01:01, DPA1*01:03/DPB1*02:01 DPA1*01/DPB1*04:01), DQ (DQA1*05:01/DQB1*03:01, DQA1*05:01/DQB1*02:01, DQA1*04:01/DQB1*04:02, DQA1*01:01/DQB1*05:01, DQA1*01:02/DQB1*06:02, DQA1*03:01/DQB1*03:02) alleles for a more prolific variety of peptides that could potentially provide a better population coverage. Peptides with predicted consensus percentile rank $\leq$10 were then considered as good binders.

2.4. Linear B-cell epitope prediction

Similar to the T-cells, B-cells can group into different categories according to the cytokine milieu that they produce in cell microenvironment. One of the functional subsets of B-cell is known as the regulatory B cells (Bregs), are known to play a crucial role in maintaining the balance required for pathogen tolerance. B cells are involved in inducing a humoral immune response and thus play an essential role in vaccine designing. We used the IEDB's BepiPred program (http://tools.iedb.org/bcell/) [40] to predict B-cell epitopes. BepiPred is a random forest algorithm technique educated on epitopes identified from antibody-antigen protein structures (a Random Forest Regression (RF) algorithm is learned using a 5-fold cross validation).

2.5. Sequence conservancy analysis

In computational vaccinology, conservancy analysis is performed in order to evaluate the percentage of epitope distribution among the various strains of the pathogen. For this purpose, we used the QIAGEN CLC Genomics Workbench software in order to evaluate the conserved epitopes of the envelope glycoprotein among the 5 common strains of the Marburgvirus retrieved from UniProt.

2.6. Evaluation and screening of epitopes

In an attempt to get the best results out of minimum but the highly favorable peptides, the selected epitopes were manually scrutinized for their antigenicity via VaxiJen v2.0 server [36], allergenicity by AllerTOP v. 2.0 server (https://www.ddg-pharmfac.net/AllerTOP/) [41]; with 85.3% prediction accuracy at fivefold cross-validation, the system uses k-nearest neighbour (kNN) algorithms, amino acid descriptors, and ACC transformation methods to separate non-allergens from allergens. The epitopes which were antigenic (threshold-0.4) were then subjected to immunogenicity calculation. Immunogenicity scores for all MHC-I related proteins were calculated using the Immune Epitope Database (IEDB) program. [42]. Toxin Pred server (http://crdd.osdd.net/raghava/toxinpred/) [43] was utilized in order to ensure that all the selected epitopes were non-toxic.

2.7. Population Coverage

One of the most important requirements for a vaccine to be effective is that it provide protection against a wide range of ethnic groups around the world. The extreme polymorphism of MHC-restricted peptides, on the other hand, results in a wide range of MHC derived frequencies in people of various races. Thus, we used IEDB Population Coverage tool (http://tools.iedb.org/population/) [44]. The fact that various HLA types are expressed at significantly varying rates in different ethnicities complicates the question of population coverage in regard to MHC polymorphism. As a result, if not carefully considered, a vaccination or diagnostic could be developed with ethnically biased population coverage. We aimed to enhance the number of best antigenic peptides in the receptor-binding domain to cover the maximum global population given the limited quantity of quality peptides. This population coverage server approach is based on the Allele Frequencies Net Database (AFND), which contains 1066 alleles of the HLA gene and covers the literature of 1081 populations. [45–47].

2.8. Vaccine construction

The final chimeric vaccine was constructed by assembling the finalized 9 epitopes. In order to build the best possible vaccine sequence, we generated 7 crude sequence combinations (S1V1, S2V1, S3V1, S4V1, S5V1, S6V1, S7V1). Initially, all the selected $CD4^{+}$, $CD8^{+}$ and B cell epitopes were separated by linkers which are known to be the short sequences of amino acids and are eventually cleaved off by the activity of proteasomes and lysosomes so the epitopes can interact with the respective HLA class. Before constructing the final vaccine sequence, we evaluated the best order out of 7 combinations through different immunological filters such as antigenicity, allergenicity and to ensure that every possible sequence was non-toxic. Various adjuvants were linked to the sequence (to assess the antigenicity of each) to improve the antigenicity of the vaccine, including 50S ribosomal protein L7/L12 (50S rib) and RS09, beta-defensin, truncated Ov-ASP-1 Protein, flagellin, and heparin-binding hemagglutinin (HBH). Beta defensin adjuvant has been chosen based on its high protein folding and flexibility resulting in better vaccine candidate [1]. Also, it has been reported to be an immunomodulator along with antimicrobial and antiviral effect [2].

2.9. Physicochemical analysis

The formulated chimeric vaccine construct was evaluated for uncompromised properties such as antigenicity (analyzed via VaxiJen v2.0 server with a default threshold of 0.4), allergenicity (via AllerTOP v. 2.0 to analyze the probability of vaccine construct to elicit any allergen reaction) and toxicity (via Toxin Pred server). Additionally, the vaccine protein was validated for its solubility via ProteinSol (https://protein-sol.manchester.ac.uk/) and SOLpro (http://scratch.proteomics.ics.uci.edu/) [48, 49]. After the basic preliminary analysis, physicochemical properties of the vaccine were computed by feeding the FASTA sequence into the to ExPASy server [50] which deploys the ProtParam tool to compute various physicochemical parameters which include the: molecular weight, estimated half-life, theoretical pI, instability index, aliphatic index, GRAVY value and amino acid composition.

2.10. Secondary and tertiary (3D) structure prediction

The secondary structure of the physicochemical-validated vaccine construct was then predicted using the different available servers: PDBsum [51], PSIPRED v4.0 [52], and SOPMA (Self-Optimized Prediction Method with Alignment) [53]. The PDBsum is a visual interface tool that depicts the contents of any sequence or 3D structure deposited in the Protein Data Bank (PDB) at a glance. It depicts the structure's molecule(s), as well as schematic schematics of their interactions (e.g., protein chains, DNA, ligands, and metal ions). Similarly, PSIPRED v4.0 and SOPMA are two other reliable sources for structure prediction. Hence, we utilized all three of the servers to get the most suitable results.

Finally, the tertiary structure was predicted by using Robetta; a protein structure prediction service that is constantly evaluated through CAMEO and includes relatively fast and accurate deep learning-based methods, RoseTTAFold and TrRosetta, and an interactive submission interface that allows custom sequence alignments for homology modelling, constraints, local fragments, and more, as well as a protein structure prediction service that is constantly evaluated through CAMEO and includes relatively fast and accurate deep learning-based methods, RoseTTAFold and TrRosetta, and an interactive submission interface that allows custom sequence alignments [54].

2.11. 3D structure refinement and validation

The highly stable and confidence level model was then refined via Galaxy Refine (https://galaxy.seoklab.org/); an online server which is based on the methods of repeated structure perturbations and overall structure relaxation of the protein model. The server generates many models which deviate from the crude protein structure by mild and aggressive relaxation methods and hence to apply the overall structural relaxation, molecular dynamics simulation is performed. For this purpose, the server performs structure perturbation only to clusters of side-chains in-case of model 1 and more agitative perturbations are performed in the secondary elements and loops of the model 2–5. Finally, for more enhanced results, we submitted the best refined model to the GalaxyRefine2, which is an advanced version of simple refinement. GalaxyRefine2 runs on intense iterative optimizations (utilizing various operators for geometric configurations) to improve the quality of query model. Compared to the previous method, GalaxyRefine, which relaxes an initial structure using MD and side chain perturbation, this algorithm runs on the application of global operators (e.g. anisotropic mode perturbations and secondary structure perturbations) and local operators (e.g. modeling of the protein loops and subsequent hybridizations of the structure), along with structural error predictions and homolog structural information, so that the accuracy of the structure could improve [55]. Finally, the ERRAT, Z-score and the 3D structure of the refined vaccine was validated [56–59].

2.12. Molecular docking

To analyze the interaction of multiepitope vaccine with the immune receptors, we retrieved the pdb structures of TLR-2, TLR-4, HLA-A, HLA-B, HLA-DRB and TIM-1 (T-cell immunoglobin mucin domain 1). For this purpose, protein-protein docking was performed via ClusPro; a widely used protein-protein docking tool. ClusPro server utilizes a numerous advanced options to modify the predicted results, some of the key features include: elimination of unstructured regions, implementation of attraction and repulsion principles, pairwise distance restrictions, homo-multimerization, and heparin-binding site targeting. For general purpose protein-protein docking, the ClusPro algorithm works by rotating the ligand with 70,000 rotations. For each rotation, the ligand is translated in x, y, z axis relative to the receptor on a grid. Thus, the translation with the best score from each rotation is shortlisted. Finally, of the 70,000 rotations, the 1000 rotation/translation combinations that have the lowest score are selected [60–63].

2.13. Binding affinity

Binding Affinity of the proposed vaccine was evaluated with the TLR-4, TLR-2, HLA-A, HLA-B, HLA-DRB using the PRODIGY server [64].

2.14. Molecular dynamics simulation

Both proteins docked with their respective TLR2/4 complex were subjected to molecular dynamics (MD) simulations for time period of 100 ns each. For this purpose, LEaP program was used to prepare topology and coordinate files by adding (Na+/Cl-) ions to neutralize the system followed by 10 Å TIP3P tetrahedral solvation box. In the next step, AMBER force fields; namely ff14SB and General Amber force field (GAFF) libraries were employed to extract required information about Phi-Psi angles. Resulting systems were then subjected to AMBER sander to prepare input files for MD simulations, which comprised 1500 steps of minimization with 200 kcal/mol/Å2 constraints followed by 5 kcal/mol/Å2 constraints only on carbon atoms. Heating was carried out at the temperature of 300 K that was controlled by Langevin thermostat using 5 kcal/mol/Å2 constraints. The system was then equilibrated for 100 ps before going to the production run of 100 ns each with non-bonded cut-off distance set to 8.0 Å with AMBER 16 [65]. Results were extracted with CPPTRAJ [66] and visualized with multiple molecular graphics programs; namely Visual Molecular Dynamics (VMD) [67], BIOVIA Discovery Studio (DS) [68], and USCF Chimera [69].

2.15. Optimization and computational cloning of the proposed vaccine

For codon optimization, Java Codon (http://www.jcat.de/) was used which efficiently performed reverse translation of the vaccine protein to produce a more appropriate DNA sequence [70]. The cDNA was then employed in *E. coli* for expression (K12 strain). This was accomplished by avoiding restriction enzyme cleavage sites, rho-independent transcription terminators, and bacterial ribosome binding sites. Using the Snap-Gene software, the chimeric vaccine's optimized cDNA sequence was subsequently put into the pET-28a (+) vector.

2.16. Immune response simulation

For immune simulation, we used the classic C-IMMSIM server. The server utilizes the machine learning based algorithm that utilizes the position-specifc scoring matrices (PSSM) for epitope prediction hence anticipates the immune profile of vaccine administered. The recommended average time span between the vaccine dose is 28 days. Thus, for 3 doses, we initialized with 1050 simulation steps (which covers the time period of a year) and subsequently we chose time steps of 1, 84, 170 for dose 1, 2 and 3 respectively (the server equates 1 simulation step to 8 h) [71].

3. Results

3.1. Retrieval of protein sequence and preliminary analysis

The sequence of envelope glycoprotein (GP) was retrieved from UniProt (Accession ID: P35253) and was subjected for further analysis. Since glycoprotein is the major immunogen which is involved in the interaction with the potential receptors, several breakthrough studies have reported the potential role and interactions of GP with various receptors. T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus, according to a study published in the Proceedings of the National Academy of Sciences in 2011. However, there is little subsequent evidence for these reports. Thus, we selected envelope glycoprotein (GP) as it showed the highest antigenicity score of 0.5281. Using the AllerTOP v. 2.0 test, the protein was likewise shown to be non-allergenic (the method is based upon the transformation of multiple protein sequences into uniform and equal length vectors by the process of auto cross covariance abbreviated as ACC).

3.2. Evaluation and screening of epitopes

An ideal vaccine candidate should mimic and fulfil the criteria propounded by the natural bodily response to any infection. Thus, CTL and HTL epitopes are of significant importance to elicit a long-lasting immune response and to stimulate the activation of B cell antibody production. Epitopes with strong binding affinities along with the most common experimentally validated and maximum covered alleles are considered the best choice for multi-epitope vaccine design. But contrary to that, in case of MARV there is not a wide spectrum of alleles that is common to the peptides predicted. Thus, in order to filter out the epitopes with potent defense capabilities, we filtered out the best-known epitopes along with the highest number of alleles and a maximum population coverage.

3.3. HLA-I epitope binding prediction

Initially, a total of 18,171 raw epitope data was generated on IEDB TepiTool. Upon further scrutinization, we filtered the highly conserved epitopes based on the percentile rank and thus the best top 52 epitopes with a significant binding affinity to the panel of 27 alleles were selected. The list of epitopes generated is based on the various immune filters which are based on: promiscuity, antigenicity and immunogenicity. Further investigation indicated that 11 of the 52 epitopes are found in the receptor-binding domain (RBD). Fortunately, we found the best accessible CTL epitope in the RBD area after further curation. Antigenic epitopes are known to elicit a large number of antibody titers in the fight against infection. All the 52 antigenic epitopes showed a significant antigenic potential which was predicted via VaxiJen v2.0 server. Upon further analysis of immunogenicity prediction, 28 epitopes which were non-immunogenic (negative scoring) were discarded. The remaining 24 epitopes were then rigorously evaluated for their quality and priority was given to the maximum allele binders residing in the RBD.

3.4. HLA-II epitope binding prediction

In this prediction, a total of 195 epitopes were generated via IEDB TepiTool, of which 9 epitopes were computed by the 7-allele method which showed a higher binding affinity towards the HLA-DRB alleles. Since, it presented epitopes with a significant median consensus percentile rank, thus this score reflects the predominant nature of the HLA-DRB allele supertype in pathogen presentation to the immune system. Due to the lack of epitopes in the receptor binding domain (RBD), we further navigated to filter out epitopes generated via the 26 panel of alleles. Initially, a total of 186 epitopes were generated, of which priority was given to the maximum number of allele-binding epitopes with a higher antigenicity and their probable location in the RBD.

Meanwhile, only 6 epitopes out of 9 fulfilled the criteria of antigenicity computed through the abovementioned 7-allele method. Out of these 6 six epitopes, only a single putative 15-mer epitope was chosen (581RTFSLINRHAIDFLL595) which presented to be highly conserved and antigenic (0.7991) and shown to be overlapping with one of the CTL epitopes. Upon further scrutinization, 4 epitopes were retained from the 186 raw epitopes generated via the panel of 26 most frequent HLA-DRB, DPA and DQA supertypes. Of these, the 3 epitopes: (^{521}VQEDDLAAGLSWIPF^{535}), (^{563}ANQTAKSLELLLRVT^{577}), (^{652}LTNLGILLLLSIAVL^{666}) were correspondents of the GP2 domain of glycoprotein and a single epitope (90KTCYNISVTDPSGKS104) presented to be part of the RBD of GP1. The binding energy between a 15-mer peptide epitope and the HLA-II receptor protein is primarily obtained from the peptide's 9-mer core. However, the presence of additional amino acid chain aid in the promiscuous binding of epitopes to the receptor groove and maintain the overall stability of the complex. Thus, 15-mer HLA-II epitopes are usually recommended. Hence, we finalized a total of 5 highly conserved and antigenic HTL epitopes which showed to be non-allergenic and non-toxins.

3.5. Prediction of linear B cell epitopes

The IEDB Bepipred linear epitope prediction algorithm predicted a total of 25 epitopes. Peptides having a mean score of more than 0.35 are automatically regarded to be the most effective. The higher is the mean score of an epitope, the greater will be its specificity and lower will the sensitivity. Thus, in an attempt to choose, the best epitope, we selected the highly antigenic 16-mer epitope (^{74}TGVPPKNVEYTEEEAK^{89}) which resides in the RBD of GP1. Thus, we further analyzed the quality of epitope by conjugating the peptide with previously passed HLA-I and II binders with maximum conservancy, antigenicity, and non-allergenicity. The shortlisted epitopes for the final vaccine construct are shown in Table 1.

3.6. Population Coverage

Global coverage was evaluated for the selected 8 CTL and HTL epitopes. With comparable HLA alleles (IC50 500 nm), population coverage was computed. The results were presented in IC50 nM units. Peptides with IC50 values of 50, 500, and 5000 nM, respectively, are considered high, intermediate, and low affinity. For the population coverage study, alleles with an IC50 value of less than 500 were chosen as the best binders. The selected epitopes of glycoprotein indicated a combined coverage 98.12% with maximum coverage being recorded in regions of West Africa (98.98%), Europe (98.84), East Africa (98.76%) and Central Africa (98.23%) shown in Fig. 3.

3.7. Vaccine design

The vaccine was formulated by using the previously selected 9 epitopes. We used the combination S4V1 since it showed a slightly higher antigenicity value out of all the 7 combinations. The main criteria in constructing the linear vaccine are: 1). It should contain common CTL and HTL epitopes, 2). There should be a higher binding affinity towards the HLA alleles, 3). It should be antigenic, immunogenic, non-allergen and non-toxic. Based on these criteria, we separated the epitopes by GPGPG linkers (shown in Fig. 4). The chosen linker was prioritized over others because it prevents the formation of junctions and non-favorable loops with an efficient immune processing of antigen. Finally, the vaccine was formulated by the addition of beta-defensin-3 adjuvant at the N-terminal (which acts as a TLR-4 agonist) and was joined to the string of epitopes via EAAAK linker to boost the long-lasting immunity and to provide a slight rigidity in the final vaccine construct. Vaccines having beta-defensin or other defensins as adjuvants have been reported, to activate the primary innate antiviral immune response both in vivo and *in vitro* and mediate other immunomodulatory activities against a number of viruses [3].

3.8. Physicochemical analysis and solubility profile

The chimeric vaccine protein was validated for the basic physicochemical analysis to proceed further validation. The sequence protein showed to be highly antigenic with a score of 0.6653 (*VaxiJen v2.0*, threshold = 0.4) and a predicted probability of antigenicity: 0.891501 on AntigenPro, non-allergenic and non-toxin. The vaccine construct also indicated a predicted scaled solubility of 0.775 on Protein SOL. Because the population average (PopAvrSol) for the experimental dataset is 0.45, any scaled solubility value

Table 1
List of finalized epitopes for chimeric vaccine construct.

No.	Epitope(s)	Binding alleles	MHC-I/II	Domain in GP
1	^{61}DSPLEASKR^{69}	HLA-A*68:01, HLA-A*33:01	MHC-I	GP1, RBD
2	^{62}SPLEASKRW^{70}	HLA-B*53:01, HLA-B*35:01, HLA-B*58:01, HLA-B*44:02, HLA-B*44:03, HLA-B*57:01, HLA-B*07:02, HLA-B*51:01	MHC-I	
3	^{129}NPHAQGIAL^{137}	HLA-B*07:02, HLA-B*35:01, HLA-B*08:01, HLA-B*53:01, HLA-B*51:01	MHC-I	GP1, RBD
4	^{90}KTCYNISVTDPSGKS^{104}	HLA-DRB1*04:01, HLA-DRB1*04:05	MHC-II	GP1, RBD
5	^{521}VQEDDLAAGLSWIPF^{535}	HLA-DQA1*03:01, HLA-DQA1*04:01	MHC-II	GP2, extracellular
6	^{563}ANQTAKSLELLLRVT^{577}	HLA-DPA1*02:01, HLA-DPA1*02:01, HLA-DPA1*03:01	MHC-II	GP2, extracellular
7	^{652}LTNLGILLLLSIAVL^{666}	HLA-DPA1*01, HLA-DPA1*02:01, HLA-DPA1*03:01, HLA-DQA1*01:02, HLA-DQA1*04:01, HLA-DRB1*01:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*11:01, HLA-DRB1*15:01, HLA-DRB4*01:01	MHC-II	GP2, helical
8	^{581}RTFSLINRHAIDFLL^{595}	HLA-DRB1*03:01, HLA-DRB1*07:01, HLA-DRB1*15:01, HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01, HLA-DRB5*01:01, HLA-A*03:01, HLA-A*30:02, HLA-A*30:01, HLA-A*31:01, HLA-A*11:01, HLA-A*32:01,	MHC-I & II	GP2, extracellular
9	^{74}TGVPPKNVEYTEEEAK^{89}		B-cell	GP1, RBD

Fig. 3. Population coverage for the combined MHC-I and MHC-II epitopes of the chimeric vaccine is shown. Generally, the African region has the highest coverage which is the top priority region due to the past outbreaks followed by Europe and America. The combined world population coverage is shown to be 98.12% with 76.78% and 91.92% for MHC-I and MHC-II respectively.

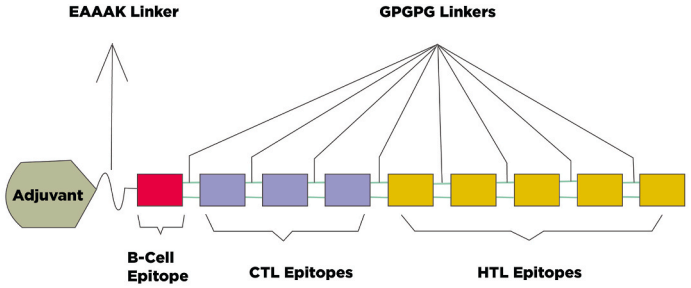
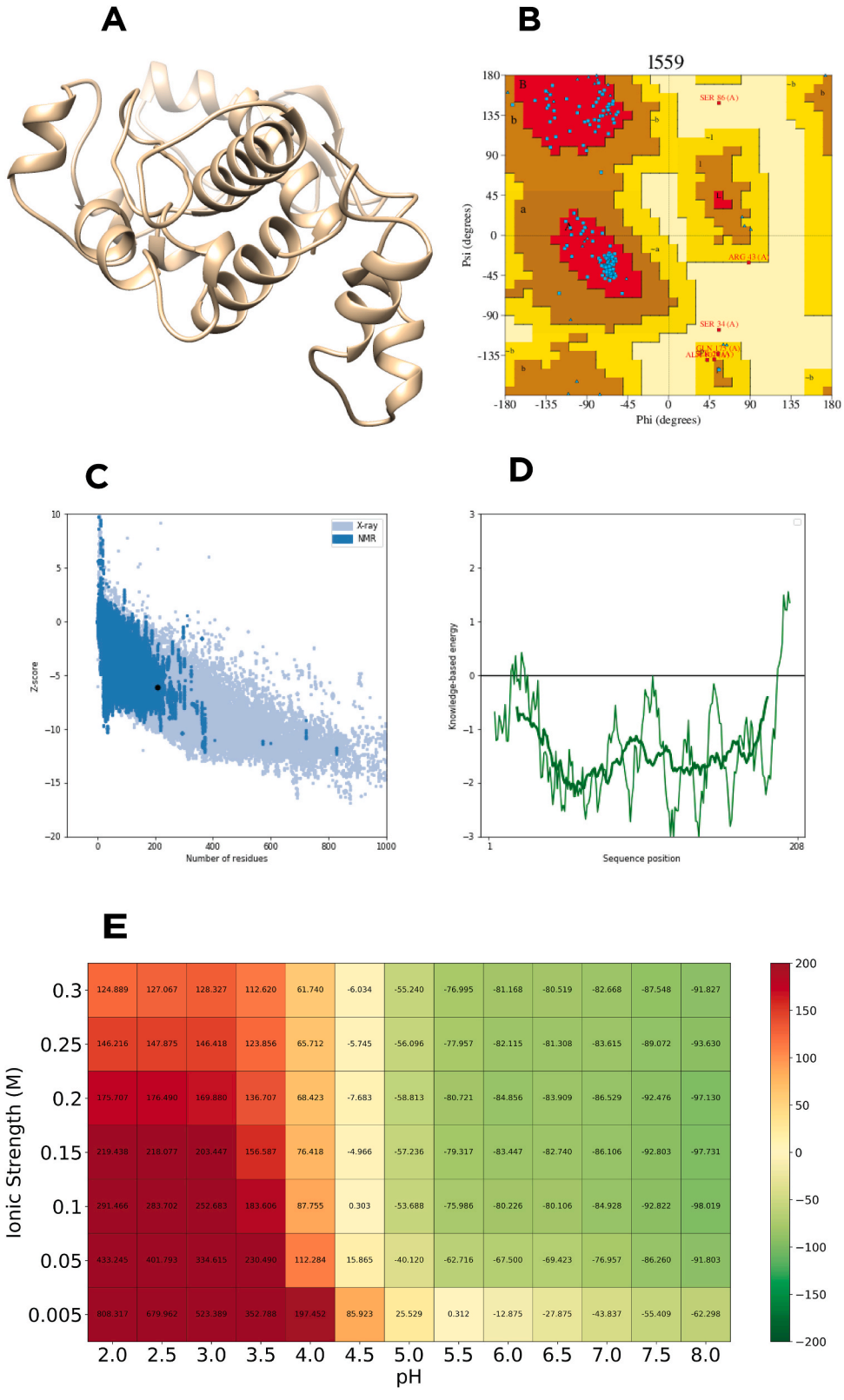


Fig. 4. Linear stretch of final MarVac chimeric vaccine: The structure depicts a final 208 amino acids long vaccine protein sequence consisting of an adjuvant (grey) at the N-terminal attached to the flanked sequence through EAAAK linkers. Followed by the linker are total 9 epitopes separated by the GPGPG linkers. There is a single 16-mer B-cell peptide (green), 3 CTL epitopes (light blue) and 5 HTL epitopes (orange). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Physicochemical profile of the proposed vaccine sequence.

Parameter		Value	Remarks
Molecular Weight		21478.69 g/mol	Appropriate
pI		9.44	Alkaline (basic)
Instability index		37.94	Stable
Aliphatic index		78.37	Thermostable
GRAVY		−0.357	Hydrophilic
Estimated half-life	mammalian reticulocytes, *in vitro*	30 h	
	yeast, in vivo	>20 h	
	Escherichia coli, in vivo	>10 h	
Antigenicity	VaxiJen v2.0	0.6653	Highly antigenic
	AntigenPro	0.891501	
Solubility	Protein SOL	0.775	Highly soluble
	SOLpro	0.886797	



(caption on next page)

Fig. 5. Refined tertiary structure of the vaccine and further structure validation plots. **A).** Robetta server generated structure of the proposed model 3. **B).** The refined vaccine protein's Ramachandran plot revealed that 93.3% of residues were in highly favored regions, 2.7% in further allowed regions, 2.0% in generously allowed regions, and 2.0% in disallowed regions. **C).** The plot depicts the overall protein quality, which has a Z-score of −6.13. Different colors distinguish groupings of structures from different sources (X-ray, NMR) in this plot. It can be used to see if the input structure's z-score falls within the range of scores found for native proteins of similar size. **D).** An energy graph depicts the local model quality in the plot. Positive values correspond to sections of the input structure that are problematic or incorrect. In general, the quality of our local model is outstanding, with very little uncertain territory. **E).** Energy heatmap of the refined 3D complex shows protein stability and charge across 91 combined pH and ionic strength conditions. Thus, it is visible that with the increase in pH and ionic strength, there is lesser free energy value, corresponding to the overall reliability of the vaccine protein. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

more than 0.45 is anticipated to have a higher solubility than the average soluble *E. coli* protein from the experimental solubility dataset. It also presented to be basic in nature with a theoretical pI of 9.44 and a molecular weight of 21478.69 g/mol. The vaccine construct's instability index (II) was calculated to be 37.94, indicating that the protein is stable. Further it indicated an aliphatic index of 78.37 which depicted that the protein is thermostable in nature and a Grand average of hydropathicity (GRAVY) as: 0.357, which represents that our vaccine construct is hydrophilic in nature and thus, it can react with water molecules. The estimated half-life was predicted to be: 30 h (mammalian reticulocytes, invitro), >20 h (yeast, in vivo), and >10 h (*Escherichia coli*, in vivo). The chemical formula of chimeric vaccine was shown to be: C946H1534N276O280S7. Table 2 represents the key physicochemical parameters of the proposed vaccine sequence.

3.9. Secondary and tertiary structure prediction

The secondary structure as predicted by the PDBsum, showed 2 antiparallel beta sheets, 2 beta hairpins, 12 helices, 8 helix-helix interactions, 15 beta turns, 2 gamma turns, 4 strands and 1 disulfide bond. A beta turn is fully defined when the distance between the C-alpha atom of residue I and the C-alpha atom of residue i+3 is less than 7, and there is no center (2 residues) residue helical. The gamma turn, on the other hand, is made up of three residues: I i+1, and i+2. If a hydrogen bond forms between residues I and i+2, the phi and psi angles of residue i+1 lie within 40° in terms of geometrical configuration. Furthermore, the secondary structure only had one beta bulge, which defined a disturbed region in the beta sheet and so interfered with the typical string of hydrogen bonds by adding an extra residue.

The secondary structure was further validated by structure prediction through SOPMA and PSIPRED which presented a reasonably well-defined structure with zero ambiguous states.

The online tertiary structure prediction server called as Robetta was further used to finally predict a 3-dimensional structure of the protein sequence. The fed sequence is parsed into putative domains and confirmations to generate 3D structures by either i) Comparative modeling or ii) de novo prediction tools. Thus, by utilizing such an approach, we received 5 models of our protein sequence. We further analyzed the structural models and chose model no. 3 as it presented the highest confidence level.

3.10. 3D structure refinement and validation

The obtained model no.3 was further refined and showed that 93.3% of the residues were in most favored regions, 2.7% in additional allowed regions, 2.0% in generously allowed regions, 2.0% in disallowed regions. Finally, the ERRAT analysis of the vaccine protein showed an overall quality factor of 97% and the 3D structure was verified and passed because 87.50% of the residues averaged 3D-1D score≥0.2 (shown in Fig. 5). Thus, the analysis authenticated the overall stability and reliability of our vaccine construct (Fig. 5A–D). A final validation of the vaccine construct was also carried by constructing the energy heatmap (Fig. 5 (E).

3.11. Molecular docking analysis

3.11.1. Docking of vaccine protein with TLR-4 & TLR-2 receptor

TLR-4 and TLR-2 are well-known toll-like receptors that play an important part in the recognition of both structural and non-structural proteins, and hence in cytokine generation. These cell surface receptors identify the surface glycoproteins as culprits from the target virion. Thus, TLR agonists play a crucial role in immune response and the subsequent viral clearance. In an attempt to study the interactions between the TL receptors with the proposed vaccine, ClusPro was utilized. The server displayed the best 10 models which depicted the binding energies between the 2 protein molecules. The complex models were created using the PIPER approach, which centers the receptor molecule in the coordinate system and analyses the ligand's likely rotational and translational positions at the set level of discretization. For four distinct coefficients, the 10 most favorable models were developed, including: I balanced models ii) electrostatically preferred models iii) vanderWaal + Electric and iv) hydrophobic-favored. $E = w1E_{rep} + w2E_{attr} + w3Ee_{lec} + w4E_{DARS}$ was the energy weight for each coefficient, where E_{elec} is an electrostatic energy term, whereas E_{rep} and E_{attr} are the repulsive and attractive contributions to the van der Waals interaction energy, respectively. The Decoys as the Reference State (DARS) approach created E_{DARS}, involved in desolvation contributions which are mostly represented by a pairwise structure-based potential.

For vaccine and TLR-4 complex, the best model was prioritized by the default top position (shown in Fig. 6(i). The models are prioritized in order of the number of members covered by each cluster. This is calculated by computing the IRMSD values for each pair of structures among the 1000 and finding the structure with the most neighbors within 9 IRMSD radius. The model zero was predicted

to have the highest number of members and a significantly low binding energy value of −1059.4, −1168.2, −1376.9, −287.7. Thus, the TLR-4 indicated a good binding interaction with our vaccine construct. To further examine the schematic illustration of the interacting residues of the docked complex, PDBsum was utilized. For our vaccine construct (chain B) and TLR-4 (chain A), 22 and 24 residues were actively involved with interface areas spanning 1368 and 1318 Å2 respectively. For the docked complex, it was observed that there was a total of 13 H-bonds involved: [chain- B(MarVac) ~ chain-A(TLR-4); 178–382, 185–456, 188–505, 187–529, 155–550, 194–555, 200–578, 59–598, 153–603, 8–603, 38–605, 4–606, 200–606. Further structural interactions revealed that ASP178 and ARG382 formed a H-bond at distance of 2.87 Å, TRP185 and HIS456 formed H-bond at 3.07 Å distance. Similarly, PHE188 and GLN505, PRO187 and HIS529, THR155 and ASP550, ARG194 and HIS555, ASN200 and GLN578, GLU59 and ARG598, GLY153 and GLU603, LYS8 and GLU603, ARG38 and GLU605, ASN4 and ARG606, ASN200 and ARG606 formed H-bonds at distance of 2.72 Å, 3.20 Å, 2.85 Å, 2.79 Å, 2.84 Å, 2.69 Å, 2.84, 2.54, 2.73, 2.67, 2.66 respectively. Finally, the complex also predicted to have 5 salt bridges shown in Fig. 6 (1A – 1C).

For vaccine and TLR-2 complex, there were a total of 29 models with differing binding energies predicted, of which 10 best models were displayed. Since, the model zero presented with the highest number of 131 members and lower binding energy values of −943.6, −964.8, −1443.7 and −283.1, thus it was considered the best of all 10 docking complexes. Schematic illustration of the interacting residues shown in Fig. 6(ii) depicted the potential interactions between 19 residues of chain-A (TLR-2) and 17 residues of chain-B (MarVac) with interface areas of 895 Å and 923 Å respectively. There were a total 7 H-bonds and 1 salt bridge interaction between the 2 chains. Hydrogen bonds were present between residues: PRO320 and GLY99, ARG321 and ASN100, ARG321 and ASP178, TYR323 and SER164, TYR323 and ASN100, LEU354 and PRO190, HIS398 and ASN156 at distance of 3.05 Å, 2.95 Å, 2.98 Å, 3.03 Å, 3.02 Å, 3.19 Å, 3.08 Å shown in Fig. 6 (2A – 2C). The interactions between the anticipated envelope glycoprotein receptor; TIM-1 (T-cell immunoglobin mucin domain-1) and the vaccine protein are shown in Fig. 7.

3.11.2. Docking of proposed vaccine with HLA alleles (HLA-A, B and DRB)

For docking analysis of vaccine with the HLA-A allele, ClusPro generated 29 clusters with the top cluster having a balanced lowest energy of −996.2 and with an area of 89 members. Similarly, the lowest energy scores for remaining 3 coefficients were computed to be −1033.5, −1150.4 and −258.0 with 76, 127 and 122 members respectively. Interacting residues for chain-A (HLA-A) and chain-B (MarVac) were 30 and 26 for chain-B (MarVac) with interface areas of 1460 Å and 1568 Å respectively. A total of 24 H-bonds were present in between residues: TYR7 and ARG194, GLU53 and ARG36, GLN54 and ARG36, GLU55 and ARG36, GLU55 and ARG36, GLU63 and ARG194, GLU63 and ARG194, ARG65 and SER197, HIS70 and PRO190, TRP107 and THR155, ARG108 and ASN156, TYR159 and ARG194, GLU166 and ASN200, GLU166 and ARG201, ARG169 and THR155, ARG169 and THR155, ARG170 and ASN4, ARG170 and ASN200, TYR171 and ARG194, GLU173 and ARG38, ASN174 and ARG38, LYS176 and GLU59, GLU177 and LEU21, GLU177 and VAL20 at varying distance. Finally, the complex also showed to contain 231 non-bonded contacts and 5 salt bridges.

For HLA-B and vaccine complex, the best scoring cluster represented a balanced lowest and center energy of −963.5 and −887.7 with 127 member structures in the set vicinity of receptor-ligand interaction. Subsequently, the lowest energy values were presented to be −906.0, −1230.9 and −279.6 with 98, 217 and 97 structural members respectively. The residue interface depicted to have 20 hydrogen bonds and 3 salt bridges with area of 1397 Å and 1489 for chain-A (HLA-B) and chain-B (MarVac) respectively. In case of HLA-DRB, top scoring cluster showed balanced center and lowest energy of −868.3 and −988.5 with 96 members in set vicinity. For the remaining 3 coefficients i.e., electrostatic-favored, hydrophobic-favored and VdW + Elec, lowest energy was shown to be: 1103.6,

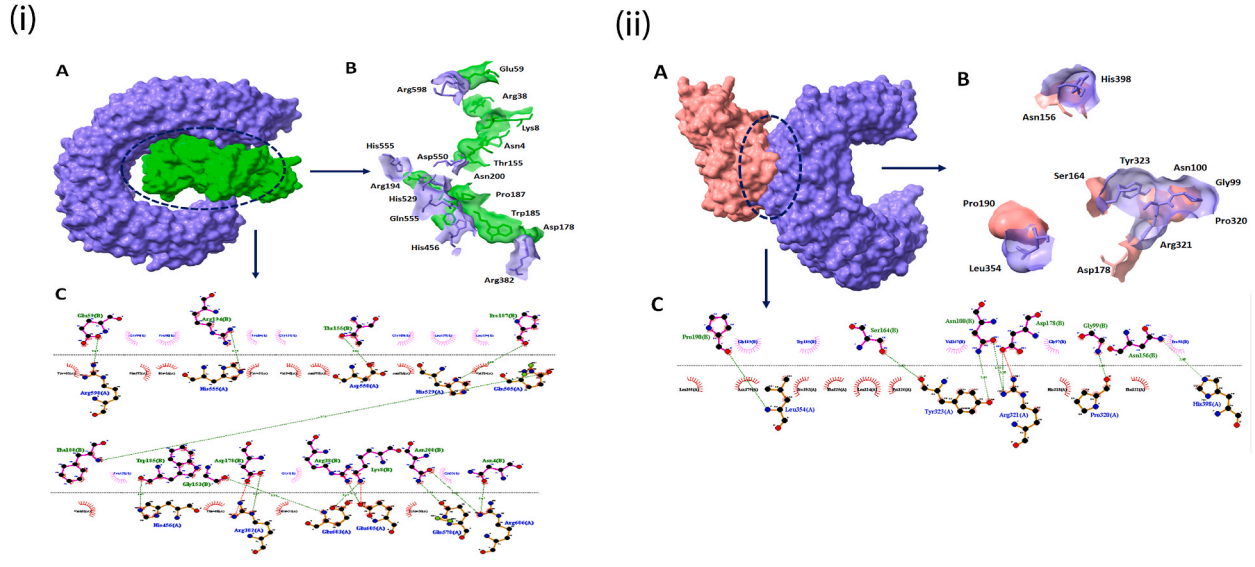


Fig. 6. (iA) Illustration shows the docked complex of vaccine and the TLR-4. **(iB).** Interactions between residues of 2 chains; chain A (TLR-4) and chain B (vaccine) are shown. **(ii).** Illustration shows the interactions between the residues of 2 chains; chain A (TLR-2) and chain B (vaccine).

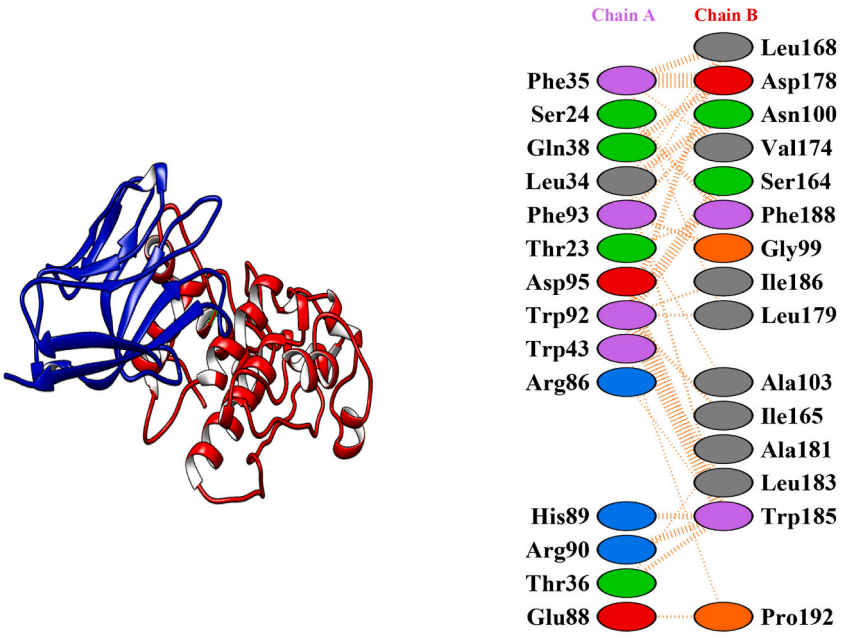


Fig. 7. A deep structural view of the structural residue interaction for TIM-1 receptor and chimeric vaccine.

−1264.6, −291.8 respectively. For HLA-DRB and vaccine complex, the top cluster represented the best lowest energy with maximum number of members. The energy values (lowest and center) of the 6 best docking complexes are depicted in Table 3. The selected docking complexes with residue interactions for the HLA-A, HLA-B, HLA-DRB and vaccine protein are shown if Fig. 8.

3.12. Analysis of binding affinity

In terms of Gibbs free energy (G), the binding affinity of docked complexes represents a critical energy number that indicates whether a suitable interaction will occur under given conditions in the body. The PRODIGY server was used to assess the binding affinity of the 5 docked complexes. HLA-A, HLA-B, HLA-DRB, TLR-2, and TLR-4 complexes had G values of −13.0 kcal mol-1, -11.4 kcal mol-1, -9.2 kcal mol-1, -8.8 kcal mol-1, and -14.5 kcal mol-1, respectively. The negative values shown here reflects the energetically feasible nature of our docking complexes and hence favorable results. The Gibbs free energy of the docking complexes along with their dissociation constant (K_d) are shown in Table 4.

3.13. Molecular dynamics simulations analysis

To infer intactness and stability of subsequent proteins with TLR2 and TLR4, both the complexes were subjected to 100 ns MD simulations. Notably distinct molecular patterns in dynamics of two systems were observed signifying their importance in mediating innate and adaptive immune responses against infectious diseases. TLR2 exhibited significant movement of attached protein over the course of 100 ns, which can be seen in Fig. 9 1A. Entire protein shifted to the opposite direction to interact with TLR2 and exhibited significant increase in number of interactions that is depicted in Fig. 9 (1B, 1C). Impact of protein movement can also be seen in RMSD graphs that displayed increasing trend for first 30 ns (Fig. 9 2A). Significant residual movement in RMSF and Beta-factor graphs is indicative of these fluctuations (Fig. 9 2B). Overall, the complex exhibited stability throughout 100 ns MD simulations that is exhibited in Radius of Gyration (Rg) graphs exhibiting rather increase in protein compactness (Fig. 9 2C). The beta factor is also depicted in Fig. 9 2D.

TLR4 on the other hand, exhibited extremely stable system during 100 ns simulations with no significant conformational shifts during the course as shown in Fig. 10 (**1A, 1B**). RMSD graphs with TLR4 exhibited stable trend over 100 ns MD simulations course compared to TLR2 complex as depicted in Fig. 10. However, minor residual fluctuations at the N-terminal of protein due to the presence of loop can be seen in RMSF graphs (Fig. 10), which do not affect overall structural dynamics of the complex under study shown in Fig. 10 (2A – 2D). The protein in complex with TLR2 exhibited an average RMSD of 7.1 Å due to residual fluctuations in residues 560–620. Despite the higher RMSD, the protein maintained compactness till the 100 ns simulations. Clearly, the protein in complex with TLR4 is more compact with and average Rg value of 36.0 Å and exhibited stability with an average RMSD value of 3.2 Å. Residual fluctuations was only observed in the N-terminal which does not affect the overall interaction and stability of protein TLR4 complex. Similarly, MMGBS and MMPBSA exhibited the same behavior. TLR4 has illustrated the higher binding potential compared to TLR2 with the values −72.80 and −13.08, respectively. Complete details of free energy values are depicted in Table 5.

Table 3
Features of the shortlisted docked complexes.

Complex type	Coefficient	Cluster number	Members	Representative	Weighted Score
HLA-Ã MarVac	Balanced	0	89	Center	−894.7
				Lowest Energy	−996.2
	Electrostatic-favored	0	76	Center	−921.2
				Lowest Energy	−1033.5
	Hydrophobic-favored	0	127	Center	−872.6
				Lowest Energy	−1150.4
	VdW + Elec	0	122	Center	−235.0
				Lowest Energy	−258.0
HLA-B ~ MarVac	Balanced	0	127	Center	−887.7
				Lowest Energy	−963.5
	Electrostatic-favored	0	98	Center	−860.0
				Lowest Energy	−906.0
	Hydrophobic-favored	0	217	Center	−1054.9
				Lowest Energy	−1230.9
	VdW + Elec	0	97	Center	−260.9
				Lowest Energy	−279.6
HLA-DRB ~ MarVac	Balanced	0	96	Center	−868.3
				Lowest Energy	−988.5
	Electrostatic-favored	0	87	Center	−911.5
				Lowest Energy	−1103.6
	Hydrophobic-favored	0	104	Center	−1048.3
				Lowest Energy	−1264.6
	VdW + Elec	0	72	Center	−262.1
				Lowest Energy	−291.8
TLR-2~MarVac	Balanced	0	131	Center	−809.7
				Lowest Energy	−943.6
	Electrostatic-favored	0	90	Center	−826.1
				Lowest Energy	−964.8
	Hydrophobic-favored	0	178	Center	−1344.8
				Lowest Energy	−1443.7
	VdW + Elec	0	92	Center	−222.1
				Lowest Energy	−283.1
TLR-4~MarVac	Balanced	0	91	Center	−888.7
				Lowest Energy	−1059.4
	Electrostatic-favored	0	83	Center	−1107.8
				Lowest Energy	−1168.2
	Hydrophobic-favored	0	92	Center	−1364.8
				Lowest Energy	−1376.9
	VdW + Elec	0	55	Center	−283.3
				Lowest Energy	−287.7
TIM-1~MarVac	Balanced	0	154	Center	−733.1
				Lowest Energy	−850.2
	Electrostatic-favored	0	145	Center	−940.6
				Lowest Energy	−944.8
	Hydrophobic-favored	0	102	Center	−1125.4
				Lowest Energy	−1165.4
	VdW + Elec	0	205	Center	−169.5
				Lowest Energy	−191.7

3.14. The vaccine's codon optimization and in silico cloning

In order to produce our vaccine protein expressed in the *E. coli* K12 expression system, *in silico* cloning was performed. Prior to the cloning step, codon optimization was performed which ensures that the target protein will be efficiently translated in the expression system by reverse translating the protein for optimizing the codon and producing a well-adapted nucleotide sequence. For this purpose, Java Codon (JCat) tool produced a DNA sequence of 624 bp. The improved DNA sequence of our proposed vaccine presented with a codon adaptation index (CAI-Value) of 0.9693114367054777 and GC-content of 56.89102564102563, reflecting a good quality of vaccine. The optimized sequence of the vaccine was then injected into the pET-28a (+) vector by choosing EcoR1 and *Bam*HI as restriction sites using the SnapGene software. Finally, the computationally cloned vaccine was prepared in order to propose effective cloning strategy for the vaccine candidates. The vaccine is illustrated in Fig. 11.

3.15. Immune simulation

The C-IMMSIM immune server was used to obtain the *in silico* immune profile of the given vaccination. The immune response shown to be highly significant and long-lasting. The tertiary response by the vaccine (3rd dose) was highly efficient when compared to the primary and secondary response (shown in Fig. 12 A-N). It is observed that during the 3rd shot of vaccine, the antigenic surge was

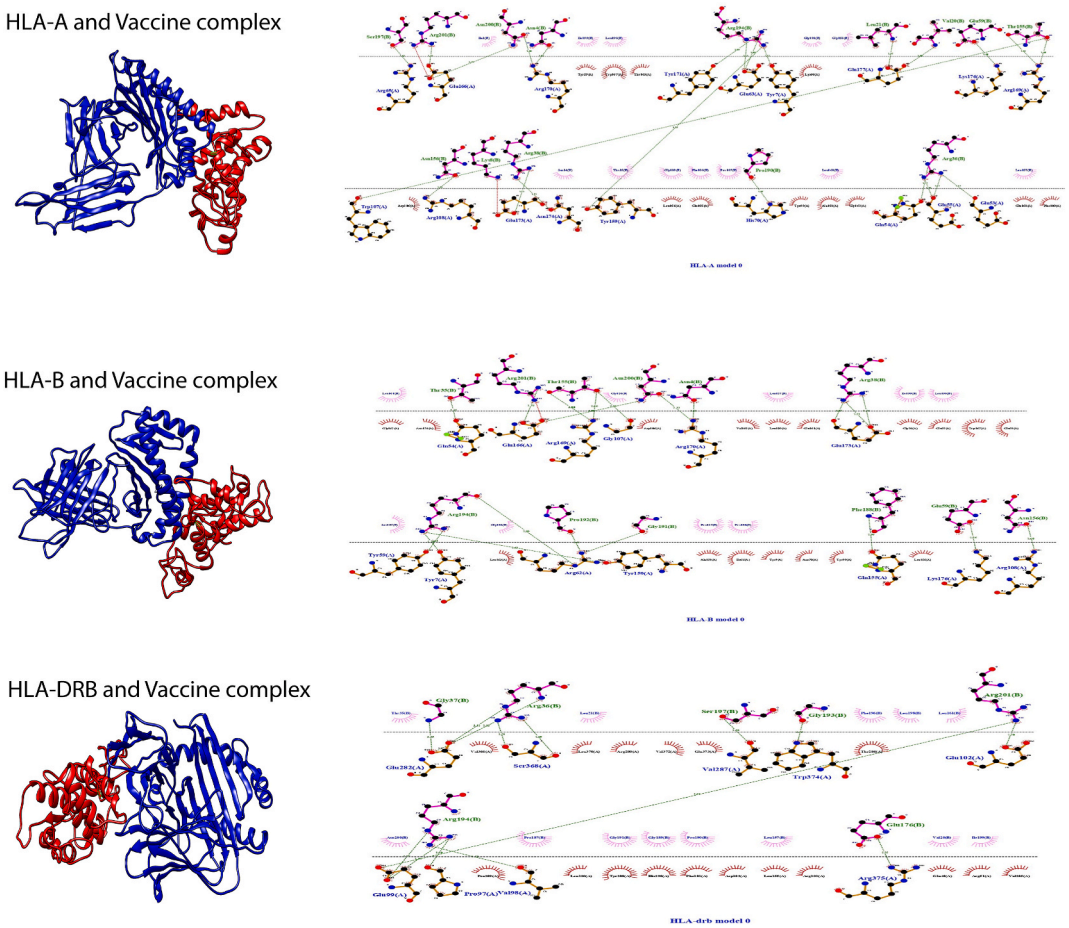


Fig. 8. Docking complexes for HLA-A (blue), HLA-B (blue), HLA-DRB (blue) and vaccine protein (red) is shown along with residue interactions. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

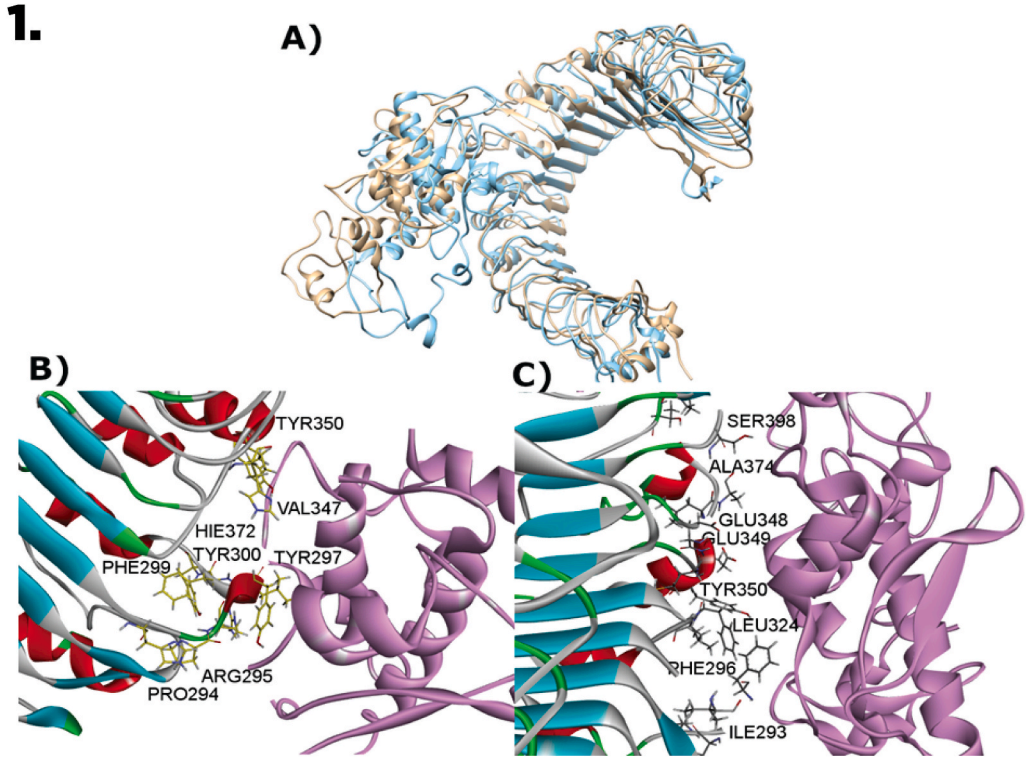
Table 4
Binding affinity of the docking complexes.

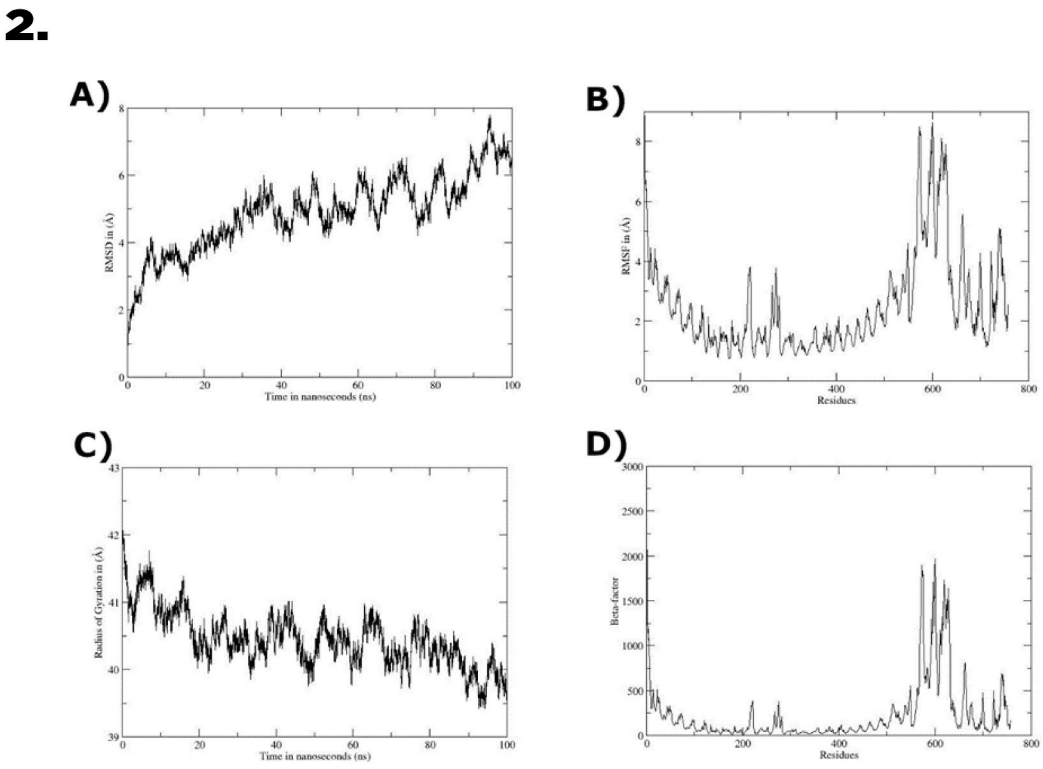
Protein-protein complex	ΔG (kcal mol^{-1})	K_d (M) at 25.0 °C
TLR-4 ¬Vaccine	−14.5	2.5E-11
TLR-2 ¬ Vaccine	−8.8	3.3E-07
HLA-A ¬ Vaccine	−13.0	2.9E-10
HLA-B ¬ Vaccine	−11.4	4.6E-09
HLA-DRB ¬ Vaccine	−9.2	1.7E-07

already plummeted and in turn the concentration of immunoglobins (IgM + IgG, IgM, IgG1+IgG2, IgG1, IgG2) indicated a significant increase. Alongside, various isotypes of the B-cells were found (isotype IgM, IgG1, IgG2) indicating the isotype switching and the memory B-cells were found to reach the maximum potential after the 3rd dose of the vaccine. In addition, our vaccine propounded the maximum number of active B-cell population and highly low concentration of anergic B-cells were indicated which reflects the immunogenic nature of our vaccine. Similarly, there was a significantly higher Th and Tc cell response right after the complete course of vaccine shots. A similar trend was also observed in the population of macrophages, natural killer cells and dendritic cells.

4. Discussion

Marburgvirus (MARV) is a WHO declared BSL-4 pathogen with no experimentally validated vaccines or therapeutics. Despite years of research and clinical trials under different phases, there has been no significant progress made to combat the infection. The virus is a potential threat throughout the African region with areas such as Uganda, Zimbabwe, the Democratic Republic of the Congo, Kenya, Angola, and South Africa being the highlighted ones due to the situated mines and caves infested by bats in these areas. Either





(caption on next page)

Fig. 9. Structural dynamics of TLR2 during 100 ns MD simulations 1A) Snapshot of overlapped structure of TLR2 at 10 ns (Brown color) and 100 ns (Blue color). **1B**) Interacting residues of TLR2 and the protein at 10 ns. **1C**) Interacting residues of TLR2 and the protein at 100 ns. **Graphs generated from 100 ns MD simulations trajectories using XMGRACE 2A).** Graph of Root mean square deviation (RMSD) of TLR2 complex for time period of 100 ns **2B).** Residual fluctuations are depicted in Root mean square fluctuations (RMSF) graph **2C**) and **2D).** Beta-factor and Radius of Gyration (Rg) graphs of 100 ns MD simulations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

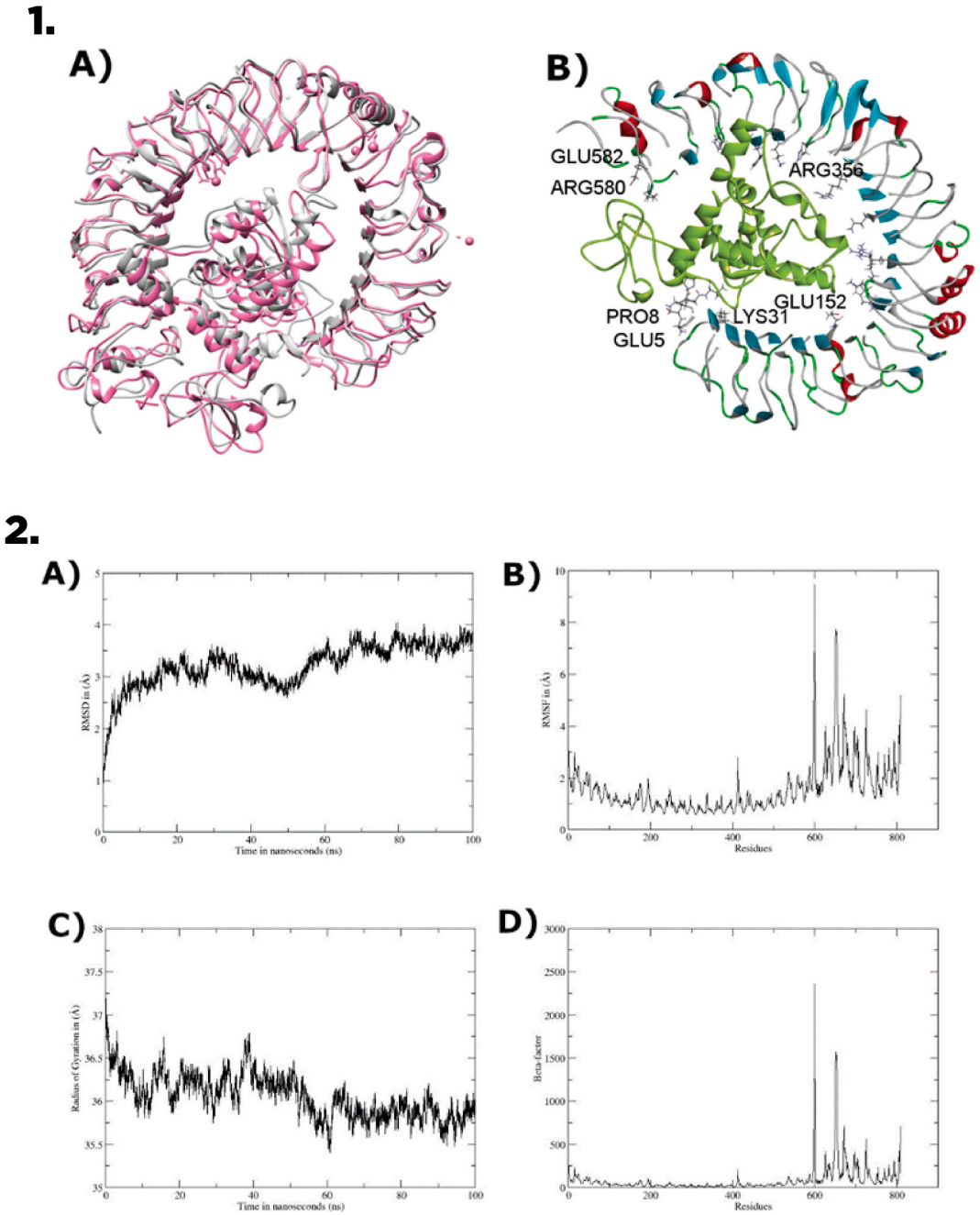


Fig. 10. Snapshots of overlapped TLR4 in complex with protein at 10 ns (Grey color) and 100 ns (Magenta color) **1B).** Interacting residues of protein and TLR4 complex throughout 100 ns. **2A).** Root mean square deviation (RMSD) graph of TLR4 complex for time period of 100 ns **2B).** Root mean square fluctuations graph (RMSF) **2C)** and **2D).** Beta-factor and Radius of Gyration (Rg) graphs of 100 ns MD simulations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 5
Binding free energy and its components in MMPBSA for TLR2 and TLR4.

Energy Components	VDW	EEL	EGB	ESURF	DELTA G gas	DELTA G solv	DELTA G TOTAL
TLR4 complex	−72.87	65.96	−63.56	−4.34	13.09	−47.90	−72.80
TLR2 complex	−13.43	−5.28	31.65	−2.87	−38.74	26.70	−9.95

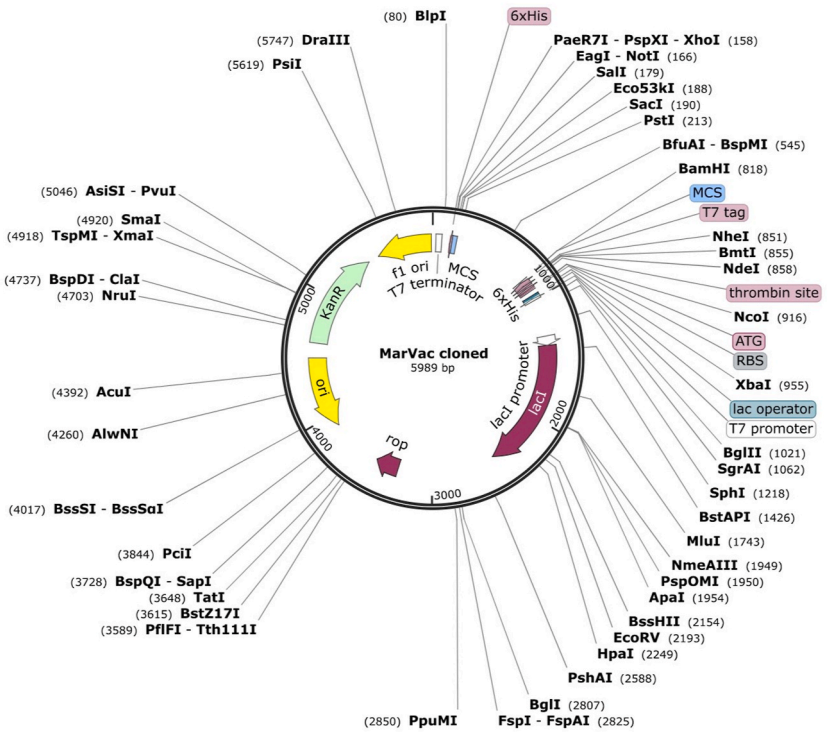


Fig. 11. Illustration of the *in silico* restriction cloning. The chimeric multi-epitope vaccine is cloned by inserting into the pET-28a (+) expression vector.

Marburgvirus disease (MVD) has been reportedly spread through the exposure to infected animals, which are the natural hosts (bats) or the spillover hosts (non-human primates). Although there are few number of reported cases yet, but there is a constant future threat of sporadic and steady outbreaks. Considering the outcomes and the severity of disease, there is always an utmost need to formulate an effective universal treatment for the infection [4]. Immunization methods comprises the major, successful and cost-effective preventive approach towards the community health and various microbial infections [5]. Since the classic vaccine types are anticipated to elicit long lasting immune responses but lead to the induction of strong allergic and autoimmune responses [6]. Thus, peptide/epitope-based vaccines offer a targeted and precise immune-protective response. Peptide based vaccines in the recent times have offered a promising alternative with highly efficacious vaccines and no allergenic or toxic reactions [7]. Although some conjoint methods have been used for peptide-based vaccine formulation of the Marburgvirus but no significant research has been done on the individual enveloped glycoprotein (env GP) of the pathogen. Considering this, we have designed a chimeric multi-epitope vaccine through computational tools with the potential to stimulate antibodies and both $CD4^+$ and $CD8^+$ T-cells. Thus, the inclusion of both T and B-cell peptides based on env GP is the novelty of our work. Although, this study is based on computational predictions, but evolution of technology has already contributed a lot in recent era of pandemic to design effective vaccines using similar approaches. Proposed vaccine constructs designed in this study have been evaluated and validated for all possible aspects and thus can be subjected to experimental validation in no time, saving much resources, time and money.

The idea of vaccine formulation rather than a combinational drug therapy comes from the consequences of antibiotic resistance. Contrary to the antibiotic resistance, antiviral resistance to the vaccine presents a rare chance, because unlike the drugs, vaccines are used for the prophylactic roles and thus coming up with broad coverage and efficacious vaccine would equip the healthcare community to combat the future challenges of the deadly infections [8].

Envelope glycoprotein (env GP), is the major culprit protein of the MARV which is responsible for the selection and entry into the target cells. Thus, targeting the env GP of the virus would be fruitful to curb the highly infectious disease [4,9]. Although, there is not much evidence on the specific receptor protein for MARV, but a recent study published in 2011, has proposed TIM-1 (T-cell immunoglobin mucin domain-1) as the potential receptor for env GP but there is little evidence to support this idea [10]. Thus, to come up

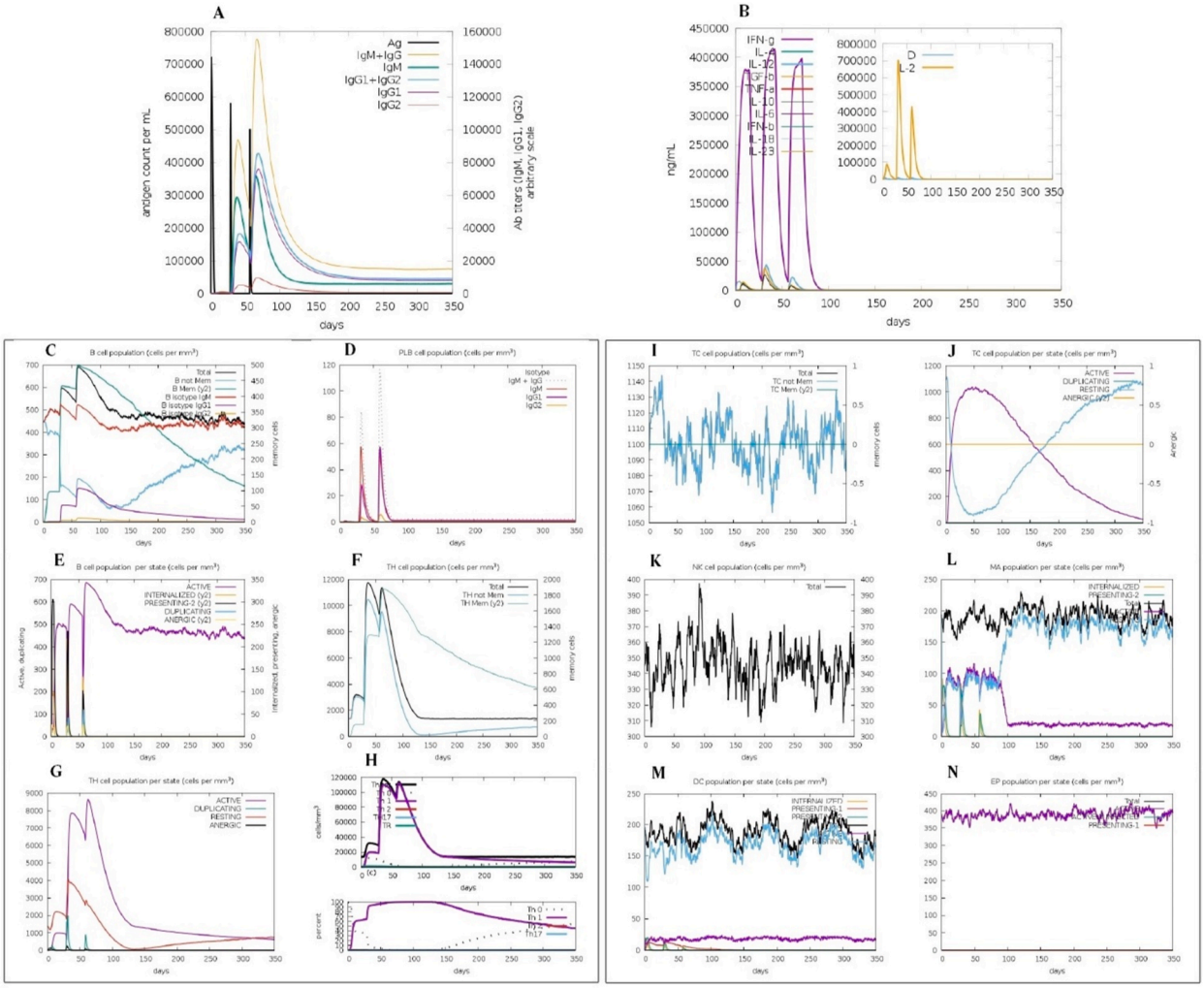


Fig. 12. A). The vaccine was injected in 3 doses (with an interval of 28 days) and it is evident from the plot that after the 3rd dose administration, the virus was almost suppressed/cleared due to the presence of protective IgGs thus showing the efficacy of the vaccination. **B).** The high peaks of indicate the presence of active interferon-gamma with a very little/declined concentration of TGF-beta and TNF-alpha. **C).** A significant rise in B-cells is recorded after the 3rd dose with a switching potential between different isotypes of the B-cell. **D). O**bvious concentration of the active B plasma cells (IgM + IgG). **G-H).** A significant response has been generated with most T-cell population in the active phase compared to the anergic phase T-cell concentration. **I–N).** Successive graphs indicate the favorable responses for other immune cell types, presenting an overall efficacy of our vaccine.

with broad-range epitope-based vaccine, rational selection was done to filter out the $CD4^+$, $CD8^+$ and B cells after navigating from a pool of env GP peptides. Selection of $CD4^+$ T cells has prioritized in the study because of their pivotal role in immune system, and being an initial mandatory factor for an effective humoral immune response. And subsequent stimulation of memory B cells and immunoglobin G. $CD4^+$ T-cells are the key cells that acquires the Th1 and Th2 phenotypes involved in the activation of the macrophages and B-cells respectively [11]. A study on West Nile virus (WNV) reported that the deficiency in $CD4^+$ T-cell and MHC-II in experimental mice models resulted in longer WNV persistence in brain and spinal cord with significantly decreasing levels of IgM and IgG [12]. Thus, the final selection of epitopes comprises of the maximum number of the $CD4^+$ T cells from both the GP1 receptor-binding domain and extracellular portion of the GP2 subunit. Although the selected GP1 subunit epitopes comprise of the receptor-binding domain but the GP2 subunit activity is crucial to the env GP. Since GP2 assists in the binding of env GP with its cytoplasmic tail involved in the enhancing the efficiency of viral entry by maintaining the structure of the ectodomain. Thus, a contrasting variety of quality epitopes have incorporated with each reflecting a significant role in the virion binding to the receptor. Our final, vaccine construct contains quality $CD4^+$ T-cell epitopes, of which epitope ^{90}KTCYNISVTDPSGKS^{104} lies in the receptor binding domain of the GP1. Since the viral infections are usually prompted by the interaction of env GP with the potential receptor, thus it plays an important role in the attachment and fusion into the target cells. Further $CD4^+$ T-cell epitopes are selected from the GP2 subunit helical and cytoplasmic domain which assists in the subsequent fusion and viral entry. To maximize the potential of our vaccine, we incorporated the 3 best

$CD8^+$ T-cell epitopes: ^{61}DSPLEASKR^{69}, ^{62}SPLEASKRW^{70}, ^{29}NPHAQGIAL^{137} with all 3 parts of the receptor binding domain of the GP1 subunit. Finally, the vaccine was modified by the addition of one 16-mer B-cell to induce the humoral immune response. The selected epitopes were highly antigenic, non-allergenic and non-toxic.

Conservancy analysis of the chosen envGP epitopes also indicated the universal candidacy of the chimeric vaccine. The world population coverage of our vaccine was shown to cover maximum population contrary to the vaccine formulation by Pervin et al. which was focused on the L-protein [13]. The adjuvant selected for our vaccine construct was beta-defensin-3 because it is an experimentally proven immune-stimulator in several wet lab studies [14–17]. Further selection of EAAAK for adjuvant attachment was due to its flexible nature and aid in vivo separation of the peptide fragments to independently act in the natural environment [18,19]. The glycine-rich GPGPG linkers were added because they are highly effective in enhancing the vaccine solubility (as evidenced by significant figure of 0.775) and allowing the epitopes to act freely [20]. Various physicochemical parameters were evaluated afterwards, which indicated that our vaccine was highly stable (with a stability index of 37.94), thermostable with an overall value of 78.37, and it was shown to be hydrophilic in nature depicting the overall reliability. Proposed vaccine construct was shown to be basic in nature with a pI of 9.44. Thus, based on this finding we constructed an energy heat map of the 3D vaccine protein which predicted that with an increasing pH and ionic strength, there would be an overall decrease in free energy.

Upon analysis of physicochemical parameters, our vaccine showed a molecular weight of 21.478 kDa with an instability index of 37.94. Generally, proteins having molecular weight of less than 110 kDa are considered stable and for the instability index, less than 40 is considered to favorable and stable [21,22]. The vaccine also displayed a half-life of more than 10 h, 20 h, 30 h in *E. coli*, yeast cells and mammalian reticulocytes respectively which indicated a desirable duration because short half-lives are the major limitation in the field of protein therapeutics [23]. Thus, the resulted half-life was quite satisfactory considering the previous studies. In terms of thermostability, our vaccine showed a value of 78.37 which indicated a quite thermostable protein. The 3D structural validation was evaluated by a Ramachandran plot drawn through PDBsum interface and our vaccine indicated to have 93.3% of the residues in most favored regions, 2.7% in additional allowed regions, 2.0% in generously allowed regions, 2.0% in disallowed regions. Our envGP based vaccine showed the highest percentage of most favored regions as compared to a study by Sami et al. [24], who employed various epitopes from multiple proteins of the virus. The ERRAT score was computed to 97% which validated the overall quality factor of our vaccine protein and a Z-score of −6.13 was determined by ProSA-web.

The envGP is the major protein involved in the receptor binding, thus it should be recognized by the toll-like receptor 2 and 4 (TLR-2 and TLR-4). Human toll-like receptor 4 is particularly expressed in various types of immune cells (macrophages, granulocytes, dendritic cells). It should be noted that besides TLR-4, TLR-2 plays an important role in eliciting an immune response. This statement is strengthened by the study published in PNAS which reported that the beta-defensin 3 is involved in activating antigen processing cells (APCs) by direct interaction with the TLR-2 [25]. Thus, selection of beta-defensin-3 as an adjuvant is crucial to our vaccine design, since it imparts effectiveness and long-lasting immune response. The interaction of our vaccine with the receptor was evaluated through molecular docking studies. The docking study of TLR4 and vaccine revealed that there was a total of 5 salt bridges, 13 hydrogen bonds and 150 non-bonded contacts. The salt bridges were formed between ARG382 and ASP178, ARG598 and GLU59, GLU603 and LYS8, GLU605 and LYS8, GLU605 and ARG38. The docking interactions of TLR-2 and vaccine protein indicated 7 hydrogen bonds and 1 salt bridge between ARG321 and ASP178. The molecular dynamics (MD) simulation was finally performed to validate our vaccine's effectiveness. MD simulation is a promising *in silico* technique to evaluate a protein's predicted response by simulating the protein under in vivo conditions. The final results of the MD simulations depicted promising results which indicated a higher confidence level for our constructed chimeric vaccine.

On these grounds, immunoinformatics and vaccinomics strategies have been utilized to design chimeric vaccines for Malaria [26], HIV [27], Zika virus [28], SARS CoV-2 [29–31], Nipah virus [32], West Nile virus [33], Tuberculosis [34], Dengue [35] and several other pathogens. Thus, in the era of advanced technology and modern vaccinology adopted methodology can be trusted to predict potential vaccine candidates. Still, it always lack experimental validation. To prove the efficiency and efficacy of the proposed vaccine construct, the study has to be extended for animal and clinical trials. Also, upon evolution of new strains, study needs to reevaluate the epitopes, their conservation pattern and population coverage. Still, along with these limitations, it is the state of the art to predict potential vaccine candidates and constructs against such stubborn viruses with great immunogenic and antigenic properties.

5. Conclusion

Marburg virus is a highly contagious virus with severe disease outcomes. After being declared a BSL-4 pathogen several years ago, no significant progress has been made to develop effective vaccines and therapeutics to combat the disease. Although a few clinical trials have been reported under study but no favorable results are achieved so far. To bridge this gap of knowledge with immunoinformatics and machine-learning, reverse vaccinology seems to be a promising method in order to develop high antigenic multi-epitope chimeric vaccines. Thus, in our study we have chosen high antigenic, non-allergic and non-toxic peptides and concatenated suitable linkers and adjuvant to design a vaccine construct. The predicted 3-dimensional structure of the vaccine was further analyzed through various immunoinformatic filters and promising results have been shown. This is the very first *in silico* study on *env*GP based vaccine for Marburg virus, has shown excellent results for the final vaccine, and thus can be subjected to further experimental validations and clinical trials.

Author contributions

AN conceived and designed the experiments. HY has performed all major computational analysis and wrote the paper. NZ and SSA

performed molecular dynamics and simulation studies. AO and FMA contributed in methodology design and manuscript writing. MH helped in generating figures and protein-protein interaction analysis. MH also contributed in analysing and interpreting data. All authors contributed in article writing and proof reading.

Funding

No funding is available for this project.

Data availability statement

All the data used from different databases have been mentioned in the manuscript. Moreover, all the generated results have been given in the tables added in the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I want to acknowledge Institute of Molecular Biology and Biotechnology (IMBB) at University of Lahore to provide us a research platform to complete this study. Also, I would like to acknowledge Computational Biology Lab, National Center for Bioinformatics, Quaid-i-Azam University, Islamabad, Pakistan for providing us with Molecular Dynamics and Simulation studies.

References

[1] A. Naz, et al., Designing multi-epitope vaccines to combat emerging coronavirus disease 2019 (COVID-19) by employing immuno-informatics approach, Front. Immunol. 11 (2020) 1663.
[2] D.M. Hoover, et al., Antimicrobial characterization of human β-defensin 3 derivatives 47 (9) (2003) 2804–2809.
[3] A. Naz, et al., Designing multi-epitope vaccines to combat emerging coronavirus disease 2019 (COVID-19) by employing immuno-informatics approach 11 (2020) 1663.
[4] K. Brauburger, et al., Forty-five years of Marburg virus research, Viruses 4 (10) (2012) 1878–1927.
[5] I. Chabot, M.M. Goetghebeur, J.-P. Grégoire, The societal value of universal childhood vaccination, Vaccine 22 (15–16) (2004) 1992–2005.
[6] M. Skwarczynski, I. Toth, Peptide-based synthetic vaccines, Chem. Sci. 7 (2) (2016) 842–854.
[7] W. Li, et al., Peptide vaccine: progress and challenges, Vaccines 2 (3) (2014) 515–536.
[8] D.A. Kennedy, A.F. Read, Why does drug resistance readily evolve but vaccine resistance does not? Proc. Biol. Sci. 284 (1851) (2017), 20162562.
[9] E.M. Janahi, et al., In silico CD4+, CD8+ T-cell and B-cell immunity associated immunogenic epitope prediction and HLA distribution analysis of Zika virus, EXCLI journal 16 (2017) 63.
[10] A.S. Kondratowicz, et al., T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus, Proc. Natl. Acad. Sci. USA 108 (20) (2011) 8426–8431.
[11] A.S. Clem, Fundamentals of vaccine immunology, J. Global Infect. Dis. 3 (1) (2011) 73.
[12] E.M. Sitati, M.S. Diamond, CD4+ T-cell responses are required for clearance of West Nile virus from the central nervous system, J. Virol. 80 (24) (2006) 12060–12069.
[13] T. Pervin, A.R. Oany, Vaccinomics approach for scheming potential epitope-based peptide vaccine by targeting l-protein of Marburg virus, In silico pharmacology 9 (1) (2021) 1–18.
[14] T. Mohan, P. Verma, D.N. Rao, Novel adjuvants & delivery vehicles for vaccines development: a road ahead, The Indian journal of medical research 138 (5) (2013) 779.
[15] T. Mohan, et al., Modulation of HIV peptide antigen specific cellular immune response by synthetic α-and β-defensin peptides, Vaccine 31 (13) (2013) 1707–1716.
[16] D. Yang, et al., Mammalian defensins in immunity: more than just microbicidal, Trends Immunol. 23 (6) (2002) 291–296.
[17] T. Mohan, D. Mitra, D. Rao, Nasal delivery of PLG microparticle encapsulated defensin peptides adjuvanted gp41 antigen confers strong and long-lasting immunoprotective response against HIV-1, Immunol. Res. 58 (1) (2014) 139–153.
[18] R.K. Pandey, S. Sundar, V.K. Prajapati, Differential expression of miRNA regulates T cell differentiation and plasticity during visceral leishmaniasis infection, Front. Microbiol. 7 (2016) 206.
[19] N. Hajighahramani, et al., Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus, Infect. Genet. Evol. 48 (2017) 83–94.
[20] M. Kavoosi, et al., Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli, Biotechnol. Bioeng. 98 (3) (2007) 599–610.
[21] A. Naz, et al., Identification of putative vaccine candidates against Helicobacter pylori exploiting exoproteome and secretome: a reverse vaccinology based approach, Infect. Genet. Evol. 32 (2015) 280–291.
[22] J.M. Walker, The Proteomics Protocols Handbook, Springer, 2005.
[23] D. Mathur, et al., PEPlife: a repository of the half-life of peptides, Sci. Rep. 6 (1) (2016) 1–7.
[24] S.A. Sami, et al., Designing of a multi-epitope vaccine against the structural proteins of Marburg virus exploiting the immunoinformatics approach, ACS Omega 6 (47) (2021) 32043–32071.
[25] N. Funderburg, et al., Human β-defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2, Proc. Natl. Acad. Sci. USA 104 (47) (2007) 18631–18635.
[26] R.K. Pandey, T.K. Bhatt, V.K. Prajapati, Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein, Sci. Rep. 8 (1) (2018) 1–11.
[27] R.K. Pandey, et al., Immunoinformatics approaches to design a novel multi-epitope subunit vaccine against HIV infection, Vaccine 36 (17) (2018) 2262–2272.
[28] A.C.B. Antonelli, et al., In silico construction of a multiepitope Zika virus vaccine using immunoinformatics tools, Sci. Rep. 12 (1) (2022) 1–20.
[29] T. Kar, et al., A candidate multi-epitope vaccine against SARS-CoV-2, Sci. Rep. 10 (1) (2020) 1–24.

[30] K. Abraham Peele, et al., Design of multi-epitope vaccine candidate against SARS-CoV-2: a in-silico study, J. Biomol. Struct. Dyn. 39 (10) (2021) 3793–3801.
[31] A. Singh, et al., Designing a multi-epitope peptide based vaccine against SARS-CoV-2, Sci. Rep. 10 (1) (2020) 1–12.
[32] R. Ojha, et al., Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus zoonotic infection, ACS Omega 4 (8) (2019) 13069–13079.
[33] M.T. Khan, et al., Immunoinformatics and molecular dynamics approaches: next generation vaccine design against West Nile virus, PLoS One 16 (6) (2021), e0253393.
[34] M. Ghandadi, An Immunoinformatic strategy to develop new Mycobacterium tuberculosis multi-epitope vaccine, Int. J. Pept. Res. Therapeut. 28 (3) (2022) 1–14.
[35] M. Ali, et al., Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection, Sci. Rep. 7 (1) (2017) 1–13.