

Immunoinformatics-guided Designing of Peptide Vaccine against Lassa Virus with Dynamic and Immune Simulation Studies

Sifat Bin Sayed¹, Zulkar Nain¹, Faruq Abdulla², Md. Shakil Ahmed Khan¹, Zahurul Haque¹, S. M. Raihan Rahman¹, Rubaia Tasmin³, Utpal Kumar Adhikari^{4*}

¹Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Islamic University, Kushtia-7003, Bangladesh. ²Department of Statistics, Faculty of Sciences, Islamic University, Kushtia-7003, Bangladesh.

³Department of Pharmacy, Faculty of Biological Sciences, Islamic University, Kushtia-7003, Bangladesh. ⁴School of Medicine, Western Sydney University, Campbelltown, NSW-2560, Australia.

*Corresponding author (u.adhikari@westernsydney.edu.au)

Abstract

Lassa virus (LASV) is responsible for a type of acute viral haemorrhagic fever referred to as Lassa fever. Lack of adequate treatment and preventive measures against LASV resulted in a high mortality rate in its endemic regions. In this study, a multi-epitope vaccine was designed using immunoinformatics as a prophylactic agent against the virus. Following a rigorous assessment, the vaccine was built using T-cell ($N_{CTL}=8$ and $N_{HTL}=6$) and B-cell ($N_{LBI}=4$) epitopes from each LASV-derived protein with suitable linkers and adjuvant. The physicochemistry, immunogenic potency and safeness of the designed vaccine (~68 kDa) were assessed. In addition, chosen CTL and HTL epitopes of our vaccine showed 97.37% worldwide population coverage. Besides, disulphide engineering also improved the stability of the chimeric vaccine. Molecular docking of our vaccine protein with toll-like receptor (TLR2) showed binding efficiency followed by dynamic simulation for stable interaction. Furthermore, higher levels of cell-mediated immunity and rapid antigen clearance were suggested by immune simulation and repeated-exposure simulation, respectively. Finally, the optimized codons were used in *in silico* cloning to ensure higher expression within *E. coli* K12 bacterium. With further assessment both *in vitro* and *in vivo*, we believe that our proposed peptide-vaccine would be potential immunogen against Lassa fever.

Keywords: Lassa fever; Immunoinformatics; Peptide vaccine; Immune simulation

Introduction

Lassa virus (LASV) is an emerging viral pathogen belongs to the *Arenaviridae* family and can cause severe viral haemorrhagic fever which is known as Lassa fever, with a 20% fatality rate [1]. Lassa fever is a viral acute zoonotic disease due to its capability to affect the highest number of people (~500,000) and causes 5,000 annual deaths in Western Africa [2]. Also, the people of Ghana (~10%), Côte d'Ivoire (~30%), Nigeria (~40%), Guinea (~50%), Sierra- Leone and Liberia (~80%) and a few areas of Mali are assumed to be affected by Lassa fever [3,4]. About 200 million people of West African regions (*i.e.*, Nigeria and Senegal) are at high risk of LASV outbreak [1]. Moreover, it also affects many areas of Europe such as the United Kingdom [5], Netherlands [6] and Germany [7]. Though it has been revealed that the virus primarily targets antigen-presenting cells (mainly dendritic cells and macrophages) and endothelial cells and interferes with their complete maturation and activation, but the pathogenesis of Lassa fever is yet not clearly

understood [8,9]. Given the high annual incidence and mortality rate, however, the development of an effective LASV vaccine is an urgent necessity. LASV is endemic to West Africa, and the genomic organization of the Lassa virus is an enveloped, ambisense and has a bisegmented, negative sense and single-stranded RNA genome consisting of large (L) segments and small (S) segments [8]. The large or L segment of the RNA encodes the 200 kDa RNA polymerase (L) protein and the small ring finger protein (matrix protein or Z protein, 11 kDa) that regulate replication and transcription [10,11]. The small segments encode the surface glycoprotein precursor (GP, 75 kDa) and the nucleoprotein (NP, 63 kDa), which is proteolytically cleaved into GP1 and GP2 (envelope glycoprotein), which bind to the alpha-dystroglycan receptor and enter mediate in the host cell [8,12].

LASV is transmitted to the human being through the rodent reservoir *Mastomys natalensis*, a typical African rat lurking in village houses [13]. Recent evidence, however, indicates that other rodent species may also be LASV recipients, like, African wood mouse *Hylomyscus pamfi* (Nigeria), and Guinea mouse *M. erythroleucus* (Nigeria and Guinea) [9]. Exchange of LASV occurs when a healthy individual comes in contact with the blood, secretion, tissue or excretion of any infected personal or by food contaminated with the host excreta. However, skin to skin contact without exchange of blood fluid cannot transmit the virus [14]. Children under ten years old are considered as the most vulnerable to LASV. For instance, a study showed 15% seropositivity in the under-aged population in West Africa [15]. Besides, pregnant patients with Lassa fever results in spontaneous abortions [16]. Ribavirin, an antiviral drug, is found to be effective at the initial phase of Lassa fever and can reduce the fatality rate [17,18].

However, the development of potential toxicity and teratogenicity when used in the later stage of disease drives us to think that Ribavirin is not a potent therapeutic against Lassa fever [19,20]. Peptide vaccines are immune stimulants where fragments of virus-derived proteins mimic natural pathogens; hence,

more influential in terms of safety, efficacy and specificity [21]. In 1985, the first epitope-based vaccine was developed using cholera toxin against *E. coli* [22]. Furthermore, peptide vaccines against many pathogenic agents (*i.e.*, HIV, malaria, swine fever, influenza, anthrax, etc.) are promptly under development [23].

Present study demonstrates the screening of whole LASV proteome followed by the grouping of viral proteins. Each protein group was evaluated separately for the identification of T-cell and B-cell epitopes along with their respective MHC alleles using vaccinomics. Subsequently, a vaccine was designed using the most persuasive epitopes from each protein with suitable adjuvant and linkers. The primary sequence was used for physicochemical analysis and immunogenic profiling, followed by the secondary and tertiary structure predictions. The predicted 3D structure was applied for refinement and validation. Besides, disulphide bridging was done to improve structural stability. The binding affinity and interactions between the vaccine protein and the receptor were calculated by molecular docking and dynamics simulation, respectively. Codon optimization and in silico cloning were taken care of for the evaluation of the expression of chimeric protein within the appropriate host. Finally, an immune simulation was performed to estimate the immunogenic potency in real-life.

Methodology

Proteome retrieval and antigenicity prediction: The whole proteome of the Lassa Mammarenavirus was retrieved from the ViPR database (Virus Pathogen Database and Analysis Resource), an integrated robust database for several virus families and their respective species [24]. Initially, the proteins from LASV proteome were isolated and classified as glycoprotein, L-protein, matrix protein, nucleocapsid protein, nucleoprotein, polymerase, ring-finger protein, Z-protein. As a measure of an immune response, the structural proteins were then applied for antigenicity prediction using the Vaxijen v2.0 server with a 0.5 threshold value [25]. This server uses auto and cross-covariance (ACC) transformation method to maintain 70-89% prediction accuracy. A protein with the best antigenic score was chosen from each class of structural proteins.

Prediction of cytotoxic t-lymphocyte (CTL) epitopes and MHC-I binding alleles: The selected protein was submitted to the NetCTL v1.2 server for the prediction of CTL epitopes (9-mer) for 12 supertypes (*i.e.*, A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62). This server predicts CTL epitopes based on three criteria, namely, MHC-I binding peptides, C-terminal cleavage, and TAP transport efficiency. MHC-I binding and C-terminal cleavage are obtained by using artificial neural

networks whereas TAP transporter efficiency is evaluated by the weight matrix [26]. In this study, the threshold was set to 0.5 which has a sensitivity and specificity of 0.89 and 0.94, respectively. Furthermore, MHC-I binding alleles for each CTL epitope were predicted using the consensus method in the IEDB analysis tool [27]. The human was selected as the source species, and percentile rank ≤ 5 was considered since lower score indicates higher affinity.

Assessment of CTL epitopes for immunogenicity, allergenicity and toxicity: Firstly, selected CTL epitopes were rechecked for antigenicity to ensure their ability to induce immune response with VaxiJen v2.0 server. Further, they were also evaluated for immunogenicity with MHC-I immunogenicity tool of IEDB server [28]. Vaccine components should be free from an allergic reaction. So, AllergenFP v.1.0 server was used for allergenicity prediction. This server can recognize both allergens and non-allergens with 88% accuracy [29]. Furthermore, toxic epitopes should be eliminated as they could compromise the functionality of the vaccine construct. Therefore, we used ToxinPred server to sort out the toxic CTL epitopes [30].

Prediction of helper T-lymphocyte (HTL) epitopes and MHC-II binding alleles: Helper T-lymphocyte (HTL) responses play an essential role in the induction of both humoral and cellular immune responses. Therefore, HTL epitopes are likely to be a significant element of preventive and immunotherapeutic vaccines. The IEDB MHC-II binding tool was applied to predict 15 amino acid long HTL epitopes using NN-align method [31]. A percentile rank was generated by comparing peptide's binding affinity with a comprehensive set of randomly selected peptides from the Swiss-Prot database. Percentile rank ≤ 5 was also considered for this analysis [32].

Identification of cytokine-inducing HTL epitopes: Innate immune system, B-lymphocytes, cytotoxic T cells and other immune cells are activated by the help of helper T-cells which further releases different types of cytokines, *i.e.*, interferon-gamma (IFN- γ), interleukin-4 (IL-4) and interleukin-10 (IL-10), HTL epitopes have the ability to overcome proinflammatory response and thus diminish tissue damage [33]. Therefore, cytokine-inducing HTL epitopes are essential in vaccine development. So, we used IFNepitope server for the prediction of IFN- γ inducing HTL epitopes using a hybrid method (Motif and SVM) along with IFN-gamma versus Non-IFN-gamma model [34]. In addition to IFN-gamma, IL-4 and IL-10 properties were also evaluated with IL4pred and IL10pred servers, respectively [35,36].

Prediction and assessment of linear B-lymphocyte (LBL) epitopes: Identification of B-cell epitopes (BCEs) is a fundamental step for epitope-based vaccine development as it produces antibodies which provide

humoral immunity. Therefore, LBL epitopes were predicted using iBCE-EL server. It uses a method that combines gradient boosting algorithms with an extremely randomized tree method [37]. The predicted LBL epitopes were then further assessed to check their antigenicity, allergenicity and toxicity with VaxiJen v2.0, AllergenFP v1.0 and ToxinPred server, respectively.

Estimation of population coverage: Different HLA alleles, as well as their expression, are sensationally distributed at various frequencies in diverse ethnicities [38]. Hence, the HLA-alleles distribution among the world population is crucial for successful multi-epitope vaccine development. In this study, the IEDB population coverage analysis tool was used for the analysis of the population coverage of the potential CTL and HTL epitopes and their MHC binding alleles [39].

Designing of multi-epitope vaccine construct: To construct multi-epitope vaccine, finally selected CTL, HTL, and LBL epitopes were linked together with the help of AAY, GPGPG and KK linkers, respectively. Peptides used for the vaccine construction are generally poorly immunogenic when used alone, therefore, requires adjuvants to boost up the immune response [23]. Consequently, the amino acid sequence of OmpA protein (GenBank: AFS89615.1) was chosen as an adjuvant and linked ahead of the first CTL epitope through EAAAK linker [40]. Using linker for joining of two epitopes is required for the effective functioning of each epitope [41].

Antigenicity, allergenicity and physicochemical evaluation: As vaccine protein should be highly antigenic, antigenicity prediction is necessary. The predicted vaccine construct was analysed using Vaxijen v2.0 [25] and cross-checked with ANTIGENpro server [42]. Vaxijen is free from any alignment and work based on various physicochemical properties of the protein, whereas ANTIGENpro server has been developed based on microarray analysis data. Screening for allergenicity is important as it indicates the potentiality of a vaccine construct to cause sensitization and allergic reaction. We used AllergenFP v1.0 [29] and AllerTOP v2.0 [43] server to check the allergenicity of the vaccine protein. Inducing an immune response after injecting the vaccine into the body is the sole purpose of vaccination. Therefore, the assessment of various physicochemical properties of the chimera protein is essential. The primary protein sequence of the vaccine was used to predict the various physicochemical features through ProtParam [44] web-server. Furthermore, the solubility of the vaccine protein upon overexpression in *E. coli* was predicted by SOLpro [45] tool in the SCRATCH suite.

Secondary and tertiary structure prediction of the vaccine construct: PSIPRED v4.0 server was used for

the prediction of vaccine's secondary structure. Two feed-forward neural networks are the basis of PSIPRED, which process the PSI-BLAST (Position-Specific Iterated-Blast) to predict a secondary structure from an amino acid sequence as an input [46]. The tertiary structure of the final subunit vaccine was predicted using the RaptorX server, which is based on three steps, specifically single-template threading, multiple-template threading and alignment quality prediction. The server predicts 3D protein model and provides some confidence scores to evaluate the quality of predicted models. *p*-value, GDT, uGDT, modelling error at each residue etc. are different confidence scores that provide a clear indication to the relative good model [47].

Refinement and validation of 3D vaccine construct: The predicted model of the vaccine protein was refined through GalaxyRefine web server. This server initially rebuilds side chains, then side chain repacking and consequently uses the molecular dynamic simulation for overall structural relaxation [48]. Structural validation is a process to identify potential errors in the predicted tertiary structure [49]. Therefore, ProSA-web was used for structural validation, which provides an overall quality score for the input structure. If the calculated score falls outside the range characteristics of the native protein, then the structure contains errors [50]. The ERRAT server was also used to analyse non-bonded atom-atom interactions [51]. Finally, Ramachandran plot was obtained using PROCHECK server. The Ramachandran plot is a way to visualize energetically allowed and disallowed dihedral angles psi (ψ) and phi (ϕ) of amino acid and is calculated based on van der Waal radius of the side chain. The result from PROCHECK include the percentage and number of residues in most favoured, additional allowed, generously allowed, and disallowed region, which defines the quality of modelled structure [52].

Conformational B-cell epitope prediction: The conformational B-cell epitope is the collection of amino acid residues on the 3-dimensional geometry of the vaccine protein which interacts directly to the immune receptor. Therefore, ElliPro tool of IEDB server was used to determine the presence of conformational B-cell epitopes in the validated tertiary structure. ElliPro uses three algorithms relying on their protrusion index (PI) values to approximate the protein structure as an ellipsoid, calculate the residue PI, and adjacent cluster residues. ElliPro provides an average PI value over each epitope residue for each generated epitope. For each epitope residue, the PI value is calculated based on the residue mass centre outside the largest possible ellipsoid [53].

Disulphide engineering of final vaccine construct: Disulphide bonds are covalent interactions that stabilizes molecular interactions and provide considerable stability by confirming precise geometric

conformations. Disulphide engineering is a novel approach for creating disulphide bonds into the target protein structure. Therefore, disulphide engineering was executed with Disulphide by Design v2.12 web tool. Initially, the refined protein model was uploaded and run for the residue-pair search that can be used for the disulphide engineering purpose. Potential residue pairs were selected for mutation, and cysteine residue was used as a final target for disulphide engineering [54].

In silico codon adaptation and cloning: Codon optimization is essential as unadapted codon may lead to the minor expression rate in the host. Codon optimization was performed using Java Codon Adaptation Tool (JCat) server with a view to improving the translational efficiency in *E. coli* K12 [55]. Three additional options were selected to avoid the rho-independent transcription termination, prokaryote ribosome binding site, and restriction enzymes cleavage sites. Codon adaptation index (CAI) value and GC content of the adapted sequence was obtained and compared with the ideal range [56]. Consequently, the received nucleotide sequence was cloned into the *E. coli* pET28a(+) vector by using SnapGene v4.2 tool.

Molecular docking between the vaccine and TLR2 receptor: Molecular docking is a computational method which involves the interaction between a ligand molecule and the receptor molecule to provide a stable adduct. Also, a calculated score was provided as a measure of the degree of binding interaction [57]. Toll-like receptor 2 (TLR2) can mediate high proinflammatory responses against LASV infection [58]. Therefore, TLR2 structure was used as the receptor (PDB ID: 3A7B) and the refined vaccine protein as the ligand [59]. Finally, the binding affinity between the multi-epitope vaccine and TLR2 receptor was calculated through the ClusPro v2.0 server [60]. This server completed the task in three consecutive steps such as rigid body docking, clustering of lowest energy structure, and structural refinement by energy minimization. The best-docked complex was selected based on the lowest energy scoring and docking efficiency.

Molecular dynamics simulation: Molecular dynamics study is critically essential for checking the stability of the protein-protein complex in any in silico analysis. Protein stability can be determined by comparing essential protein dynamics to their normal modes [61,62]. The iMODS server was used to explain the collective protein motion in the internal coordinates through normal mode analysis (NMA) [63]. The server estimated the direction and extent of the immanent motions of the complex in terms of deformability, eigenvalues, B-factors, and covariance. The deformability of the main chain depends on whether a specified molecule can deform at each of its residues. The eigenvalue of each normal mode describes the

rigidity of motion. This value is related directly to the energy required for the structural deformation, and the deformation is much easier if the eigenvalue is low.

In silico evaluation of immune response: To estimate the immunogenic potential of the final vaccine, in silico immune simulations were conducted using the C-ImmSim server. This immune simulator uses a position-specific scoring matrix (PSSM) and machine learning techniques for the prediction of epitope prediction and immune interactions, respectively [64]. The minimum suggested interval between dose 1 and dose 2 is four weeks, according to most vaccines in current use [65]. All parameters were set at default with time steps set at 1, 84, and 168 where each time step is equal to 8 hours. Therefore, three injections were given four weeks apart. Moreover, six doses of injections of the designed vaccine were given in the same manner to simulate repeated exposure interaction to the antigen seen in a typical endemic area to probe for clonal selection.

Results

Protein retrieval and highest antigenic protein selection: For the construction of candidate vaccine, a total of 1665 protein sequences of different structural (glycoprotein, matrix protein, nucleocapsid protein, nucleoprotein, ring finger protein and Z protein) and non-structural (L protein, Polymerase, RNA directed RNA-polymerase) protein sequences of the LASV were retrieved from ViPR database. Structural protein enables viruses to invade and assemble viral particles in the host, while non-structural proteins secrete various enzymes that assist in viral replication and development of structural proteins. Vaxijen v2.0 server revealed the highest antigenic protein from each type and as we set a threshold of 0.5, only glycoprotein, matrix protein, ring finger protein and Z protein showed antigenicity (Table 1).

Table 1: The most antigenic protein of each type, along with their GenBank accession ID, antigenicity score, and length.

Proteins Name	GenBank Accession	Protein Length	Antigenic Score
Z protein	AAT48997	99	0.7645
Matrix protein	ANH09745	99	0.7368
Glycoprotein	AAT49014	254	0.6956
Ring finger	AAO59514	99	0.5849

Prediction and appraisal of CTL, HTL and LBL epitopes: A total of 180 unique CTL epitopes (9-mer) were predicted from the four LASV highest antigenic proteins. Here 97, 22, 32 and 29 CTL epitopes were predicted from the glycoprotein, matrix protein, ring finger protein and Z protein, respectively, using the NetCTL v1.2 server. Among them, only 42 epitopes were found as antigenic, immunogenic, and non-toxic (Supplementary Table 1). Of the 42 epitopes, 22 were

found to be non-allergenic. These non-allergenic epitopes were further used to predict their MHC-I binding alleles using MHC-I allele prediction tool of the IEDB server. Similarly, a total of 72 unique HTL epitopes (15-mer) and their MHC-II binding molecules were predicted using the IEDB MHC-II prediction tool. The cytokine (*i.e.*, IFN- γ , IL-4 and IL-10) inducing ability of these HTL epitopes were also evaluated (Supplementary Table 2). B-cell epitopes are antigenic regions of a protein that can trigger antibody formation. The iBCE-EL tool was used to predict Linear B lymphocyte (LBL) from the LASV proteins.

cytokines were selected in case of other proteins (Table 3). As we get a numerous number of B-cell epitopes with higher antigenicity, non-toxicity and non-allergenicity, we took the epitope with the best probability score (obtained from iBCE-EL tool) from each type of protein (Table 4). Therefore, 6 CTL, 8 HTL and 4 LBL epitopes (Table 1) are merged by AAY, GPGPG and KK linkers, respectively (Figure 1). OmpA agonist (GenBank ID: AFS89615.1), which is 352 amino acid residues long, was used as an adjuvant for TLR2 receptor using EAAAK linker. The final vaccine construct comprises 642 amino acid residues.

Table 2: Finally selected cytotoxic T lymphocyte (CTL) epitopes for multi-epitope vaccine construction.

Protein	CTL epitopes	Position	Combined	Antigenicity	Immunogenicity	Allergenicity	Toxicity
Glycoprotein	RPSPIGYLG	234-242	0.6122	2.3297	0.08704	No	No
	CGIIGLVAF	41-49	0.5494	1.4004	0.25738	No	No
	ATCGIIGLV	39-47	0.7889	1.3749	0.2936	No	No
Matrix	APPARTGEN	83-91	0.527	1.1219	0.20981	No	No
Ring finger	FENRGLVEC	36-44	0.7557	1.228	0.16588	No	No
Z protein	GDRTRPPPY	89-97	0.6183	1.4294	0.07998	No	No

We found a total of 101 LBL epitopes and after evaluation, only 37 unique epitopes (14, 5, 9 and 9 LBL epitopes were found from the glycoprotein, matrix protein, ring finger protein and z protein, respectively) were found to be non-allergenic and non-toxic and considered for vaccine construction (Supplementary Table 3).

Construction of multi-epitope vaccine: For multi-epitope vaccine designing, we have considered highly antigenic CTL epitopes from every type of proteins that are immunogenic, non-allergenic and non-toxic (Table 2). On the other hand, at HTL epitope selection, we screened cytokine-inducing properties and found that only two epitopes from the glycoprotein have the capacity to induce all three type of cytokine, while epitopes from the other proteins were positive for a maximum of two cytokines.

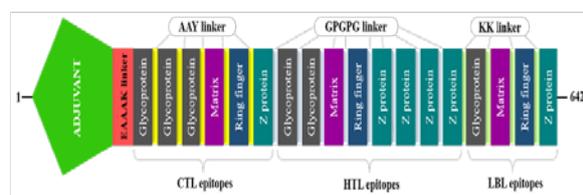


Figure 1: Graphical presentation of the multi-epitope vaccine construct. A 642 amino acid residues long vaccine construct consisting of an adjuvant (green) at the N-terminal end is linked with the whole multi-epitope sequence through EAAAK linker (red). CTL, HTL and BL epitopes are fused with the help of AAY (yellow), GPGPG (light blue) and KK (light green) linkers, respectively.

Hence, we selected two epitopes from the glycoprotein and epitopes having inducing feature for at least two

Population coverage analysis: The distribution of HLA allele varies between different geographical and ethnic regions around the globe. Therefore, population coverage during the development of an efficient vaccine must be taken into account. In this study, selected CTL and HTL epitopes, which were used to construct the vaccine and their corresponding HLA alleles (Supplementary Table 4 and 5) were obtained for population coverage analysis both individually and in combination.

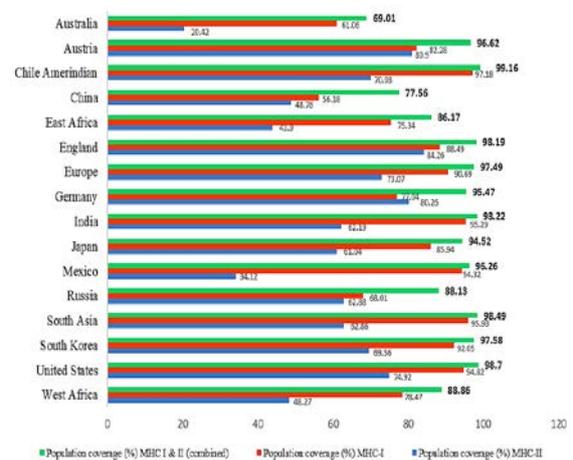


Figure 2: Population coverage of the selected T-cell epitopes and their respective HLA alleles. Regions of particular importance were considered in this graph with their MHC-I (red), MHC-II (blue) and combined (green) coverage rate.

Our selected CTL and HTL epitopes were found to cover 91.74% and 68.15% of the world population, respectively. Importantly, CTL and HTL epitopes showed 97.37% population coverage worldwide when

used in combination. The highest population coverage was found to be 99.16% in the South American country Chile Amerindian. In West Africa, where the bacteria first appeared and had several outbreaks, the population coverage was 88.86%. LASV caused several epidemics in different countries of the world, especially in England, Germany, and the Netherlands where the population coverage was found to be 98.19%, 95.47%, and 69.80%, respectively (Supplementary Table 6). In addition, the population coverage in China, Europe, France, India, Japan, South Korea, Russia, South Asia, and United States were found to be 77.56%, 97.49%, 94.65%, 98.22%, 94.52%, 97.58%, 88.13%, 98.49% and 98.70%, respectively (Figure 2).

Table 3: Finally selected helper T lymphocyte (HTL) epitopes with their cytokine inducing properties

Protein	HTL epitopes	Position	Antigenicity	IL4pred	IL10pred	IFNepitope
Glycoprotein	GYLGLLSQRTRDIYI	239-253	1.8077	IL4 inducer	IL10 inducer	Positive
	GVLQTFMRMAWGGRY	185-199	0.4769	IL4 inducer	IL10 inducer	Positive
Matrix	NCLTLLLSVSSRCPI	52-66	0.6165	IL4 inducer	IL10 inducer	Negative
Ring finger	NCLSLLLSVSSRCPI	52-66	0.7793	IL4 inducer	IL10 inducer	Negative
Z protein	LTLILLSVSNRPCICK	54-68	0.2654	IL4 inducer	IL10 inducer	Negative
	NCLTLLLSVSNRPCI	52-66	0.7145	IL4 inducer	IL10 inducer	Negative
	CLTLLLSVSNRPCIC	53-67	0.8906	IL4 inducer	IL10 inducer	Negative
	CLNCLTLLLSVSNRC	50-64	0.9551	Non-inducer	IL10 inducer	Positive

Table 4: Finally selected linear B lymphocyte (LBL) epitopes for multi-epitope vaccine construction

Protein	Peptide Sequence	Position	Probability	Allergenicity	Antigenicity	Toxicity
Glycoprotein	GKSCSLTLKGGY	54-65	0.8007	Non-allergen	Antigen	Non-Toxic
Matrix	KQTKAPEVEDGP	4-15	0.7259	Non-allergen	Antigen	Non-Toxic
Ring finger	WFENRGLVECEN	35-46	0.7579	Non-allergen	Antigen	Non-Toxic
Z protein	APETKNSPRASL	8-19	0.7927	Non-allergen	Antigen	Non-Toxic

Physicochemical analysis and solubility prediction:

Multiple physicochemical properties were calculated from the ProtParam server by inserting the whole vaccine construct as an amino acid sequence. The molecular weight of the construct was calculated as ~68 kDa and the antigenicity prediction showed that the construct has good antigenic properties. The analysis shows 9.31 pI (Isoelectric point) value which indicates the vaccine construct is basic in nature. The instability index (II) was computed to be 29.05, which implies that the sequence of the construct will remain stable after expression. The aliphatic index was calculated as 84.98 which indicates the construct's thermostability.

The grand average of hydropathicity (GRAVY) was calculated to be negative (-0.118). This negative value indicates the hydrophilic nature of protein; therefore this protein tends to have better interaction with other proteins. Estimated half-life in mammalian reticulocyte (*in vitro*) was found to be 30 hours, while in yeast and *Escherichia coli* the estimated half-life (*in vivo*) are >20 hours and >10 hours, respectively. The immunogenic appraisal also revealed that our vaccine construct is highly antigenic, non-allergenic and showed higher solubility rate calculated through

SOLpro server. These evaluations suggested that our vaccine construct might be an ideal vaccine against LASV (Table 5).

Secondary structural feature prediction: PSIPRED v4.0 workbench was used to predict the secondary structure of the final vaccine construct. The final vaccine construct (642 amino acid long) was analysed in which 352 amino acids involved in random coil formation while 149 amino acids involved in α -helix creation and β -strands are formed only by 141 amino acids. So, secondary structural feature prediction results indicate 54.83% are random coils, 23.21% forms α -helix and 21.96% are β -strands (Figure 3).

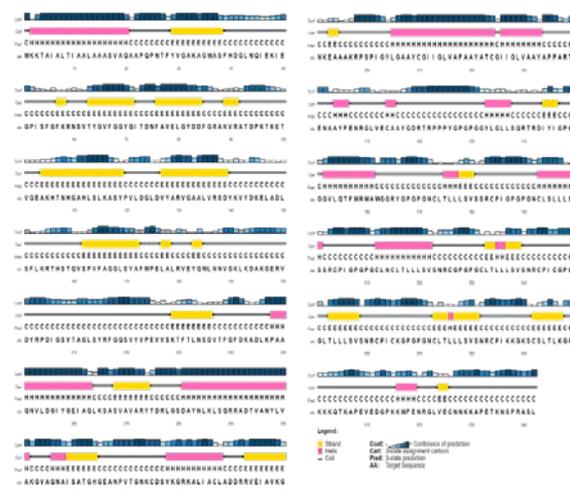


Figure 3. Graphical representation of secondary structure prediction of the multi-epitope vaccine. Here, the β -strands, α -helix and random coils are indicated by yellow, pink and blue colour, respectively.

Tertiary structure modelling and refinement: The tertiary structure of the multi-epitope vaccine was

modelled from the RaptorX server. In RaptorX server, 100% (642) amino acid residues were modelled as six domains. The protein structure (PDB ID: 1bxwA) was used as the best template for modelling. The relative quality of the structure was evaluated by *p*-value. Calculated *p*-value for the modelled structure was 5.39×10^{-06} which is very low. Here, lower the *p*-value higher the quality of the model.

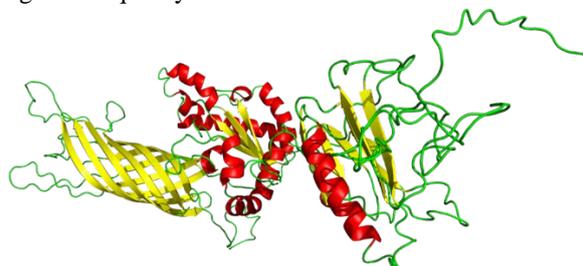


Figure 4: Tertiary structure of the multi-epitope vaccine after refinement indicating α -helix (red), β -strand (yellow) and random coil (green) in colour.

Furthermore, uGDT is also used as a parameter for 3D structure evaluation and a construct with >100 residues, $uGDT > 50$ is a good indicator. As we got 354 as uGDT score, it indicates the tertiary model as acceptable protein model for further analysis (Supplementary Table 7). The predicted tertiary (3D) structure was further refined using GalaxyRefine server, leads to an increase in the number of residues in the favoured region, generated five refined models.

Table 5: Antigenic, allergenic and physiochemical assessments of the final vaccine protein

Features	Assessment
Number of amino acids	642
Molecular weight	67909.01 Dalton
Chemical formula	$C_{3018}H_{4809}N_{845}O_{886}S_{25}$
Theoretical pI	9.31
(+) Residues: Asp + Glu	49
(-) Residues: Arg + Lys	76
Total number of atoms	9583
EC_{280nm} (in H_2O)	$64970 M^{-1}cm^{-1}$
Instability index	29.05
Aliphatic index	84.98
GRAVY	-0.118
Antigenicity	0.7223
Allergenicity	Non-allergen

The refined best model showed 93.6% residues in the most favoured region in the Ramachandran plot, GDT-HA score 0.9276, RMSD 0.490, MolProbity 2.052, Clash score 13.2 and Poor rotamers 0.8 (Supplementary Table 8), which indicate the quality of the refined model among all the five models after comparison and finally helped us to select the model for further studies (Figure 4).

Validation of refined tertiary structure: Validation of the refined tertiary structure was checked by using PROCHECK, ProSA-Web and ERRAT server.

Ramachandran plot analysis of the crude structure, by PROCHECK server, revealed that 85.9% of the structure was located in the most favoured region, 11.1% in additional allowed regions, 2.0% in generously allowed regions and 1.0% of the residues were the in disallowed regions. Whereas, after refinement PROCHECK generated a better result, 89.7% of residues were located in the most favoured regions, 8.6% in additionally allowed, 1.0% in generously allowed and 0.8% of residues were found in the disallowed region (Figure 5A).

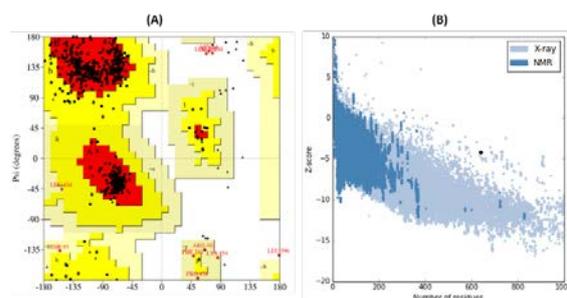


Figure 5: Structural validation of the tertiary structure of the vaccine construct. (A) Represents the Ramachandran plot of the refined model where most favoured, allowed and disallowed regions are 89.7%, 9.6% and 0.8% respectively. (B) Indicates ProSA-web validation of 3D structure showing Z-score (-4.23).

ProSA-web and ERRAT verified the quality and potential errors in a crude 3D model. The selected best model after refinement had an overall quality factor of 78.2% with ERRAT while ProSA-web gave a Z-score of -4.23 for the input vaccine protein model, indicating the model is slightly in the range of native protein conformation (Figure 5B).

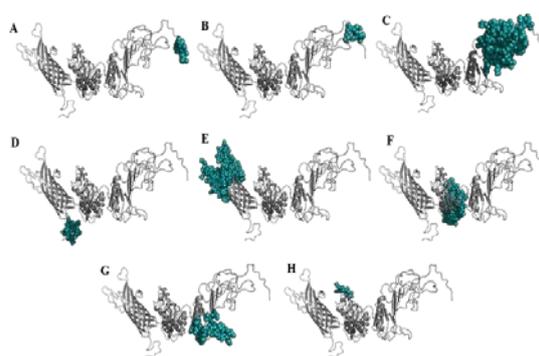


Figure 6: Graphical presentation of eight conformational B-cell epitopes of the peptide vaccine. Here the predicted epitopes are indicated by cyan colour (spheres) and the rest of the residues are in grey (cartoon).

Conformational B-cell epitope prediction: Conformational B-cell epitopes were predicted using ElliPro, an online web server that predicts epitope

based on the tertiary structure. A total of 305 residues with scores varying from 0.566 to 0.944 were predicted to be located in eight conformational B-cell epitopes (Figure 6). The epitopes ranged in size from 4 to 93 residues.

Disulphide bridging for vaccine stability: Disulphide engineering was performed using Disulfide by Design v2.12 to stabilize the modelled structure of the final vaccine construct. In total, 54 pairs of residues could be used in disulphide engineering have been discovered (Supplementary Table 9). However, only two pairs of residuals have been concluded after the evaluation of other parameters such as energy score and χ^3 angle, as their value falls below the allowed range, i.e. energy should be less than 2.2 kcal/mol, and χ^3 angles are expected to be between -87 and $+97$ degrees [54].

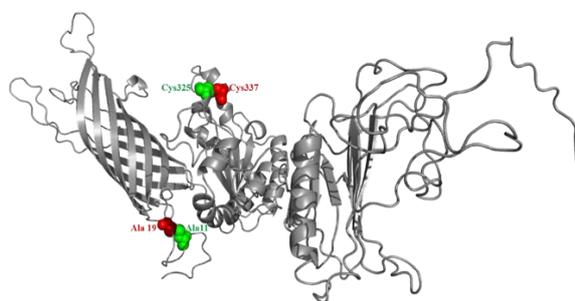


Figure 7: Disulphide engineering of the vaccine construct to improve stability. Two mutated pairs are shown in green and red colour which are selected based on their energy, χ^3 value, and B-factor.

Therefore, a total of four mutations were generated on the residue pairs. For Ala11-Ala19 residual pairs, the energy score is 0.98 kcal/mol, and the χ^3 angle is 85.64 degree. Whereas, for Cys325-Cys337, the χ^3 angle and the energy were -74.33 degree and 1.93 kcal/mol (Figure 7).

Codon optimization and *in silico* cloning: Expressing the LASV-derived vaccine protein epitope into the *E. coli* expression system was the primary purpose of *in silico* cloning. Therefore, according to the codon usage of the *E. coli* expression system, it was necessary to adapt the codon respectively to the subunit vaccine construct. To optimize codon usage of the vaccine construct in *E. coli* K12 for maximal protein expression, the Java Codon Adaptation Tool (JCat) was used. The length of the optimized codon sequence was 1,926 nucleotides. Codon optimization evaluates the sequence and tells about GC content of the cDNA sequence and codon adaptive index (CAI) where GC content was calculated as 53.63% which lies in the optimum range of (30-70) %. CAI was calculated as 0.98, which also lies in the range (0.8-1.0), which indicates the possibility of good expression of the vaccine candidate in the *E. coli* host. XhoI and NdeI

restriction sites were later created and cloned using SnapGene software into the pET28a (+) vector (Figure 8). Thus, the total length of the clone was 7.22 kbp.

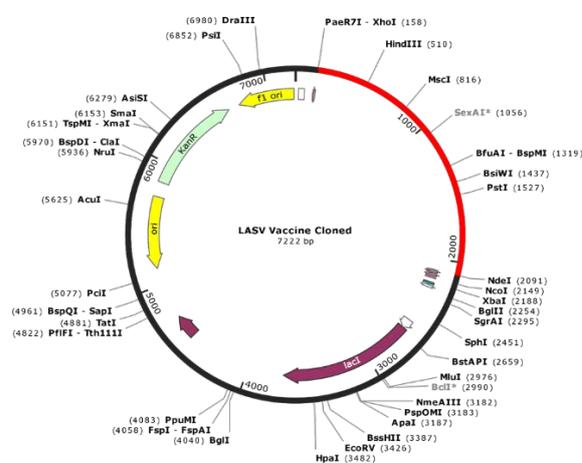


Figure 8: *In silico* cloning of the final vaccine construct into pET28a(+) expression vector where the red part indicates the coding gene for the vaccine surrounded between XhoI (158) and NdeI (2091) while the vector backbone has shown in a black circle.

Molecular docking of vaccine with immune receptor: To assess the interaction between the refined model and the TLR 2 (PDB ID-3A7B) immune receptor, molecular docking was performed by using online server ClusPro 2.0 and a total of 30 models were generated (Supplementary Table 10). Among them, only that model was selected, which occupied the receptor properly and having the lowest energy score. Since model number 1 fulfils the desired criteria, was chosen as the best-docked complex (Figure 9). The energy score obtained for the model 1 was found to be -1406 , which is lowest among all other predicted docked complex confirming the highest binding affinity.

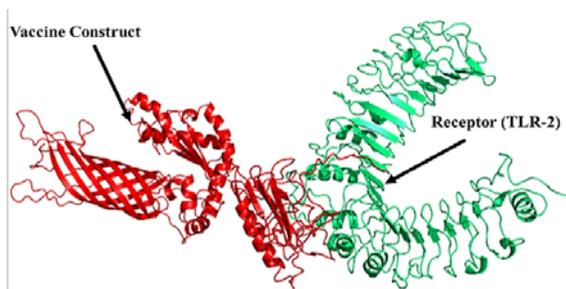


Figure 9: The docked complex of vaccine protein and TLR-2 (PDB ID- 3A7B) receptor. Here the receptor is represented in green colour whereas the vaccine protein is in red.

Molecular dynamics simulation of the vaccine-TLR2 complex: Normal mode (NMA) analysis was conducted to scrutinize protein stabilization and their

large-scale mobility. This assessment was conducted by iMODS server depending on the internal coordinates of the docked complex. The complex's deformability depends on the individual distortion of each residue, depicted by chain hinges (Figure 11B). The eigenvalue found for the complex was $9.857553e-08$ (Figure 11A). The variance correlated with each normal mode was inverted to the eigenvalue [66]. The B-factor values generated from normal mode analysis were proportional to RMS (Figure 11C). Covariance matrix showing various pairs of related, anti-correlated or uncorrelated motions represented by red, blue and white colours, has stated the coupling of pairs of residues, respectively (Figure 11D). The result also provided an elastic model of the network that distinguished the pairs of atoms linked through springs (Figure 11E). Each dot in the diagram showing one spring, coloured by the degree of stiffness, between the corresponding atom pairs. The darker the greys, the more rigid the springs were.

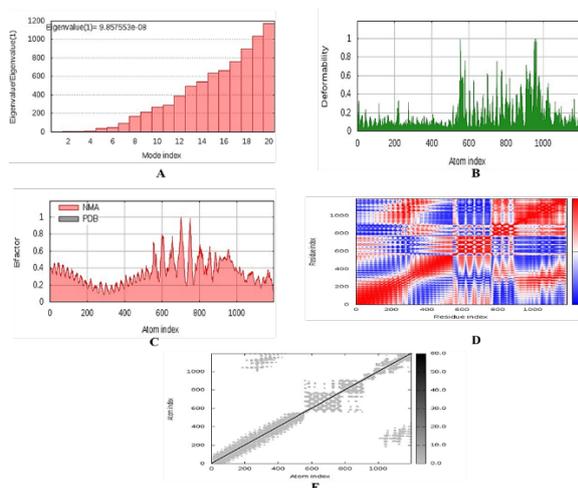


Figure 10: Molecular dynamics simulation of the vaccine-TLR2 complex, showing (A) eigenvalue; (B) deformability; (C) B-factor; (D) covariance matrix; and (E) elastic network analysis.

In silico immune simulation: The simulated immune response was compatible with actual immune responses (Figure 11). For instance, the secondary and tertiary responses were higher than the primary response. High concentrations of IgM was characterized at the primary response. In both the secondary and tertiary reactions, the typical high levels of immunoglobulin activities (i.e., IgG1 + IgG2, IgM, and IgG + IgM antibodies) were evident with concomitant antigen reduction (Figure 11A). This indicates the emergence of immune memory and thus increased antigen clearance upon subsequent exposures (Figure 11E). Furthermore, several long-lasting B-cell isotypes were observed, suggesting the potential for isotype switching and memory formation (Figure 11B-11C). In the TH (helper) and TC (cytotoxic) cell populations with the respective

memory development, a similarly elevated response was noticed (Figure 11D-11F).

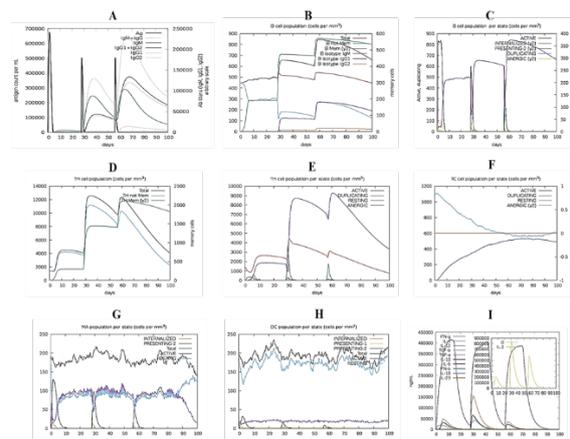


Figure 11: C-ImmSim presentation of an in silico simulation of immune response using the chimeric peptide as antigen, showing (A) Immunoglobulin production in response to antigen injection, (B) B cell population after three injections, (C) B cell population per state, (D) The evolution of T-helper cell, (E) T-helper cell population per state, (F) Cytotoxic T-cell population per state, (G) Macrophages population per state, (H) Dendritic cell population per state, and (I) Production of cytokine and interleukins with Simpson index (D).

During exposure, increased macrophage activity was demonstrated, with continuously proliferating dendritic cells (Figure 11G-11H). High levels of IFN- γ and IL-2 were also evident. Besides, a lower Simpson index (D) indicates greater diversity (Figure 11I). Moreover, multiple exposure ($n=6$) simulation as an encounter to endemic regions led to an increase of concentrations of IgG1 and decreasing IgM and IgG2 levels, while maintaining a high concentration of IFN- γ , IL-2, TC, and TH cell populations. (Supplementary Figure 1). This profile suggests immune memory development and, therefore, natural immune protection against the virus in question.

Discussion

LASV is a virus with a higher mortality rate and has the potential to bring upon catastrophe among the endemic region like West Africa. So, for developing a prevention method to fight against LASV is an obligation. Nowadays, vaccination is the most dynamic approach to improve the immunity system to fight against infectious diseases. Efficient development and manufacturing of live or attenuated vaccine, however, is expensive and can take years to complete. Nonetheless, the incorporation of excessive antigenic load in the attenuated vaccine appears not only to contribute little to the protective immune response but to complicate the state by causing allergic reactions [23]. Compared to traditional vaccines, multi-epitope vaccines decrease unwanted parts, which can either

cause pathological immune responses or adverse effects [67]. Increased safety, cost-effectiveness, the opportunity to rationally engineer the epitopes for increased potency and breadth, and the ability to focus immune responses on conserved epitopes also include potential benefits of epitope-based vaccines [68]. For many years, researchers have sought to minimize the cost, time and side-effects of vaccine development. Different strategies are readily available at this moment for designing and developing efficient and competent new generation epitope-based vaccines depending on immunoinformatic approaches [69,70]. Researchers also used immunoinformatics methods as a tool to provide futuristic models of multi-epitope driven vaccine against Ebola virus, Hepatitis C virus, Oropouche virus, Dengue virus, etc [71–74]. Knowing all the pros of multi-epitope based vaccine, our first and foremost concern was to construct a vaccine which will be able to elicit a robust immune response after vaccination. Though there had been a few attempts to suggest candidates for peptide vaccine against LASV, this is the very first approach recommending a fully functional multi-peptide based vaccine that has been evaluated by in silico approaches [75–77].

ViPR database was used to retrieve the whole complete sequence of LASV and after screening different types of protein, 4 protein sequences were selected due to their higher antigenicity. Through different servers and databases CTL, HTL and LBL epitopes were chosen as a vaccine candidate. An effective multi-epitope vaccine should be designed to include epitopes capable of producing CTL, HTL and B cells epitopes and inducing efficient reactions to a specific tumour or virus [67]. Because of its function in inducing antibody manufacturing and mediating its effective features, we have been interested in incorporating B cell epitopes [78]. Over time, however, the humoral response from memory B cells can easily be overcome by the emergence of antigens, whereas cell-mediated immunity (T-cell immunity) often leads to lifelong immunity [79]. CTL limits pathogen spread by identifying and destroying infected cells and by secreting unique antiviral cytokines [80]. Therefore, B and T-cell epitopes were predicted for the vaccine construct.

The vaccine candidates were selected from HTL, CTL and B-cell epitopes based on their antigenicity, allergenicity, immunogenicity and toxicity. Helper T-cells that release other types of cytokines such as interferon-gamma (IFN- γ), interleukin-4 (IL-4) and interleukin-10 (IL-10) have the potential to overcome pro-inflammatory response and therefore reduce tissue damage. The innate immune system, B-lymphocytes, cytotoxic T cells and other immune cells are activated with the help of helper T-cells. Thus, the cytokine (i.e., IFN- γ , IL-4 and IL-10) inducing ability of specific HTL epitopes were also evaluated for candidate choosing. The vaccine construction was completed

after joining the CTL, HTL and B-cell epitopes with AAY, GPGPG and KK linkers, respectively. To enhance expression, folding and stabilization, linkers are implemented as an indispensable element in the development of vaccine protein [81]. Furthermore, the OmpA agonist (GenBank ID: AFS89615.1) was used as a TLR2 adjuvant and joined to the first CTL epitope using EAAAK linker [40]. When used alone, multi-epitope-based vaccines are poorly immunogenic and require coupling to adjuvants [82]. Adjuvants are ingredients added to vaccine formulations that affect particular immune responses to antigens, their development, stability, and longevity and are protective against infection [83]. Also, they gain great attention because the immune response to humoral and cell-mediated immune responses can be selectively modulated [84]. When assessing the vaccine construct, we observed that the non-adjuvant construct showed less antigenicity (0.624) than the adjuvant construct (0.7223) with the aid of Vaxijen server, which clearly states that the adjuvant is significant for the chimera.

The molecular weight of our vaccine candidate is ~68 kDa which is an average molecular weight for a multi-epitope vaccine. One of the fundamental requirements of many biochemical and functional analysis is the solubility of overexpressed recombinant protein within the *E.coli* host [49]. The constructed vaccine protein was found to be soluble which secured their easy access to the host. The basic nature of the vaccine is indicated by theoretical pI value. In addition, the predicted instability index shows that the protein will remain stable after expression, thus enhancing usage capacity further. The GRAVY score and aliphatic index depict the hydrophilicity and thermostability, respectively. The 3-dimensional structure modelling includes sufficient information on the spatial arrangement of crucial protein components and excellent support in the investigation of protein function, dynamics, ligand interactions and other proteins. The desirable properties of the vaccine construct enhanced significantly after refinement. The Ramachandran plot demonstrates that most residues are discovered in the favored and allowed regions (99.2%) with very few residues in the disallowed region; which depicts that the quality of the overall model is satisfactory. Besides, GDT-HA, RMSD value, MolProbity, Clash Score and Poor Rotamers values indicate the good quality of our designed vaccine construct. Different structure validation tools were used to identify errors in the modelled vaccine construct. The Z-score (-4.23) and ERRAT quality factor (78.2%) showed that the overall structure of the refined vaccine is adequate.

The HLA alleles maintain the response to T-cell epitopes, and in different ethnic communities, these alleles are highly polymorphic. The T-cell epitope should bind with more HLA alleles to obtain more population coverage. So, we selected the CTL and HTL

epitopes with their respective HLA alleles to predict the allele distribution worldwide. The findings showed that the chosen epitopes and their individual alleles cover ideally in numerous geographic regions of the globe. The highest population coverage was recorded at 99.16% in Chile Amerindian, and those epitopes and their respective HLA alleles cover 97.37% of the world population when used in combination. In Western Africa, in particular, Nigeria, Senegal and Mali, the epidemic of the LASV happened in most significant measure. Therefore, in these geographical regions, vaccine candidates are essential to protect people against LASV infection. The population coverage was found to be 88.86% at West Africa, where the virus first appeared and had several outbreaks.

Data-Driven protein-receptor docking analysis and molecular dynamics simulation was carried out to evaluate a potential immune interaction and stability between TLR2 and the vaccine protein, considering the use of a TLR2 agonist as an adjuvant in the constructed chimera. Energy minimization was conducted to minimize the potential energy of the whole system for the complete conformational stabilization of the vaccine protein-TLR2 docked complex. The energy minimizes the inappropriate structural geometry by replacing individual protein atoms, thus making the structure more stable with adequate stereochemistry. The derived eigenvalue indicates the stiffness of motion and the required energy for the complex deformability.

The immunoreactivity testing through serological assessment is one of the first steps in validating a candidate vaccine [85]. The expression of the recombinant protein in a suitable host is required. E.coli expression systems are determined for recombinant protein manufacturing [86,87]. Codon optimization had been performed with a view to achieving a high level of expression of our recombinant vaccine protein in E. coli K12. Both the codon adaptability index (0.98) and the GC content (53.63%) were promising for high-level protein expression in bacteria. Enhancing the stability of proteins is an indispensable objective in various biomedical and mechanical applications. In this study, we have introduced a disulphide bridging into the multi-epitope vaccine construct to improve protein thermostability, modify its practical features and assist in the analysis of genetic components.

The immune simulation revealed results consistent with typical immune responses. There was an overall increase in immune responses following repeated exposure of the antigen. The development of memory B-cells and T-cells were visible, with several months lasting memory B-cells. Another intriguing finding is that after the first injector concentrations of IFN- γ and IL-2 increased and were maintained at peaks after repeated antigen exposure. This finding shows elevated

TH cell concentrations and therefore effective Ig production that support a humoral response. Both dendritic and macrophage cells activity were satisfactory in our study. Besides, components like epithelial cells of the innate immune system were active. The Simpson Index, D suggests a possibility of different immune responses for clonal specificity analysis.

Conclusion

Lassa virus is an emerging viral pathogen which is characterized by severe haemorrhagic fever with a higher mortality rate, hence, become an increasing concern. Though antiviral drug Ribavirin showed some promises earlier, excessive toxicity and teratogenicity rendered its effectivity questionable. Knowing all the merits that a peptide vaccine has to offer, immunoinformatics strategies have been taken into consideration for designing a multi-epitope vaccine. Both T-cell and B-cell epitopes derived from different LASV proteins were included in the vaccine to produce an effective immune response. We believe that our vaccine will hopefully generate cell-mediated and humoral immune responses. The binding potential and interaction between vaccine protein and receptor were higher and stable. Besides, effective immune responses in real life were observed in immune simulation. However, further investigations both in vitro and in vivo are warranted to ensure its true potential to fight against Lassa fever.

Author Contributions

ZN and UKA conceived and designed the analysis; SBS and ZN performed immunoinformatic analyses; SBS prepared illustrations and wrote the manuscript; FA contributed in the molecular docking and dynamic simulation; SBS, SAK, ZH, SMRR and RT performed the antigenicity and epitope predictions; UKA and ZN contributed to the critical revision of the manuscript; UKA supervised the whole work; and all authors approved the final manuscript.

References

1. R.N. Charrel, X. De Lamballerie, Arenaviruses other than Lassa virus, *Antiviral Res.* 57 (2003) 89–100.
2. O. Ogbu, E. Ajuluchukwu, C.J. Uneke, Lassa fever in West African sub-region: An overview, *J. Vector Borne Dis.* 44 (2007) 1–11.
3. E. Fichet-Calvet, D.J. Rogers, Risk maps of lassa fever in West Africa, *PLoS Negl. Trop. Dis.* 3 (2009).
4. D. Safronetz, J.E. Lopez, N. Sogoba, S.F. Traore, S.J. Raffel, E.R. Fischer, H. Ebihara, L. Branco, R.F. Garry, T.G. Schwan, H. Feldmann, Detection of lassa virus, mali, *Emerg. Infect. Dis.* 16 (2010) 1123–1126.
5. A. Kitching, S. Addiman, S. Cathcart, L. Bishop, D. Krahe, M. Nicholas, J. Coakley, G. Lloyd, A fatal case of Lassa fever in London, January 2009, *Rapid Commun.* 14 (2009) 12–14.
6. W.E.R. WHO, Imported case of Lassa fever in the Netherlands, 2000.
7. W.H. Haas, T. Breuer, G. Pfaff, H. Schmitz, P. Kohler, M. Asper, P. Emmerich, C. Drosten, U. Golnitz, K. Fleischer, S. Gunther, Imported Lassa Fever in Germany: Surveillance

- and Management of Contact Persons, *Clin. Infect. Dis.* 36 (2003) 1254–1258.
- Victor B. Oti, A Reemerging Lassa Virus: Aspects of Its Structure, Replication, Pathogenicity and Diagnosis, *IntechOpen*. (2018).
- H.J. Hallam, S. Hallam, S.E. Rodriguez, A.D.T. Barrett, D.W.C. Beasley, A. Chua, T.G. Ksiazek, G.N. Milligan, V. Sathiyamoorthy, L.M. Reece, Baseline mapping of Lassa fever virology, epidemiology and vaccine research and development review-article, *Npj Vaccines*. 3 (2018).
- M. Djavani, I.S. Lukashevich, A. Sanchez, S.T. Nichol, M.S. Salvato, Completion of the Lassa Fever Virus Sequence and Identification of a RING Finger Open Reading Frame at the L RNA 5' End, *Virology*. 235 (1997) 414–418.
- T.I. Cornu, J.C. de la Torre, RING finger Z protein of lymphocytic choriomeningitis virus (LCMV) inhibits transcription and RNA replication of an LCMV S-segment minigenome, *J. Virol.* 75 (2001) 9415–9426.
- W. Cao, M.D. Henry, P. Borrow, H. Yamada, J.H. Elder, E. V Ravkov, S.T. Nichol, R.W. Compans, K.P. Campbell, M.B. Oldstone, Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus., *Science*. 282 (1998) 2079–2081.
- P.C. Bonner, W.P. Schmidt, S.R. Belmain, B. Oshin, D. Baglolle, M. Borchert, Poor housing quality increases risk of rodent infestation and lassa fever in refugee camps of sierra leone, *Am. J. Trop. Med. Hyg.* 77 (2007) 169–175.
- R.A. Keenlyside, J.B. McCormick, P.A. Webb, E. Smith, L. Elliott, K.M. Johnson, Case-Control Study of Mastomys Natalensis and Humans in Lassa Virus-Infected Households in Sierra Leone, *Am. Soc. Trop. Med. Hyg.* 32 (1983).
- S. Kernéis, L. Koivogui, N. Magassouba, K. Koulemou, R. Lewis, A. Aplogan, R.F. Grais, P.J. Guerin, E. Fichet-Calvet, Prevalence and risk factors of lassa seropositivity in inhabitants of the Forest Region of Guinea: A cross-sectional study, *PLoS Negl. Trop. Dis.* 3 (2009).
- M.E. Price, S.P. Fisher-Hoch, R.B. Craven, J.B. McCormick, A prospective study of maternal and fetal outcome in acute Lassa fever infection during pregnancy, 297 (1988) 584–587.
- P.B. Jahrling, R.A. Hesse, G.A. Eddy, K.M. Johnson, R.T. Callis, E.L. Stephen, Lassa virus infection of rhesus monkeys: Pathogenesis and treatment with ribavirin, *J. Infect. Dis.* 141 (1980) 580–589.
- J.B. McCormick, I.J. King, P.A. Webb, C.L. Scribner, R.B. Craven, K.M. Johnson, L.H. Elliott, R. Belmont-williams, Effective therapy with Ribavirin, *N. Engl. J. Med.* 314 (1986).
- D.M. Kochhar, Effects of exposure to high concentrations of ribavirin in devloping embryos, *Pediatr. Infect. Dis. J.* 9 (1990).
- S.P. Fisher-Hoch, S. Gborie, L. Parker, J. Huggins, Unexpected adverse reactions during a clinical trial in rural West Africa, *Antiviral Res.* 19 (1992) 139–147.
- M. Skwarczynski, I. Toth, Peptide-based synthetic vaccines, *Chem. Sci.* 7 (2016) 842–854.
- C.O. Jacob, M. Leitner, A. Zamir, D. Salomon, R. Arnon, Priming immunization against cholera toxin and E. coli heat-labile toxin by a cholera toxin short peptide-beta-galactosidase hybrid synthesized in E. coli., *EMBO J.* 4 (1985) 3339–3343.
- W. Li, M.D. Joshi, S. Singhanian, K.H. Ramsey, A.K. Murthy, Peptide vaccine: Progress and challenges, *Vaccines*. 2 (2014) 515–536.
- B.E. Pickett, E.L. Sadat, Y. Zhang, J.M. Noronha, R.B. Squires, V. Hunt, M. Liu, S. Kumar, S. Zaremba, Z. Gu, L. Zhou, C.N. Larson, J. Dietrich, E.B. Klem, R.H. Scheuermann, ViPR: an open bioinformatics database and analysis resource for virology research, *Nucleic Acids Res.* 40 (2012) D593–D598.
- I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, *BMC Bioinformatics*. 8 (2007) 4.
- M. V Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction, *BMC Bioinformatics*. 8 (2007) 424.
- M. Moutafsi, B. Peters, V. Pasquetto, D.C. Tschärke, J. Sidney, H.H. Bui, H. Grey, A. Sette, A consensus epitope prediction approach identifies the breadth of murine TCD8+ cell responses to vaccinia virus, *Nat. Biotechnol.* 24 (2006) 817–819.
- J.J.A. Calis, M. Maybeno, J.A. Greenbaum, D. Weiskopf, A.D. De Silva, A. Sette, C. Keşmir, B. Peters, Properties of MHC class I presented peptides that enhance immunogenicity, *PLoS Comput. Biol.* 9 (2013) e1003266.
- I. Dimitrov, L. Naneva, I. Doytchinova, I. Bangov, AllergenFP: Allergenicity prediction by descriptor fingerprints, *Bioinformatics*. 30 (2014) 846–851.
- S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, Open Source Drug Discovery Consortium, G.P.S. Raghava, In silico approach for predicting toxicity of peptides and proteins, *PLoS One*. 8 (2013) e73957.
- M. Nielsen, O. Lund, NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction, *BMC Bioinformatics*. 10 (2009) 296.
- S. Paul, J. Sidney, A. Sette, B. Peters, TepiTool: A pipeline for computational prediction of T cell epitope candidates, *Curr. Protoc. Immunol.* 2016 (2016) 18.19.1-18.19.24.
- R.V. Luckheeram, R. Zhou, A.D. Verma, B. Xia, CD4 +T cells: Differentiation and functions, *Clin. Dev. Immunol.* 2012 (2012).
- S.K. Dhanda, P. Vir, G.P.S. Raghava, Designing of interferon-gamma inducing MHC class-II binders, *Biol. Direct*. 8 (2013) 30.
- S.K. Dhanda, S. Gupta, P. Vir, G.P.S. Raghava, Prediction of IL4 Inducing Peptides, *Clin. Dev. Immunol.* 2013 (2013).
- G. Nagpal, S.S. Usmani, S.K. Dhanda, H. Kaur, S. Singh, M. Sharma, G.P.S. Raghava, Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential, *Sci. Rep.* 7 (2017) 1–10.
- B. Manavalan, R.G. Govindaraj, T.H. Shin, M.O. Kim, G. Lee, iBCE-EL: A New Ensemble Learning Framework for Improved Linear B-Cell Epitope Prediction, *Front. Immunol.* 9 (2018) 1695.
- U.K. Adhikari, M.M. Rahman, Overlapping {CD8} + and {CD4} + T-cell epitopes identification for the progression of epitope-based peptide vaccine from nucleocapsid and glycoprotein of emerging Rift Valley fever virus using immunoinformatics approach, *Infect. Genet. Evol.* 56 (2017) 75–91.
- H.-H. Bui, J. Sidney, K. Dinh, S. Southwood, M.J. Newman, A. Sette, Predicting population coverage of T-cell epitope-based diagnostics and vaccines, *BMC Bioinformatics*. 7 (2006) 153.
- R. Arai, H. Ueda, A. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein, *Protein Eng.* 14 (2001) 529–532.
- N. Nezafat, Y. Ghasemi, G. Javadi, M.J. Khoshnoud, E. Omidinia, A novel multi-epitope peptide vaccine against cancer: An in silico approach, *J. Theor. Biol.* 349 (2014) 121–134.
- C.N. Magnan, M. Zeller, M.A. Kayala, A. Vigil, A. Randall, P.L. Felgner, P. Baldi, High-throughput prediction of protein

- antigenicity using protein microarray data, *Bioinformatics*. 26 (2010) 2936–2943.
- I. Dimitrov, I. Bangov, D.R. Flower, I. Doytchinova, AllerTOP v.2 - A server for in silico prediction of allergens, *J. Mol. Model.* 20 (2014).
- M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein identification and analysis tools in the ExpASY server., *Methods Mol. Biol.* 112 (1999) 531–552.
- C.N. Magnan, A. Randall, P. Baldi, SOLpro: Accurate sequence-based prediction of protein solubility, *Bioinformatics*. 25 (2009) 2200–2207.
- D.W.A. Buchan, F. Minneci, T.C.O. Nugent, K. Bryson, D.T. Jones, Scalable web services for the PSIPRED Protein Analysis Workbench, *Nucleic Acids Res.* 41 (2013) W349–57.
- M. Källberg, H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu, J. Xu, Template-based protein structure modeling using the RaptorX web server, *Physiol. Behav.* 176 (2017) 139–148.
- L. Heo, H. Park, C. Seok, GalaxyRefine: Protein structure refinement driven by side-chain repacking., *Nucleic Acids Res.* 41 (2013) 384–388.
- N. Khatoun, R.K. Pandey, V.K. Prajapati, Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach, *Sci. Rep.* 7 (2017) 1–12.
- M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res.* 35 (2007) W407–10.
- C. Colovos, T.O. Yeates, Verification of protein structures: Patterns of nonbonded atomic interactions, *Protein Sci.* 2 (1993) 1511–1519.
- R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- J. Ponomarenko, H.-H. Bui, W. Li, N. Fusseder, P.E. Bourne, A. Sette, B. Peters, ElliPro: a new structure-based tool for the prediction of antibody epitopes, *BMC Bioinformatics*. 9 (2008) 514.
- D.B. Craig, A.A. Dombkowski, Disulfide by Design 2.0: A web-based tool for disulfide engineering in proteins, *BMC Bioinformatics*. 14 (2013) 0–6.
- A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel, D. Jahn, JCat: A novel tool to adapt codon usage of a target gene to its potential expression host, *Nucleic Acids Res.* 33 (2005) 526–531.
- P.M. Sharpl, W. Li, Codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications, *Nucleic Acids Res.* 15 (1987) 1281–1295.
- T. Lengauer, M. Rarey, Computational methods for biomolecular docking, *Curr. Opin. Struct. Biol.* (1996) 402–405.
- M. Hayes, M. Salvato, Arenavirus evasion of host anti-viral responses, *Viruses*. 4 (2012) 2182–2196.
- H.M. Berman, W.F. Bluhm, E. Philip, J. Marvin, H. Weissig, D. John, The Protein Data Bank, *Biol. Crystallogr.* D58 (2002) 899–907.
- D. Kozakov, D.R. Hall, B. Xia, K.A. Porter, D. Padhorny, C. Yueh, D. Beglov, S. Vajda, The ClusPro web server for protein-protein docking, *Nat. Protoc.* 12 (2017) 255–278.
- D.M.F. Van Aalten, B.L. De Groot, J.B.C. Findlay, H.J.C. Berendsen, A. Amadei, A comparison of techniques for calculating protein essential dynamics, *J. Comput. Chem.* 18 (1997) 169–181.
- K. Wüthrich, G. Wagner, R. Richarz, W. Braun, Correlations between internal mobility and stability of globular proteins, *Biophys. J.* 32 (1980) 549–560.
- J.R. López-Blanco, J.I. Aliaga, E.S. Quintana-Ortí, P. Chacón, IMODS: Internal coordinates normal mode analysis server, *Nucleic Acids Res.* 42 (2014) 271–276.
- N. Rapin, O. Lund, M. Bernaschi, F. Castiglione, Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system, *PLoS One.* 5 (2010).
- F. Castiglione, F. Mantile, P. De Berardinis, A. Prisco, How the interval between prime and boost injection affects the immune response in a computational model of the immune system, *Comput. Math. Methods Med.* 2012 (2012).
- J.A. Kovacs, P. Chacón, R. Abagyan, Predictions of protein flexibility: First-order measures, *Proteins Struct. Funct. Genet.* 56 (2004) 661–668.
- L. Zhang, Multi-epitope vaccines: A promising strategy against tumors and viral infections, *Cell. Mol. Immunol.* 15 (2018) 182–184.
- R.A. Shey, S.M. Ghogomu, K.K. Esoh, N.D. Nebangwa, C.M. Shintouo, N.F. Nongley, B.F. Asa, F.N. Ngale, L. Vanhamme, J. Souopgui, In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases, *Sci. Rep.* (2019) 1–18.
- [69] R.R. María, C.J. Arturo, J.A. Alicia, M.G. Paulina, A.O. Gerardo, The Impact of Bioinformatics on Vaccine Design and Development, in: F. Afrin, H. Hemeg, H. Ozbak (Eds.), *Vaccines*, InTech, Rijeka, 2017.
- K.L. Seib, X. Zhao, R. Rappuoli, Developing vaccines in the era of genomics: a decade of reverse vaccinology, *Clin. Microbiol. Infect.* 18 (2012) 109–116.
- R. Dash, R. Das, M. Junaid, M.F.C. Akash, A. Islam, S.M.Z. Hosen, In silico-based vaccine design against Ebola virus glycoprotein, *Adv. Appl. Bioinform. Chem.* 10 (2017) 11–28.
- A. Ikram, T. Zaheer, F.M. Awan, A. Obaid, A. Naz, R. Hanif, R.Z. Paracha, A. Ali, A.K. Naveed, H.A. Janjua, Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches, *Sci. Rep.* 8 (2018) 1–14.
- U.K. Adhikari, M. Tayebi, M. Mizanur Rahman, Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus, *J. Immunol. Res.* 2018 (2018). doi:10.1155/2018/6718083.
- M. Ali, R.K. Pandey, N. Khatoun, A. Narula, A. Mishra, V.K. Prajapati, Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection, *Sci. Rep.* 7 (2017) 1–13.
- S.K. Verma, S. Yadav, A. Kumar, In silico prediction of B- and T- cell epitope on Lassa virus proteins for peptide based subunit vaccine design., *Adv. Biomed. Res.* 4 (2015) 201.
- A. Faisal, S.H. Imtiaz, T. Zerín, T. Rahman, H.U. Shekhar, Computer aided epitope design as a peptide vaccine component against Lassa virus, *Bioinformation.* 13 (2017) 417–429.
- M.U. Hossain, T.M. Omar, A.R. Oany, K.M.K. Kibria, A.Z. Shibly, M. Moniruzzaman, S.R. Ali, M.M. Islam, Design of peptide-based epitope vaccine and further binding site scrutiny led to groundswell in drug discovery against Lassa virus, *3 Biotech.* 8 (2018) 1–14. doi:10.1007/s13205-018-1106-5.
- N.R. Cooper, G.R. Nemerow, The role of antibody and complement in the control of viral infections, *J. Invest. Dermatol.* 83 (1984) S121–S127.

- R. Bacchetta, S. Gregori, M.G. Roncarolo, CD4+ regulatory T cells: Mechanisms of induction and effector function, *Autoimmun. Rev.* 4 (2005) 491–496.
- K.C. Garcia, L. Teyton, I.A. Wilson, Structural Basis of T Cell Recognition, *Annu. Rev. Immunol.* 17 (1999) 369–397. doi:10.1146/annurev.immunol.17.1.369.
- S. Shamriz, H. Ofoghi, N. Moazami, Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application, *Comput. Biol. Med.* 76 (2016) 24–29.
- B. Meza, F. Ascencio, A.P. Sierra-Beltrán, J. Torres, C. Angulo, A novel design of a multi-antigenic, multistage and multi-epitope vaccine against *Helicobacter pylori*: An in silico approach, *Infect. Genet. Evol.* 49 (2017) 309–317.
- S. Lee, M.T. Nguyen, Recent Advances of Vaccine Adjuvants for Infectious Diseases, *Immune Netw.* 15 (2015) 51.
- S.R. Bonam, C.D. Partidos, S.K.M. Halmuthur, S. Muller, An Overview of Novel Adjuvants Designed for Improving Vaccine Efficacy, *Trends Pharmacol. Sci.* 38 (2017) 771–793.
- A. Gori, R. Longhi, C. Peri, G. Colombo, Peptides for immunological purposes: Design, strategies and applications, *Amino Acids.* 45 (2013) 257–268.
- R. Chen, Bacterial expression systems for recombinant protein production: *E. coli* and beyond, *Biotechnol. Adv.* 30 (2012) 1102–1107.
- G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: Advances and challenges, *Front. Microbiol.* 5 (2014) 1–17.