

Conglomeration of highly antigenic nucleoproteins to inaugurate a heterosubtypic next generation vaccine candidate against Arenaviridae family

Kazi Faizul Azim^{a,b}, Tahera Lasker^a, Rahima Akter^a, Mantasha Mahmud Hia^a, Omar Faruk Bhuiyan^c, Mahmudul Hasan^{a,d}, Md. Nazmul Hossain^{a,b*}

^a*Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet-3100, Bangladesh*

^b*Department of Microbial Biotechnology, Sylhet Agricultural University, Sylhet-3100, Bangladesh*

^c*Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh*

^d*Department of Pharmaceuticals and Industrial Biotechnology, Sylhet Agricultural University, Sylhet-3100*

***Corresponding author:**

Md. Nazmul Hossain

Assistant Professor

Department of Microbial Biotechnology

Sylhet Agricultural University, Sylhet-3100, Bangladesh

E-mail: nhossain.mib@sau.ac.bd

Telephone: +8801718826131

Abstract

Arenaviral infections often resulting in lethal Hemorrhagic Fevers (HF) affect primarily African and South American regions. To date, there is no FDA-approved licensed vaccine against arenaviruses and treatments have been limited to supportive therapies. Hence, the study was employed to design a highly immunogenic heterosubtypic vaccine candidate against Arenaviridae family. The whole proteomes of Lassa virus (LASV), Lymphocytic Chorio Meningitis Virus (LCMV), Lujo virus and Guanarito virus were retrieved from NCBI database and assessed to determine the most antigenic viral proteins. Only the conserved sequences were used for T cell and B cell epitope prediction to ensure protective response against a wide range of viral strains. For each virus, nucleoproteins were identified as most antigenic which generated a plethora of antigenic epitopes. The proposed epitopes were highly conserved (up to 100%) and showed high cumulative population coverage. Moreover, results revealed that among the top epitopes, T cell epitope GWPYIGSRS were conserved in Argentine mammarenavirus (Junin virus) and Brazilian mammarenavirus (Sabia virus), while B cell epitope NLLYKICLSG were conserved in Bolivian mammarenavirus (Machupo virus) and Brazilian mammarenavirus (Sabia virus), indicating the possibility of final vaccine constructs to confer broad range immunity in the host. A total 3 constructs were designed by the combination of top epitopes from each protein along with suitable adjuvant and linkers. Different physicochemical properties revealed the superiority of construct V1 in terms of safety and efficacy. Docking analysis of the refined vaccine structure with different MHC molecules and human immune receptors were also biologically significant. The vaccine receptor complex (V1-TLR3) showed minimal deformability at molecular level. Moreover, construct V1 was compatible for insertion into pET28a(+) vector and heterologous cloning in *E. coli* strain K12. However, the results were based on different sequence analysis and various immune databases. Further wet lab based studies using model animals are highly recommended for the experimental validation of the designed vaccine candidates.

Keywords: Arenavirus, hemorrhagic fevers, nucleoprotein, recombinant vaccine, molecular docking, *in silico* cloning

Conglomeration of highly antigenic nucleoproteins to inaugurate a heterosubtypic next generation vaccine candidate against Arenaviridae family

Introduction

Arenaviral infections, primarily affecting South American and African regions, have traditionally been neglected as tropical diseases (Brisse and Ly, 2019). Although most infections are mild, transmission occurs through rodents to humans, sometimes resulting in severe Hemorrhagic Fevers (HF) with high death rates (Salvato, 2012). They are divided into 2 categories, Old World (eastern hemisphere) and New World (western hemisphere) arenavirus. Old World arenaviruses and the diseases they cause include Lassa virus (Lassa fever), Lujo virus, and Lymphocytic Chorio Meningitis Virus (LCMV) (meningitis, encephalitis, congenital fetal infection, severe disease with multiple organ failure). New World arenaviruses include Junin (Argentine hemorrhagic fever), Machupo (Bolivian hemorrhagic fever), Guanarito (Venezuelan hemorrhagic fever), Sabia (Brazilian hemorrhagic fever) and Chapare virus (Shao et al., 2015). Though, arenaviral outbreaks have been restricted to certain geographic areas, known cases of exportation of arenaviruses from endemic regions and socioeconomic challenges to control the rodent reservoirs locally raised serious concerns about the potential for larger outbreaks in the future. Currently, there are no FDA-approved vaccines for arenaviruses and treatments have been limited to supportive therapy and use of non-specific nucleoside analogs like Ribavirin (Brisse and Ly, 2019; Shoemaker et al., 2015). Though, investigational vaccines exist for Argentine hemorrhagic fever (Ambrosio et al., 2011) and Lassa fever (Carrion et al., 2007), no preventive strategies have been available to treat diseases caused by LCMV, Lujo and Gunarito virus (Daniel et al., 2012; Cheng et al., 2015).

Lassa fever is an animal-borne, or zoonotic, acute viral illness which is endemic in parts of West Africa including Sierra Leone, Liberia, Guinea and Nigeria. Neighboring countries are also at risk, as the animal vector for Lassa virus, “multimammate rat” (*Mastomys natalensis*) is distributed throughout the region. It was estimated that this virus infects roughly 300,000 to 500,000 individuals per year yielding approximately 5,000 deaths (Ogbu et al., 2007; Houlihan et al., 2017). In some areas of Sierra Leone and Liberia, it is known that 10-16% of people

admitted to hospitals annually have Lassa fever, demonstrating the serious impact the disease has on the region. Case fatality rates in infected patients can reach more than 80% (Houlihan et al., 2017). Making a correct diagnosis of Lassa fever is made difficult by the wide spectrum of clinical effects that manifest, ranging from asymptomatic to multi-organ system failure and death. Common symptoms associated with cases of LASV are fever, sore throat, red eyes, headache, weakness, retrosternal pain, facial edema, generalized abdominal pain, epistaxis and haemoptysis. In severe cases bleeding from mucousal membranes such as the mouth can also be observed. The only available drug, ribavirin is only effective if administered early in infection (within the first 6 days after disease onset). A fundamental understanding of the mechanisms of antibody-mediated neutralization of Lassa virus may have significant implications for the generation epitope-targeted vaccines. There is no approved vaccine for humans against LASV as of 2019 (Yun et al., 2012).

LCMV is another rodent-borne (*Mus musculus*) prototypic virus of Arenaviridae family that can cause substantial neurological problems, including meningitis, encephalitis, and neurologic birth defects, particularly among prenatal and immune compromised humans (Daniel et al., 2012). LCMV Infections have been reported in Europe, America, Australia, Japan, and may occur wherever infected rodent hosts of the virus are found. Several serologic studies conducted in urban areas have shown that the prevalence of LCMV antibodies in human populations range from 2% to 5% (Wright et al., 1997). A meta-analysis of all reported cases of congenital LCMV infection revealed a mortality rate of 35% by 21 months of age (Bonthius et al., 2007; Wright et al., 1997). Most of the survivors have severe neuro-developmental disorders, including microcephaly, poor somatic growth, profound vision impairment, severe seizure disorders, spastic weakness, and substantial mental retardation (Bonthius et al., 2007; Larsen et al., 2001). An effective antiviral therapy for LCMV infection has not yet been developed and still, there is no vaccine to prevent LCMV infection (Daniel et al., 2012). Although Ribavirin has had mixed success in the treatment of severe infections, but is limited to off-label use and can cause muscular toxicity (Mendenhall et al., 2011). Recently, Lujo virus was isolated as a newly discovered novel arenavirus associated with a VHF outbreak in southern Africa in 2008. It was found to cause a fulminant viral hemorrhagic fever (LUHF) syndrome characterized by nonspecific symptoms such as fever, malaise, myalgias, sore throat, nausea, vomiting and non-

bloody diarrhea followed with variable retrosternal or epigastric pain, usually progressing to bleeding, shock and multiorgan failure (Sewlall et al., 2017). This virus has been associated with an outbreak of five cases in September and October 2008 in cities namely Lusaka (Zambia) and Johannesburg (Republic of South Africa). The case fatality rate was 80% (Paweska et al., 2014). In essence, the distribution and prevalence of LUHF in humans and rodents is unknown, as are the ecology, distribution, and mode of transmission from reservoir host to humans. Ribavirin's effectiveness against Lujo virus remains unknown as well (Paweska et al., 2009).

Guanarito (GTO) virus is the etiological agent of Venezuelan haemorrhagic fever, a rodent borne zoonosis which is endemic in the Northern America (Manziona et al., 1998). Infections are characterized by having the onset of pulmonary congestion and edema, renal and cortical necrosis and haemorrhage in different sites like mucous membranes, major internal organs, digestive and urinary tracts (Craighead, 2000). The short tailed cane mouse, *Zygodontomys brevicauda* (a grassland rodent) acts as a reservoir of Guanarito virus (Fulhorst et al., 1999). Rodent to human transmission may occur via inhalation of virus in aerosolized droplets of secretions or excretions from infected rodents or via contact with the virus through mucosal or cutaneous routes (Ter Meulen et al., 1996). Venezuelan hemorrhagic fever was first recognized as a distinct clinical entity in 1989 during an outbreak of hemorrhagic fever that began in the Municipality of Guanarito in southern Portuguesa (Salas et al., 1991, Tesh et al., 1994). Results of an epidemiologic study of 165 cases of Venezuelan haemorrhagic fever indicated that the disease is seasonal and the number of affected people peaks in november to january. The overall fatality rate was 33.3 % among the 165 cases despite hospitalization and vigorous supportive care (Manziona, 1998). Although successful in many cases traditional vaccines are associated with several demerits (Stratton et al., 2003; Hasan et al., 2019a). Developing vaccines for the organisms not grown in culture is very costly and the yield of vaccines is very low. There is also danger of non-virulent organisms getting converted to virulent ones (Hasson et al., 2015; Kaufmann et al., 2014). Vaccinations by such organisms may cause the disease itself. Recombinant vaccines produced using immunoinformatic approaches, on the contrary, offer some advantages while reducing the time and cost for production (Azim et al., 2019). Hence, the study was conducted to design a highly antigenic polyvalent vaccine against the viruses of Arenaviridae family which are responsible for severe hemorrhagic fevers in human.

Materials and Methods

In the present study, reverse vaccinology approach was employed to design a novel multiepitope subunit vaccine against the most deadly viruses of Arenaviridae family requiring urgent need for effective medications and preventive measures. The flow chart summarizing the entire protocol of *in silico* strategy for developing a chimeric polyvalent vaccine has been illustrated in Fig. 1.

Retrieval of viral proteomes and antigenic protein selection

NCBI server was used for the selection of viral strains and the whole proteomes of LASV, LCMV, Lujo virus and Guanarito virus were retrieved from the database (<https://www.ncbi.nlm.nih.gov/genome/>). Study of viral genus, family, host, disease, transmission, and genome were performed by using ViralZone (<https://viralzone.expasy.org/>). The most potent immunogenic proteins were identified individually for all the viruses after determining the antigenicity score via VaxiJen v2.0 server (Doytchinova & Flower, 2007). Different physiochemical parameters of the proteins were analyzed through ProtParam tool (Gasteiger et al., 2003).

Retrieval of homologous protein sets and identification of conserved regions

The selected proteins from each virus were used as query and homologous sequences were retrieved using BLASTp tool from NCBI. Multiple sequence alignment (MSA) was further performed to find out the common fragments for each set of proteins using CLUSTAL Omega (Sievers et al., 2011) along with 1000 bootstrap value and other default parameters to fabricate the alignment.

Antigenicity prediction and transmembrane topology analysis of the conserved fragments

The conserved regions from the selected proteins were again screened to demonstrate their antigenicity via VaxiJen v2.0 (Doytchinova & Flower, 2007). The fragments were subjected

TMHMM v0.2 server (Krogh et al., 2001) for transmembrane topology prediction. Only the common fragments were used to predict the highly immunogenic T-cell epitopes.

Prediction of T-cell epitopes, transmembrane topology screening and antigenicity analysis

From the IEDB database (Immune Epitope Database), both MHC-I (<http://tools.iedb.org/mhci/>) and MHC-II prediction tools (<http://tools.iedb.org/mhcii/>) were used to predict the MHC-I binding and MHC-II binding peptides respectively (Vita et al., 2014). The TMHMM server (<http://www.cbs.dtu.dk/services/TMHMM/>) predicted the transmembrane helices in proteins (Krogh et al., 2001). Again, VaxiJen v2.0 server (<http://www.ddg-pharmfac.net/vaxijen/>) was used to determine the antigenicity of predicted CTL (Cytotoxic T-lymphocytes) and HTL (Helper T-lymphocytes) epitopes.

Population coverage analysis, allergenicity assessment and toxicity analysis

Population coverage analysis is crucial due to variation of HLA distribution among different ethnic groups and geographic regions around the world. In this study, population coverage for each individual epitope was analyzed by IEDB population coverage calculation tool analysis resource (<http://tools.iedb.org/population/>). The allergenicity pattern of the predicted epitopes were determined through four servers i.e. AllerTOP (Dimitrov et al., 2013), AllergenFP (Dimitrov et al., 2014), Allergen Online (Goodman et al., 2016) and Allermatch (Fiers et al., 2004) were, while the toxicity level was demonstrated using ToxinPred server (Hasan et al., 2019b).

Epitope conservancy analysis

To determine the extent of desired epitope distributions in the homologous protein set epitope conservancy analysis is a vital step. IEDB's epitope conservancy analysis tool (<http://tools.iedb.org/conservancy/>) was selected for the analysis of conservancy level by considering the identities of the selected proteins.

Designing three-dimensional (3D) epitope structure and molecular docking analysis

The top epitopes were allowed for the docking study after analyzing through different bioinformatics tools. PEP-FOLD server was used to design and retrieve the 3D structure of most potent selected epitopes (Maupetit et al., 2010). The docking was conducted using AutoDOCKVina program at 1.00°A spacing (Morris et al., 2009). The exhaustiveness parameters were kept at 8.00 and the numbers of outputs were set at 10. OpenBabel (version 2.3.1) was used to convert the output PDBQT files in PDB format. The docking interaction was visualized with the PyMOL molecular graphics system, version 1.5.0.4 (<https://www.pymol.org/>).

B-Cell epitope prediction and Screening

Three different algorithms from IEDB were used to identify the most potent B cell (BCL) epitopes of the selected antigenic proteins. The algorithms include Bepipred linear epitope prediction, Emini surface accessibility (Emini et al., 1985) and Kolaskar & Tongaonkar antigenicity scale analysis (Kolaskar & Tongaonkar, 1990). The top B cell epitopes were selected based on their allergenicity pattern and VaxiJen score.

Epitope cluster analysis and vaccine construction

Epitope cluster analysis tool from IEDB (Dhanda et al., 2018) was used to identify the overlapping peptides among the top CTL, HTL and BCL epitopes at minimum sequence identity threshold of 100%. The identified clusters and singletons (unique epitopes) were utilized in a sequential manner to design the final vaccine constructs. Each vaccine proteins started with an adjuvant followed by the top epitopes. Interactions of adjuvants with toll like receptors (TLRs) induce robust immune reactions (Rana & Akhter, 2016). Hence, three different adjuvants were utilized in the study including beta defensin (a 45 mer peptide), L7/L12 ribosomal protein and HABA protein (*M. tuberculosis*, accession number: AGV15514.1). PADRE sequence was also incorporated along with the adjuvant and peptides with a view to overcome the problem caused

by highly polymorphic HLA alleles. EAAAK, GGGS, GPGPG and KK linkers were used to conjugate the adjuvant, CTL, HTL and BCL epitopes respectively.

Allergenicity, antigenicity and solubility prediction of different vaccine constructs

Allergenicity pattern of the designed vaccines were determined by AlgPred v.2.0 (Saha and Raghava, 2000). VaxiJen v2.0 server (Doytchinova and Flower, 2007) was further used to evaluate the probable antigenicity of the constructs in order to suggest the superior vaccine candidate. Protein-sol software (Hebditch et al., 2017) analyzed the solubility score of the proposed vaccine candidates by calculating the surface distribution charge, hydrophobicity and the stability at 91 different combinations of pH and ionic strength.

Physicochemical characterization and secondary structure analysis

ProtParam, a tool provided by ExPASy server (Hasan et al., 2015a) was used to functionally characterize the vaccine proteins. Molecular weight, aliphatic index, isoelectric pH, hydropathicity, instability index, GRAVY values, estimated half-life and other physicochemical properties were analyzed. The Prabi server (<https://npsa-prabi.ibcp.fr/>) predicted the alpha helix, beta sheet and coil structure of the vaccine constructs through GOR4 secondary structure prediction method.

Vaccine tertiary structure prediction, refinement, validation and disulfide engineering

I-TASSER server (Zhang, 2010) performed 3D modeling of the designed vaccines depending on the level of similarity between target protein and available template structure in PDB (Hasan et al., 2015b). Refinement was conducted using ModRefiner (Xu and Zhang, 2011) to improve the accuracy of the predicted 3D modeled structure. The refined protein structure was further validated by Ramachandran plot assessment through MolProbity software (Davis et al., 2004). DbD2, an online tool was used to design disulfide bonds for the designed construct (Craig & Dombkowski, 2013).

Conformational B-cell and IFN- γ inducing epitopes prediction

The conformational B-cell epitopes in the vaccine were predicted via ElliPro server (<http://tools.iedb.org/ellipro/>) with minimum score 0.5 and maximum distance 7 angstrom (Ponomarenko et al., 2004). Moreover, IFN- γ inducing epitopes within the vaccine were predicted using IFNepitope server (Hajighahramani et al., 2017) with motif and SVM hybrid prediction approach (<http://crdd.osdd.net/raghava/ifnepitope/scan.php>).

Protein-protein docking

Different Pattern Recognition Receptors (PRRs) including both membrane associated Toll-like receptors (TLR-3, TLR-7) and cytoplasmic RIG-I-like receptors (RIG-I, MDA5) can recognize infections caused by the members of Arenaviridae (Borrow et al., 2010). Subsequent studies also revealed that α -dystroglycan (α DG) expressed at high levels in skeletal muscle (Ibraghimov-Beskrovnaya et al., 1993) acts as a primary receptor for Old World arenaviruses. The 3D structure of different MHC molecules and human receptors (TLR-3, RIG-I, MDA5, α DG) were retrieved from RCSB protein data bank. Protein-protein docking was conducted to determine the binding affinity of designed vaccines with different HLA alleles and human immune receptors via PatchDock (Hasan et al., 2019c). Docked complexes from PatchDock were subjected to the FireDock server to refine the complexes.

Molecular dynamics simulation

Molecular dynamics study was performed to strengthen the in silico prediction via iMODS server (Lopez-Blanco et al., 2017). The structural dynamics of protein complex (V1-Toll like receptor 3) was investigated by using this server due to its much faster and effective assessments than other molecular dynamics (MD) simulations tools (Awan et al., 2017). The iMODS server explained the collective motion of proteins by analyzing the normal modes (NMA) in internal coordinates (Tama & Brooks, 2006). Stability was determined by comparing the essential dynamics of proteins to their normal modes (Aalten et al., 1997). The direction of the complex and extent of the motions was predicted in terms of deformability, eigenvalues, B-factors and covariance.

Codon adaptation and in silico cloning

Codon adaptation was performed to accelerate the expression of construct V1 in *E. coli* strain K12. JCAT server was used for this purpose, while Rho independent transcription termination, prokaryote ribosome-binding site and cleavage sites of several restriction enzymes (i.e. BglIII and BglI) were avoided during the operation (Grote et al., 2005). The optimized sequence of vaccine protein V1 was reversed and then conjugated with BglIII and BglI restriction site at the N-terminal and C-terminal sites respectively. To insert the adapted sequence into pET28a(+) vector between the BglIII (401) and BglI (2187), SnapGene (Solanki & Tiwari, 2018) restriction cloning module was utilized.

Results

Retrieval of viral proteomes and antigenic protein selection

The entire viral proteomes of LASV, LCMV, Lujo virus and Guanarito virus were retrieved from NCBI protein database in FASTA format. Only the structural proteins were prioritized for vaccine candidacy. Among the rescued proteins, Nucleoproteins from Lassa virus (AAX49342) and LCMV (ADY11071), and Nucleocapsid proteins from Guanarito virus (AAS55657) and Lujo virus (AFP21515) were selected for vaccine development on the basis of highest antigenicity score (Supplementary file 1). Different physiochemical parameters of the proteins analyzed via ProtParam tool are shown in Table 1.

Table 1: ProtParam analysis of selected antigenic proteins

<i>Virus</i>	<i>Selected proteins</i>	<i>Accession ID</i>	<i>Molecular weight</i>	<i>Instability index</i>	<i>Half life</i>	<i>pI</i>	<i>Amino acids</i>	<i>Extinction coefficient</i>
Lassa virus	Nucleoprotein	ADY11071	62998.24	36.64	>10h	8.6	569	53860
LCMV	Nucleoprotein	AAX49342	62177.47	40.20	>10h	8.8	558	56380
Lujo virus	Nucleocapsid protein	AFP21515	63104.55	39.35	30h	8.7	558	51630
Guanarito virus	Nucleocapsid protein	AAS55657	62345.44	34.93	>10h	8.2	560	46410

Retrieval of homologous protein sets and identification of conserved regions

Different homologous protein sets for each protein were generated after BLASTp search using NCBI BLAST tools. A total of 5, 12, 3, and 9 conserved fragments were found among the LASV, LCMV, Lujo and Guanarito viral proteins respectively (Table 2).

Table 2: Identified conserved regions among different homologous protein sets of LASV, LCMV, Lujo virus and Guanarito virus

<i>Virus</i>	<i>Protein</i>	<i>Conserved region</i>	<i>Vaxijen</i>	<i>Topology</i>
Lassa	Nucleoprotein	YCSNIKLQVVKDAQALLHGLDFSEV	0.5300	inside
		NNQFGTMPSLTLACLTKQGQVDLND	1.0428	outside
		SGYNFSLGAAVKAGACMLDGGNMLET	0.6470	outside
		PGERNPYENILYKICLSGDGWPYIASRTSI	0.0358	outside
		WMDIEGRPEDPVE	1.0407	outside
LCMV	Nucleoprotein	KSFQWTQALRRELQ	0.5501	inside
		LLNGLDFSEVSNVQRIMRKE	0.5746	outside
		DKDLQRLRSLNQTV	0.3230	inside
		DSSLLNNQFGTMPSLTM	0.4840	outside
		LNDVVQALTDLGLLYTVKYPNL	0.6929	outside
		KLNMFVSDQVGDRNPYE	0.3663	outside
		NFSLGAAVKAGAALL	1.2041	outside
		GGNMLESILIKPSN	0.2419	outside
		NILYKVCLSGEGWPYIACRT	0.2895	outside
		MLLKDLMSGIDPN	0.5306	outside
		GALPQGMVLSCQGS	0.1074	outside
		TPHCALMDCIIFESA	0.0003	outside
Lujo virus	Nucleocapsid protein	VVRVWDVKDSSLLNNQFGTMPSLT	0.5378	inside
		HFYREPTDQKQFK	-0.2593	inside
		KKEITPHCALMDCI	0.4836	inside
Guanarito virus	Nucleocapsid	DAKLIADSLDFTQVS	0.9498	outside
		PRMYMGNLTQSQLEKRAGILR	0.1864	inside
		AGGVVRLWDVSDPSKLNNQFGSMPALTIA	0.5374	outside
		CMTVQGGETMNNVVQALTSGLLYTVKYPN	0.5873	outside
		LDDLEKLTLEHDCLQITKDESALNISGYN		
		FLSAAVKAGASLIDGGNMLETI	0.5427	outside
		LLYKLCLSGEGWPYIGRSQI	0.3678	outside
		TIWMDIEGPPTDPVELAVFQPSSG	-0.0436	outside
		TVQGADDIKKLFDIHGRKDLKLV DVRLTGE	0.9209	inside
VWEKFGHLCRAHNGVIVPKKK	0.9542	outside		

Antigenicity prediction and transmembrane topology analysis of the conserved fragments

Results showed that 4, 6, 2, 6 and conserved sequences from identified Lassa, LCMV, Lujo, and Guanarito viral proteins respectively met the criteria of default threshold level in VaxiJen (Table 2). Moreover, transmembrane topology screening revealed, among the immunogenic conserved sequences 3, 2, 5 and 5 sequences from the corresponding proteins fulfilled the criteria of exomembrane characteristics (Table 2).

Prediction of T-cell epitopes, transmembrane topology screening and antigenicity analysis

Numerous immunogenic epitopes from the conserved sequences were generated that could bind maximum number of HLA cells with high binding affinity (Supplementary file 2 and Supplementary file 3). Top epitopes from each of the protein were selected as putative T cell epitope candidates based on their transmembrane topology screening and antigenicity score (Table 3). Epitopes with a positive score of immunogenicity exhibited potential to elicit effective T-cell response.

Population coverage, allergenicity assessment and toxicity analysis of T-cell epitopes

Results showed that population from the most geographic areas can be covered by the predicted T-cell epitopes. Population coverage results for the epitopes of four different viral proteins are shown in Fig. 2. Through the allergenicity assessment by four servers (i.e. AllerTOP, AllergenFP, Allergen online, Allermatch), epitopes that were found to be non-allergen for human were identified (Supplementary file 4 and Supplementary file 5). Epitopes those were indicated as allergen for human and classified as toxic or undefined were removed from the predicted list of epitopes (Table 3).

Table 3: Predicted T-cell (CTL and HTL) epitopes of Nucleoproteins (Lassa virus and LCMV) and Nucleocapsid proteins (Lujo virus and Guanarito virus)

<i>Types</i>	<i>Virus</i>	<i>Epitope</i>	<i>Start</i>	<i>End</i>	<i>VaxiJen Score</i>	<i>No. of interacting HLA's</i>	<i>Conservancy (identity <=100%)</i>
<i>MHC-I peptides (CTL epitopes)</i>	Lassa Virus (LASV)	IEGRPEDPV	4	12	1.1593	81	100.00% (100/100)
		NFSLGAAVK	4	12	1.8014	27	100.00% (100/100)
		PSLTLAHLT	8	16	1.5130	27	100.00% (100/100)
		YNFSLGAAV	3	11	1.5124	54	100.00% (100/100)
		SLGAAVKAG	6	14	1.4140	27	100.00% (100/100)
		MPSLTLACL	7	15	1.3792	81	97.00% (97/100)
		FSLGAAVKA	5	13	1.3262	81	100.00% (100/100)
		GYNFSLGAA	2	10	1.2890	54	100.00% (100/100)
		DIEGRPEDP	3	11	1.2557	27	100.00% (100/100)
		SLGAAVKAGA	6	15	1.1754	27	100.00% (100/100)
	LCMV	NFSLGAAVK	1	9	1.8014	54	67.00% (67/100)
		LDFSEVSNV	5	13	1.7221	81	91.00% (91/100)
		NGLDFSEVS	3	11	1.6145	81	40.00% (40/100)
		GLDFSEVSN	4	13	1.5849	27	81.00% (81/100)
		DLGLLYTVK	10	18	1.4806	81	41.00% (41/100)
		SLGAAVKAG	3	12	1.1754	27	69.00% (69/100)
		LNLDFSEV	2	10	1.3833	81	40.00% (40/100)
		DFSEVSNVQ	6	14	1.3277	81	91.00% (91/100)
		FSLGAAVKA	2	10	1.3262	81	69.00% (69/100)
		TDLGLLYTV	9	17	1.2339	81	40.00% (40/100)
	Lujo Virus	WDVKDSSL	5	13	1.7861	54	100.00% (4/4)
		VWDVKDSSL	4	12	1.7406	27	100.00% (4/4)
		VWDVKDSSL	4	13	1.4658	27	100.00% (4/4)
		DVKDSSLN	6	14	1.2269	54	100.00% (4/4)
		RVWDVKDSSL	3	12	0.9679	27	100.00% (4/4)
		KDSSLNMQ	8	16	0.8111	54	100.00% (4/4)
		RVWDVKDSS	3	11	0.7777	54	100.00% (4/4)
		QFGTMPSLT	16	24	0.7616	27	100.00% (4/4)
		LNNQFGTMP	13	21	0.7367	27	100.00% (4/4)
		KDSSLNMQF	8	17	0.6510	27	100.00% (4/4)
	Guanarito Virus	DSLDFTQVS	7	15	1.8576	27	24.00% (24/100)
		ADSLDFTQVS	6	15	1.5975	27	24.00% (24/100)
		GWPYIGSRS	11	19	1.5322	27	99.00% (99/100)
		LKLVDRVRLTG	20	29	1.4348	27	24.00% (24/100)
		ADSLDFTQV	6	14	1.4384	27	24.00% (24/100)
		KYPNLDDLE	27	35	1.4366	27	26.00% (26/100)
VWEKFGHLC		1	9	1.3953	27	29.00% (29/100)	

		<i>MHC-II peptides (HTL epitopes)</i>					
		Peptide	HLA-A*02:01	HLA-B*07:01	HLA-C*01:02	HLA-DP*02:01	HLA-DQ*02:01
	Lassa Virus (LASV)	LVDVRLTGE	22	30	1.3388	27	24.00% (24/100)
		TSLGLLYTVK	18	27	1.3006	27	78.00% (78/100)
		SLGLLYTVK	19	27	1.2711	27	79.00% (79/100)
	LCMV	PSLTLACLTKQGQVD	8	22	1.1813	27	100.00% (100/100)
		NFSLGAAVKAGACML	4	18	1.1070	27	100.00% (100/100)
		SGYNFSLGAAVKAGA	1	15	1.0922	27	100.00% (100/100)
		NNQFGTMPSLTLACL	1	15	1.0193	27	91.00% (91/100)
		FSLGAAVKAGACMLD	5	19	0.9323	27	100.00% (100/100)
		LTLACLTKQGQVDLN	10	24	1.4060	27	100.00% (100/100)
		SLTLACLTKQGQVDL	9	23	1.3579	27	100.00% (100/100)
		YNFSLGAAVKAGACM	3	17	1.1719	27	100.00% (100/100)
		TMPSTLACLTKQGQ	6	20	1.1134	27	97.00% (97/100)
		MPSLTLACLTKQGQV	7	21	1.0459	27	97.00% (97/100)
		NFSLGAAVKAGAALL	1	15	1.2041	27	34.00% (34/100)
		LTDLGLLYTVKYPNL	8	22	1.1861	27	40.00% (40/100)
	ALTDLGLLYTVKYPN	7	21	0.9756	27	40.00% (40/100)	
	LNGLDFSEVSNVQRI	2	16	0.9237	27	40.00% (40/100)	
	LLNGLDFSEVSNVQR	1	15	0.8913	27	40.00% (40/100)	
	VQALTDLGLLYTVKY	5	19	0.8413	27	40.00% (40/100)	
	VVQALTDLGLLYTVK	4	18	0.8009	27	40.00% (40/100)	
	NGLDFSEVSNVQRIM	3	17	0.7633	27	40.00% (40/100)	
	QALTDLGLLYTVKY	6	20	0.6954	27	40.00% (40/100)	
	DVVQALTDLGLLYTV	3	17	0.5867	27	40.00% (40/100)	
	Lujó Virus	VWDVKDSSLLNNQFG	4	18	1.1337	27	100.00% (4/4)
		WDVKDSSLLNNQFGT	5	19	1.1169	27	100.00% (4/4)
		DVKDSSLLNNQFGTM	6	20	0.8891	27	100.00% (4/4)
		RVWDVKDSSLLNNQF	3	17	0.7122	27	100.00% (4/4)
		VRVWDVKDSSLLNNQ	2	16	0.6717	27	100.00% (4/4)
		VKDSSLLNNQFGTMP	7	21	0.6619	27	100.00% (4/4)
		KDSSLLNNQFGTMPS	8	22	0.5533	27	100.00% (4/4)
		VVRVWDVKDSSLLNN	1	15	0.4941	27	100.00% (4/4)
		SLLNNQFGTMPSLT	10	24	0.4797	27	100.00% (4/4)
		DSSLLNNQFGTMPSL	9	23	0.3935	27	100.00% (4/4)
	Guanarito Virus	LFDIHGRKDLKLV DV	11	25	1.5311	27	24.00% (24/100)
		TSLGLLYTVKYPNLD	18	32	1.1723	27	26.00% (26/100)
		LGLLYTVKYPNLDDL	20	34	1.1312	27	28.00% (28/100)
SLGLLYTVKYPNLDD		19	33	1.1124	27	26.00% (26/100)	
LTSGLLYTVKYPNL		17	31	1.0659	27	74.00% (74/100)	
DAKLIADSLDFTQVS		1	15	0.9498	27	24.00% (24/100)	
GLLYTVKYPNLDDLE		21	35	0.9362	27	25.00% (25/100)	
ALTSLGLLYTVKYPN		16	30	0.8800	27	71.00% (71/100)	
FSLSAAVKAGASLID		1	15	0.8734	27	23.00% (23/100)	
LTLEHDCLQIITKDE		37	51	0.8608	27	22.00% (22/100)	

Epitope conservancy analysis

Putative epitopes generated from 4 diverse viral proteins were found to be highly conserved within different strains (Table 3). The top epitopes showing conservancy at a superior level (ranging from 40% to 100%) were allowed for further docking study and used to design the final vaccine constructs to ensure a broad spectrum vaccine efficacy (Table 4).

Table 4: Proposed CTL and HTL epitopes for vaccine construction

<i>Virus</i>	<i>Proteins Interacting</i>	<i>Top epitopes</i>	<i>Allergenicity</i>	<i>Toxicity</i>	<i>VaxiJen score</i>	<i>Conservancy (%)</i>
LASV	Nucleoprotein	IEGRPEDPV	Non-Allergen	Nontoxic	1.1593	100.00%
		NNQFGTMPSLTLACL	Non-Allergen	Nontoxic	1.0193	91.00%
LCMV	Nucleoprotein	LDFSEVSNV	Non-Allergen	Nontoxic	1.7221	91.00%
		ALTDLGLLYTVKYPN	Non-Allergen	Nontoxic	0.9756	40.00%
Lujo Virus	Nucleocapsid protein	KDSSLLNNQ	Non-Allergen	Nontoxic	0.8111	100.00%
		VWDVKDSSLLNNQFG	Non-Allergen	Nontoxic	1.1337	100.00%
Guanarito Virus	Nucleocapsid protein	GWPYIGSRS	Non-Allergen	Nontoxic	1.5322	99.00%
		LTSLGLLYTVKYPNL	Non-Allergen	Nontoxic	1.0659	74.00%

Designing three-dimensional (3D) epitope structure and molecular docking analysis

A total of 8 T-cell (4 CTL and 4 HTL) epitopes were subjected to PEP-FOLD server for 3D structure conversion and their interactions with HLA molecules were analyzed. HLA-A*11:01 (Class-I) and HLA-DRB1*01:01 (Class-II) were selected for docking analysis based on available PDB structures deposited in the database. Results confirmed that all the predicted epitopes bound in the groove of MHC molecules with a negative binding energy which were biologically significant (Table 5).

B-cell epitope prediction and screening

B-cell epitopes were predicted according to the analysis via three different algorithms (i.e. Bepipred Linear Epitope prediction, Emini Surface Accessibility, Kolaskar & Tongaonkar Antigenicity prediction) from IEDB. Top BCL epitopes for selected four proteins were further screened based

on their antigenicity scoring and allergenicity pattern which were used to design the final vaccine construct (Table 6).

Table 5: Binding energy of predicted epitopes and MHC alleles generated from molecular docking by AutoDock.

<i>CTL Epitopes</i>	<i>MHC allele</i>	<i>Binding energy (kcal/mol)</i>	<i>HTL Epitopes</i>	<i>MHC allele</i>	<i>Binding energy (kcal/mol)</i>
IEGRPDPV	HLA-A*11:01	-7.4	NNQFGTMPSLTLACL	HLA-DRB1*01:01	-6.6
LDFSEVSNV		-6.5	ALTDLGLLYTVKYPN		-7.9
KDSLLNNQ		-7.3	VWDVKDSSLLNNQFG		-7.2
GWPYIGSRS		-7.0	LTSLGLLYTVKYPNL		-8.7

Table 6: Allergenicity assessment antigenicity analysis of the predicted B-cell epitopes

<i>Virus</i>	<i>Proteins</i>	<i>Algorithm</i>	<i>Top CTL epitopes</i>	<i>Allergenicity</i>	<i>Vaxijen score</i>
<i>LASV</i>	Nucleoprotein	Linear Epitope	PNAKTWMDIEGRPDPVE	Non-allergen	0.6123
		Surface Accessibility	LMRKERRDDNDLKRLRDL	Non-allergen	0.9451
		Antigenicity	SLTLACLTKQ	Non-allergen	1.5660
<i>LCMV</i>	Nucleoprotein	Linear Epitope	SERPQASGV	Non-allergen	0.9344
		Surface Accessibility	IMRKEKRDDKDLQRLR	Non-allergen	0.9494
		Antigenicity	PPQVGLSYSQTM	Non-allergen	1.0820
<i>Lujó Virus</i>	Nucleocapsid protein	Linear Epitope	KVVKDAVSLINGLDFSMV	Non-allergen	0.5035
		Surface Accessibility	PNMDDLKDKLKNK	Non-allergen	0.7144
		Antigenicity	NLLYKICLSG	Non-allergen	0.7217
<i>Guanarito Virus</i>	Nucleocapsid protein	Linear Epitope	DERPGNRNPY	Non-allergen	0.9986
		Surface Accessibility	DERPGNRNPY	Non-allergen	0.9986
		Antigenicity	PGLLSYVIGLLPQGSVITVQ	Non-allergen	0.6398

Epitope cluster analysis and vaccine construction

Epitope cluster analysis tool from IEDB predicted the clusters among the top epitopes (4 CTL, 4 HTL and 12 BCL epitopes) proposed in Table 4 and Table 6. A total 19 clusters were identified which were utilized to design vaccine constructs in the study. Each construct designed were composed of a protein adjuvant followed by T-cell and B-cell epitopes with their respective linkers. PADRE sequence was incorporated to maximize the efficacy and potency of the peptide vaccine. A total 3 vaccine molecules (i.e. V1, V2, and V3) of 394, 479 and 508 amino acid residues were constructed (Table 7) and further investigated to evaluate their immunogenic potential.

Allergenicity, antigenicity and solubility prediction of different vaccine constructs

Results showed that construct V1 and V2 were non-allergic in behavior, while V3 exhibited allergenic pattern (Table 7). However, V1 was best in terms of safety and found superior as potential vaccine candidate with better antigenicity (0.703) and solubility score (0.65) (Fig. 3).

Physicochemical characterization and secondary structure analysis of vaccine protein

Vaccine construct V1 was characterized on the basis of physical and chemical properties. The computed instability index of the protein was 35.87 which classified it as a stable one. The theoretical pI 9.96 indicated that the protein will have net negative charge above the pI. Assuming all cysteine residues are reduced at 0.1% absorption, the extinction coefficient evaluated was 43890. The estimated half-life of the vaccine was expected to be 1 h in mammalian reticulocytes in vitro, while >10 h in *E. coli* and 30 minutes in yeast in vivo. Thermostability and hydrophilic nature of the vaccine protein was represented by aliphatic index (75.25) and GRAVY value (-0.625). Secondary structure of the construct V1 confirmed to have 26.14% alpha helix, 20.56% sheet and 53.30% coil structure (Fig. 4).

Table 7: Allergenicity, antigenicity and solubility prediction of the designed vaccine constructs

Constructs	Composition	Complete Sequence of Vaccine Constructs	Antigenicity (Threshold 0.4)	Allergenicity (Threshold -0.4)	Solubility (Threshold 0.45)
V1	Predicted CTL, HTL & BCL epitopes with β defensin adjuvant and PADRE sequence	EAAAKGIINTLQKYYCRVRRGRCVLSCLPKKEEQIGKCS TRGRKCCRRKKEAAAKAKFVAAWTLKAAAAGGGS ^L DF SEVSNVGGGSI ^E GRPEDPVGGGSKDSSLLNNQGGGSG WPIYIGSRSGPGPGAL ^T DLGLLYTVKYPNPGPGNNQF GTMPSLTLACL GPGPG VWDVKDSSLLNNQFGG PGPGL TSLGLLYTVKYPNL KK SERPQASGV KK IMRKEKRDDK DLQRLR KK PPQVGLSYSQTM KK PNAKTWMDIEGRPED PVE KK LMRKERRDDNDLKRRLDL KK SLTLACLTK QKK KVVKDAVSLINGLDFSMV KK PNMDDLKLNK KK NLL YKICLSG KK DERPGNRNPY KK PGLLSYVIGLLPQGSVIT V QKK AKFVAAWTLKAAAAGGGS	0.703	-0.45	0.65
V2	Predicted CTL, HTL & BCL epitopes with L7/L12 ribosomal protein adjuvant & PADRE sequence	EAAAKMAKLSTDELLDAFKEMTLELSDFVKKFEETFE VTAAAPVAVAAAGAAPAGAAVEAAEEQSEFDVILEAA GDKKIGVIKVVREIVSGLGLKEAKDLVDGAPKPLEKVA KEAADEAKAKLEAAGATVTVKEAAAKAKFVAAWTLK AAAAGGGS ^L DFSEVSNVGGGSI ^E GRPEDPVGGGSKDSSL LNNQGGGSGWPIYIGSRSGPGPGAL ^T DLGLLYTVKYPN GPGPG NNQFGT MP SLTLACL GPGPG VWDVKDSSLLNN QFGG PGPGL TSLGLLYTVKYPNL KK SERPQASGV KK IM RKEKRDDKDLQRLR KK PPQVGLSYSQTM KK PNAKTW MDIEGRPEDPVE KK LMRKERRDDNDLKRRLDL KK SLTL ACLTK QKK KVVKDAVSLINGLDFSMV KK PNMDDLKLN KNK KK NLLYKICLSG KK DERPGNRNPY KK PGLLSYVIG LLPQGSVITV QKK AKFVAAWTLKAAAAGGGS	0.630	0.010	0.67
V3	Predicted CTL, HTL & BCL epitopes with HABA adjuvant & PADRE sequence	EAAAKMAENPNIDDLPAPLLAALGAADLALATVNDLIA NLRERAEE ^T RAETRTRVEERRARLTKFQEDLPEQFIELR DKFTTEELRKA ^E GYLEAATNRYNELVERGEAALQRLR SQTAFEDASARAEGYVDQAVELTQEALGTVASQTRAVG ERA ^A KLVGIELEAAAKAKFVAAWTLKAAAAGGGS ^L DFS EVS ^N VGGGSI ^E GRPEDPVGGGSKDSSLLNNQGGGSGW PYIGSRSGPGPGAL ^T DLGLLYTVKYPNPGPGNNQFGT MPSLTLACL GPGPG VWDVKDSSLLNNQFGG PGPGL TS LGLLYTVKYPNL KK SERPQASGV KK IMRKEKRDDKDL QRLR KK PPQVGLSYSQTM KK PNAKTWMDIEGRPEDPV E KK LMRKERRDDNDLKRRLDL KK SLTLACLTK QKKK V VKDAVSLINGLDFSMV KK PNMDDLKLNK KK NLLYK ICLSG KK DERPGNRNPY KK PGLLSYVIGLLPQGSVITV Q KK AKFVAAWTLKAAAAGGGS	0.650	-1.004	0.63

Vaccine tertiary structure prediction, refinement, validation and disulfide engineering

I-TASSER generated 5 tertiary structures of the designed construct V1 using top 10 threading templates by LOMETS. TM score and RMSD were estimated based on C score which was minimum for Model 1 (-0.77), thus ensuring its better quality (Fig. 5A). The refined structure

was validated through Ramachandran plot analysis which revealed that 94.6% residues were in the allowed and 5.36% residues in the outlier region (Fig. 5B). Modeled tertiary structure of vaccine construct V2 and V3 have been shown in Fig. 6. A total 29 pairs of amino acid residue were identified having the capability to form disulfide bond by DbD2 server. However, after evaluation of the residue pairs in terms of energy, chi3 and B-factor parameter, only 2 pairs (ALA 24-CYS28 and GLU77-GLU87) satisfied the criteria for disulfide bond formation which were replaced with cysteine. The value of chi3 considered for the residue screening was between -87 to +97 while the energy value was less than 2.5.

Conformational B-cell and IFN- γ inducing epitopes prediction

A total of 5 conformational B-cell epitopes were predicted by using 3D structure of the proposed vaccine as an input. Epitopes No. 2 and 4 were considered as the broadest and smallest conformational B- cell epitopes with 85 and 13 amino acid residues (Table 8). Results also revealed that most of the residues, which were located in the multi-epitope region in our designed vaccine were included in the predicted conformational B-cell epitopes Moreover, the sequence of the final vaccine was applied for prediction of 15-mer IFN- γ inducing epitopes. Results showed that there were 100 positive IFN- γ inducing epitopes from which 8 had a score ≥ 1 (Supplementary file 6). Residues of 375-390 regions in the vaccine showed highest score of 1.209 (Table 9).

Protein-protein docking

The affinity between the constructed vaccines and HLA alleles were evaluated using molecular docking via Patchdock server. The server ranked the docked complexes based on Atomic Contact Energy (ACE), complementarity score and approximate interface area of the complex. Results revealed that construct V1 was superior in terms of free binding Energy (Table 10). Moreover, binding affinity of construct V1 with different human immune receptors was also demonstrated (Table 10). The lowest binding energy of the complexes indicated the highest binding affinity between receptor and vaccine construct.

Table 8: Conformational B-cell epitopes from 3D model of the vaccine construct V1

No.	Residues	No. of residues	Score
1	A:A66, A:A67, A:A68, A:G69, A:G70, A:G71, A:S72, A:L73, A:D74, A:F75, A:S76, A:E77, A:V78, A:S79, A:N80, A:V81, A:G82, A:I86, A:D276, A:L277, A:K278, A:R279, A:L280, A:R281	24	0.758
2	A:S98, A:K99, A:D100, A:S101, A:S102, A:L103, A:L104, A:N105, A:N106, A:Q107, A:G108, A:G109, A:G110, A:L127, A:T128, A:D129, A:L130, A:I307, A:N308, A:G309, A:L310, A:D311, A:F312, A:S313, A:M314, A:V315, A:K316, A:K317, A:P318, A:N319, A:M320, A:D321, A:D322, A:L323, A:D324, A:N328, A:K329, A:K330, A:K331, A:N332, A:L333, A:L334, A:S340, A:G341, A:K342, A:K343, A:D344, A:E345, A:R346, A:P347, A:G348, A:N349, A:R350, A:N351, A:P352, A:Y353, A:K354, A:K355, A:L366, A:P367, A:Q368, A:G369, A:S370, A:V371, A:I372, A:T373, A:V374, A:K377, A:A378, A:K379, A:F380, A:V381, A:A382, A:A383, A:W384, A:T385, A:L386, A:K387, A:A388, A:A389, A:A390, A:G391, A:G392, A:G393, A:S394	85	0.716
3	A:E1, A:A2, A:A3, A:A4, A:K5, A:G6, A:I8, A:N9, A:T10, A:K13, A:G163, A:P164, A:G165, A:V166, A:W167, A:D168, A:V169, A:K170, A:D171, A:S172, A:S173, A:L174, A:L175, A:N176, A:N177, A:Q178, A:F179, A:G180, A:G181, A:P184, A:T187, A:S188, A:G190, A:L191, A:L192, A:T194, A:V195, A:K196, A:Y197, A:P198, A:N199, A:L200, A:K201, A:K202, A:S203, A:E204, A:P206, A:Q207, A:A208, A:S209, A:G210, A:V211, A:K212, A:K213, A:I214, A:M215, A:R216, A:K217, A:E218, A:K219, A:R220, A:D221, A:D222, A:K223, A:Q226, A:K230, A:T250, A:W251, A:D253, A:I254, A:E255, A:G256, A:R257, A:P258, A:E259	75	0.671
4	A:D282, A:L283, A:K284, A:K285, A:S286, A:L287, A:T288, A:L289, A:A290, A:C291, A:L292, A:T293, A:G357	13	0.591
5	A:C28, A:K31, A:E32, A:E33, A:Q34, A:I35, A:G36, A:K37, A:C38, A:T40, A:R41, A:K44, A:C45, A:R47	14	0.56

Molecular dynamics simulation

The high ranked complex between vaccine molecules and TLR-3 were selected for analysis through Normal mode analysis (NMA). NMA was performed to describe the stability of proteins and large scale mobility. Results showed that the mobility of vaccine protein V1 and α -

dystroglycan were oriented towards each other. Probable deformability of the complex was indicated by hinges in the chain due to distortion of the individual residues (Fig. 7A). The B-factor values which was equivalent to RMS inferred via NMA (Fig. 7B). Eigenvalue for the complex demonstrated was $2.2023e-06$ (Fig. 7C). Colored bars showed the individual (red) and cumulative (green) variances which were inversely related to eigenvalue (Fig. 7D). The indicated coupling between Different interactions between residues were revealed by covariance matrix i.e. correlated, uncorrelated and anti-correlated motions indicated by red, white and blue colors respectively (Fig. 7E). The elastic network model (Fig. 7F). It identified the pairs of atoms those were connected via springs.

Table 9: Predicted IFN- γ inducing epitopes from the proposed vaccine

Start-End	Epitopes	Methods	Prediction	Score
375-390	KKAKFVAAWTLKAAA	SVM	POSITIVE	1.20947
48-63	KKEAAAKAKFVAAWT	SVM	POSITIVE	1.188251
50-65	EAAAKAKFVAAWTLK	SVM	POSITIVE	1.12401
53-68	AKAKFVAAWTLKAAA	SVM	POSITIVE	1.111491
106-121	QGGGSGWPYIGSRSG	MERCI	POSITIVE	1.00
316-331	KPNMDDLKLNKKNK	MERCI	POSITIVE	1.00
321-336	DLDKLNKKNKKNLLYK	MERCI	POSITIVE	1.00
368-383	GSVITVQKKAKFVAA	MERCI	POSITIVE	1.00
370-385	VITVQKKAKFVAAWT	MERCI	POSITIVE	1.00
374-389	QKKAKFVAAWTLKAA	MERCI	POSITIVE	1.00

Codon adaptation and in silico cloning

The Codon Adaptation Index for the optimized codons of construct V1 was 0.969 determined via JCAT server. The GC content of the adapted codons was also significant (50.93%). An insert of 1194 bp was obtained which lacked restriction sites for BglII and BglI ensuring, thus ensuring safety for cloning purpose. The codons were inserted into pET28a(+) vector along with BglII and BglI restriction sites and a clone of 4777 base pair was produced (Fig. 8).

Table 10: Docking scores of vaccine constructs with different HLA alleles i.e. HLA-DRB1*03:01 (1a6a), HLA-DRB5*01:01 (1h15), HLA-DRB1*01:01 (2fse), HLA-DRB3*01:01 (2q6w), HLA-DRB1*04:01 (2seb) and HLA-DRB3*02:02 (3c5j) and receptors i.e. TLR 3 (2a0z), MDA 5 (3b6e), α dystroglycan (5llk), RIG-1 (2qfd)

<i>Vaccine construct</i>	<i>PDB IDs of HLAs/receptors</i>	<i>Global Energy</i>	<i>Hydrogen bond energy</i>	<i>ACE</i>	<i>Score</i>	<i>Area</i>
V1	1a6a	-2.65	-3.20	5.18	16280	2111.50
	1h15	-6.40	-2.75	7.08	17246	2552.90
	2fse	-26.12	-2.94	0.28	16062	2181.20
	2q6w	-2.59	-4.48	9.33	15968	2084.90
	2seb	-28.54	-3.71	5.51	17904	2308.00
	3c5j	-28.29	-5.12	9.40	16956	2397.50
V2	1a6a	-0.55	-2.61	12.31	16394	2349.80
	1h15	-0.55	-4.21	10.79	17406	2402.80
	2fse	12.04	-5.21	12.26	18306	2646.60
	2q6w	-0.33	-3.06	6.47	18262	2940.70
	2seb	-5.52	-0.82	-1.77	21390	3942.60
	3c5j	-22.27	-3.87	9.04	16846	2299.50
V3	1a6a	-0.23	-4.64	10.85	17830	2324.60
	1h15	-16.48	-3.48	15.41	19260	2658.10
	2fse	-9.38	-1.68	12.37	19720	2651.90
	2q6w	-23.12	-4.09	10.44	19724	2701.10
	2seb	-19.67	-0.52	-3.87	18596	3097.60
	3c5j	1.78	-5.71	12.71	18192	2430.00
V1	2a0z	-10.82	-3.34	2.33	14764	2172.70
	3b6e	-3.91	-3.33	16.69	15762	1889.60
	5llk	-17.17	-6.88	4.72	15410	1906.70
	2qfd	-0.89	-3.03	9.80	14636	1953.30

Discussion

Arenaviruses exhibit a catastrophic potential to destroy the public health scenario in several regions of the world. Few members of Arenaviridae family pose a credible bioterrorism threat (Hauer et al., 2002), while six of them, including LASV and LCMV are classified as Category A agents by the National Institute of Allergy and Infectious Diseases (Borio et al., 2002; Charrel and de Lamballerie, 2003). Despite the significance of arenaviruses in public health and biodefense readiness, to date there are no vaccines approved by the Food and Drug

Administration (FDA) (Brisse and Ly, 2019; Buchmeier et al 2007; Cheng et al., 2015). Current anti-arenavirus therapy is partially effective and requires an early and intravenous administration of nucleoside analog ribavirin that may cause significant side effects (Bausch et al., 2010; Moreno et al., 2011). Therefore, there is an unmet need to develop both safe and protective vaccines to combat pathogenic arenavirus infections in humans.

The live attenuated Candid#1 strain of Junin virus has been shown to be an effective vaccine against Argentine Hemorrhagic Fever (Ambrosio et al., 2011; Maiztegui et a., 1998). Researchers also found that Junin Virus Vaccine Antibodies are effective against Machupo Virus as well (Clark et al., 2011). However, the preventive measures to combat infections caused by LASV, LCMV, Lujo and Machupo virus has not yet gained considerable success. Although, ML29, a live attenuated vaccine has been shown to provide effective protection *in vivo* against Lassa virus (Carrion et al., 2007; Lukashevich et al., 2008), the mechanism for ML29 attenuation remains unknown (Olschlager and Flatz, 2013). Therefore, the incorporation of a limited number of additional mutations into the ML29 genome may result in viruses with enhanced virulence (Greenbaum et al., 2012). In the present study attempts were taken to develop a vaccine candidate using the most antigenic viral proteins of Arenaviridae family that could elicit broad spectrum immunity in the host. To the best of our knowledge, similar genome based screening and reverse vaccinology approach have not yet been explored to design an arenavirus vaccine.

Nucleoprotein (NP) specific CD8⁺ T cells play a major role in virus control and immune stimulation in the host (Schildknecht et al., 2008). Meulen and coworkers (2000) found that Lassa fever survivors had strong CD4⁺ T cell responses against LASV Nucleoprotein (ter Meulen et al., 2000). In another study, a single inoculation of a plasmid encoding full-length Lassa nucleoprotein induced CD8⁽⁺⁾ T cell responses in mice model and protected against LCMV (Rodriguez-Carreno et al., 2005). All these findings suggest that anti-NP response at an early stage effectively controls and contributes to cross-protective immunity against arenavirus infections. In the present study, the superiority of Nucleoproteins of LASV, LCMV, Lujo and Guanarito virus in terms of antigenicity score were also revealed (Supplementary File 1). Hence, a multivalent vaccine strategy was implemented to protect against nucleoprotein antigens from different arenavirus species. To ensure protective response against a longer range of virus strains for a longer period, the candidate epitopes must remain in the highly conserved region. Thus,

only the conserved sequences from each of the viral proteins (Table 2) were used for epitope prediction, which we believe, would generate more acceptable and universally effective vaccine constructs. Determination of potent T cell epitopes is a pivotal step during vaccine design, since T-cells play the key role in antibody production through antigen presentation (Amorim et al., 2016). Both MHC-I (CTL) and MHC-II restricted (HTL) epitopes were predicted using IEDB T cell epitope prediction tools and top candidates were screened based on transmembrane topology screening, antigenicity scoring, allergenicity pattern, toxicity profile and conservancy analysis (Table 3). The proposed epitopes showed a high cumulative population coverage in most of the geographic regions including 88% and 86% MHC-I coverage in South America and Africa respectively (Fig. 2). The top epitopes were highly conserved among different viral strains ranging from 40% to 100% (Table 4). In the next step, the predicted CTL and HTL epitopes were checked for their ability to bind with MHC-I and MHC-II alleles. Results showed that all the predicted epitopes were efficient binders of corresponding MHC allele (Table 5). B-cell epitopes boost neutralizing-antibody responses in the host. To achieve protective immunity against arenavirus infection, B-cell (BCL) epitopes were identified using three different algorithms (i.e. Linear Epitope, Surface Accessibility, Antigenicity) from IEDB (Table 6). Moreover, our results revealed that among the top epitopes, T cell epitope GWPYIGSRS were conserved in Argentine mammarenavirus (Junin virus) and Brazilian mammarenavirus (Sabia virus), while B cell epitope NLLYKICLSG were conserved in Bolivian mammarenavirus (Machupo virus) and Brazilian mammarenavirus (Sabia virus), indicating the possibility of final vaccine constructs to confer broad range immunity in the host against the arenaviruses.

The screened CTL, HTL and BCL epitopes were combined using respective linker to design the final vaccine constructs. Three different constructs (i.e. V1, V2, V3) were generated using 3 distinct linkers and investigated in terms of safety and efficacy (Table 7). Vaccine protein V1 was identified superior considering the immunogenicity score, allergenicity, hydrophobicity and other physicochemical properties. To strengthen our prediction, all three constructs were subjected to 3D modeling and interactions between designed vaccine constructs with different HLA molecules (i.e. HLA-DRB1*03:01, HLA-DRB5*01:01, HLA-DRB1*01:01, HLA-DRB3*01:01, HLA-DRB1*04:01 and HLA-DRB3*02:02) were determined (Table 7). Again construct V1 was found to be best in terms of free binding energy. IFN- γ is the signature cytokine of both the innate and adaptive immune systems with ability to provok antiviral

immune responses and protection against reinfection. The release of IFN- γ enhances the magnitude of antiviral cytotoxic T lymphocytes (CTLs) responses and aids in production of neutralizing IgG (Nosrati et al., 2019). Construct V1 was screened to identify such IFN- γ producing epitopes which showed positive results (Table 9). Moreover, docking analysis was also performed to explore the binding affinity of construct V1 and different human immune-receptors (TLR 3, MDA 5, α dystroglycan, RIG-1) to evaluate the efficacy of used adjuvants (Table 10). α -dystroglycan (α -DG), a peripheral membrane protein, acts as an anchor between the submembranous cytoskeleton and the extracellular matrix which is widely expressed in most cells (Durbeej et al., 1998). It has been identified as a cellular receptor for LASV, certain strains of LCMV and the New World arenaviruses (Kunz et al., 2005, Spiropoulou et al., 2002). Molecular dynamics study by iMOD server ensured the stability of V1-TLR3 complex at molecular level. Finally, the designed construct V1 was reverse transcribed and inserted within pET28a(+) vector for heterologous expression in *E. coli* strain K12.

The predicted results were based on different sequence analysis and various immune databases. Due to the encouraging findings of the study, we suggest further wet lab based analysis using model animals for the experimental validation of our predicted vaccine candidates. Moreover, novel antigen delivery systems such as Nano-delivery platforms could enhance the efficacy of the proposed vaccine (Hojo 2014; Trovato and Berardinis 2015).

Conclusion

In-silico bioinformatics study could be considered as a promising strategy to accelerate vaccine development against highly pathogenic organisms. In the present study, such approach was employed to design a novel heterosubtypic peptide vaccine against the most deadly viruses of Arenaviridae family requiring urgent need for effective medications and preventive measures. The study suggests, the proposed vaccine could stimulate both humoral and cellular mediated immune responses and serve as a potential vaccine against arenaviruses. However, *in vitro* and *in vivo* immunological experiments are highly recommended to validate the efficacy of designed vaccine constructs.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The Authors declare that they have no conflicts of interest.

Acknowledgments

Authors would like to acknowledge the Department of Microbial Biotechnology and Department of Pharmaceuticals and Industrial Biotechnology at Sylhet Agricultural University for the technical support of the project.

References

- Aalten DMF, Groot BL, Findlay JBC, Berendsen HJC, Amadei A (1997) A Comparison of Techniques for Calculating Protein Essential Dynamics. *Journal of Computational Chemistry* 18(2):169-181
- Ambrosio A, Saavedra M, Mariani M. Argentine hemorrhagic fever vaccines. *Human Vaccin.* 2011;7(6):694–700.
- Amorim KN, Rampazo EV, Antonialli R, Yamamoto MM, Rodrigues MM, Soares IS, et al. The presence of T cell epitopes is important for induction of antibody responses against antigens directed to DEC205+ dendritic cells. *Scientific reports.* 2016;6:39250.
- Awan FM, Obaid A, Ikram A, Janjua HA (2017) Mutation-structure function relationship based integrated strategy reveals the potential impact of deleterious missense mutations in autophagy related proteins on hepatocellular carcinoma (HCC): a comprehensive informatics approach. *Int J Mol Sci* 18(1):139

Azim KF, Hasan M, Hossain MN, Somana SR, Hoque SF, Bappy MN, Chowdhury AT, Lasker T. Immunoinformatics approaches for designing a novel multi epitope peptide vaccine against human norovirus (Norwalk virus). *Infection, Genetics and Evolution*. 2019 Jun 21:103936.

Bausch DG, Hadi CM, Khan SH, Lertora JJ. 2010. Review of the literature and proposed guidelines for the use of oral ribavirin as postexposure prophylaxis for Lassa fever. *Clin Infect Dis* 51:1435–1441.

Bonthius DJ, Wright R, Tseng B, Barton L, Marco E, Karacay B, Larsen PD. Congenital lymphocytic choriomeningitis virus infection: spectrum of disease. *Ann Neurol*. 2007;62:347–355.

Borio L, Inglesby T, Peters CJ, Schmaljohn AL, Hughes JM, Jahrling PB, Ksiazek T, Johnson KM, Meyerhoff A, O'Toole T, Ascher MS, Bartlett J, Breman JG, Eitzen EM Jr, Hamburg M,

Borrow, P., Martinez-Sobrido, L., delaTorre, J.C.,2010 .Inhibition of the type I interferon antiviral response during arenavirus infection. *Viruses* 2, 2443–2480.

Brisse ME, Ly H. Hemorrhagic Fever-Causing Arenaviruses: Lethal Pathogens and Potent Immune Suppressors. *Frontiers in immunology*. 2019;10:372.

Buchmeier MJ, de la Torre JC, Peters CJ. 2007. Arenaviridae: the viruses and their replication, p 1791–1827. In Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE. (ed), *Fields virology*, 5th ed, vol II Lippincott Williams & Wilkins, Philadelphia, PA, USA.

Carrion R, Jr, Patterson JL, Johnson C, Gonzales M, Moreira CR, Ticer A, Brasky K, Hubbard GB, Moshkoff D, Zapata J, Salvato MS, Lukashevich IS. A ML29 reassortant virus protects guinea pigs against a distantly related Nigerian strain of Lassa virus and can provide sterilizing immunity. *Vaccine*. 2007;25:4093–4102

Charrel RN, de Lamballerie X. 2003. Arenaviruses other than Lassa virus. *Antiviral Res* 57:89–100.

Cheng BY, Ortiz-Riaño E, de la Torre JC, Martínez-Sobrido L. Arenavirus genome rearrangement for the development of live attenuated vaccines. *Journal of virology*. 2015 Jul 15;89(14):7373-84.

Clark LE, Mahmutovic S, Raymond DD, Dilanyan T, Koma T, Manning JT, Shankar S, Levis SC, Briggiler AM, Enria DA, Wucherpfennig KW. Vaccine-elicited receptor-binding site antibodies neutralize two New World hemorrhagic fever arenaviruses. *Nature communications*. 2018 May 14;9(1):1884.

Craig DB, Dombkowski AA (2013) Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinform* 14:346

Craighead, J. E. (2000). *Pathology and pathogenesis of human viral disease*. Academic Press.

Daniel J. Bonthius ; Lymphocytic Choriomeningitis Virus: An Underrecognized Cause of Neurologic Disease in the Fetus, Child, and Adult; *Seminars in Pediatric Neurology* Volume 19, Issue 3, September 2012, Pages 89-95

Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic acids research*. 2004 Jul 1;32(suppl_2):W615-9.

Dhanda SK, Vaughan K, Schulten V, Grifoni A, Weiskopf D, Sidney J, Peters B, Sette A. Development of a novel clustering tool for linear peptide sequences. *Immunology*. 2018 Nov;155(3):331-45.

Dimitrov I, Bangov I, Flower DR, Doytchinova IA (2013) AllerTOP v.2- a server for in silico prediction of allergens. *J of Molecular Modeling* 20: 2278

Dimitrov I, Naneva L, Doytchinova IA, Bangov I (2014) AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* 30: 846–851

Doytchinova IA, Darren R Flower (2007) Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine* 25:856-866

Durbeej M, Henry MD, Campbell KP. Dystroglycan in development and disease *Curr. Opin. Cell Biol.* 10 (1998) 594-601

Emini EA, Hughes JV, Perlow D, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of virology* 55:836-9

Fiers M WEJ, Kleter GA, Nijland H, Peijnenburg A, Nap JP, Ham R (2004) Allermatch™, a webtool for the prediction of potential allergenicity according to current FAO/WHO Codex alimentarius guidelines BMC Bioinformatics 5:133

Fulhorst, C. F., Bowen, M. D., Salas, R. A., Duno, G., Utrera, A., Ksiazek, T. G. & Tesh, R. B. (1999). Natural rodent host associations of Guanarito and pirital viruses (Family Arenaviridae) in central Venezuela. *The American journal of tropical medicine and hygiene*, 61(2), 325-330.

Gasteiger E, Hoogland C, GattikerA, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2003) Protein Identification and Analysis Tools on the ExPASy Server. *Nucleic Acids Res* 31:3784-3788

Goodman RE, Ebisawa M, Ferreira F, Sampson HA, van Ree R, Vieths S, Baumert JL, Bohle B, Lalithambika S, Wise J, Taylor SL. AllergenOnline: a peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Molecular nutrition & food research*. 2016 May;60(5):1183-98.

Greenbaum BD, Li OT, Poon LL, Levine AJ, Rabadan R. Viral reassortment as an information exchange between Viral segments. *Proc Natl Acad Sci USA*. 2012;109:3341–3346.

Grote A et al (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host *Nucleic acids research* 33:W526–531

Hajighahramani, N., Nezafat, N., Eslami, M., Negahdaripour, M., Rahmatabadi, S.S. and Ghasemi, Y., 2017. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against *Staphylococcus aureus*. *Infection, Genetics and Evolution*, 48, pp.83-94.

Hasan M, Ghosh P, Azim K, Mukta S, Abir RA, Nahar J, Khan MM (2019b) Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus. *Microbial pathogenesis* 130:19-37

Hasan M, Hakim A, Iqbal A, Bhuiyan FR, Begum MK, Sharmin S and Abir RA. Computational study and homology modeling of phenol hydroxylase: key enzyme for phenol degradation. *Int J Comput Bioinfo In Silico Model*. 2015b;4(4):691-69

Hasan M, Hakim A, Iqbal A, Bhuiyan FR, Begum MK, Sharmin S and Abir RA (2015b) Computational study and homology modeling of phenol hydroxylase: key enzyme for phenol degradation. *Int J Comput Bioinfo In Silico Model* 4(4):691-698

Hasan M, Joy ZF, Bhuiyan EH, Islam MS. In Silico Characterization and Motif Election of Neurotoxins from Snake Venom. *American Journal of Biochemistry & Biotechnology*. 2015a;11(2):84.

Hasan, M., Islam, M.S., Chakraborty, S., Mustafa, A.H., Azim, K.F., Joy, Z.F., Hossain, M.N., Foysal, S.H. and Hasan, M.N., 2019. Contriving a chimeric polyvalent vaccine to prevent infections caused by Herpes Simplex Virus (Type-1 and Type-2): an exploratory immunoinformatic approach. *Journal of biomolecular Structure and Dynamics* 12:1-18

Hasson SS, Al-Busaidi JK, Sallam TA (2015) The past, current and future trends in DNA vaccine immunisations. *Asian Pacific Journal of Tropical Biomedicine* 5(5):344-53

Hauer J, Henderson DA, Johnson RT, Kwik G, Layton M, Lillibridge S, Nabel GJ, Osterholm MT, Perl TM, Russell P, Tonat K. 2002. Hemorrhagic fever viruses as biological weapons: medical and public health management. *JAMA* 287:2391–2405

Hojo H. Recent progress in the chemical synthesis of proteins. *Current opinion in structural biology*. 2014;26:16-23.

Houlihan, Catherine; Behrens, Ron (12 July 2017). "Lassa fever". *BMJ*:j2986.

Ibraghimov-Beskrovnyaya, O.,Milatovich, A.,Ozcelik, T.,Yang,B., Koepnick,K., Francke,U., Campbell,K.P., 1993. Human dystroglycan: skeletal muscle cDNA, genomic structure,origin of tissue specific isoforms and chromosomal localization. *Hum.Mol.Genet.*2,1651–1657.

Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic acids research* 2017; 45: W24-W9

Kaufmann SH, McElrath MJ, Lewis DJ, Del Giudice G (2014) Challenges and responses in human vaccine development. *Current opinion in immunology* 28:18-26

Kolaskar A, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS letters*; 276: 172-4

Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. *Journal of molecular biology* 305:567-80

Kunz S, Rojek JM, Kanagawa M, Spiropoulou CF, Barresi R, Campbell KP, Oldstone MB. Posttranslational modification of alpha-dystroglycan, the cellular receptor for arenaviruses, by the glycosyltransferase LARGE is critical for virus binding. *J. Virol.* 79 (2005) 14282-14296

Larsen PD, Wright R. Early clinical manifestations and long-term outcome in children with symptomatic congenital lymphocytic choriomeningitis virus infection. *Neurology.* 2001;56:A39–A40

Lopez-Blanco JR, Aliaga JI, Quintana-Orti ES, Chacon P (2014) iMODS: internal coordinates normal mode analysis server. *Nucleic Acids Res* 42:W271–W276

Lukashevich IS, Carrion R, Jr, Salvato MS, Mansfield K, Brasky K, Zapata J, Cairo C, Goicochea M, Hoosien GE, Ticer A, Bryant J, Davis H, Hammamieh R, Mayda M, Jett M, Patterson J. Safety, immunogenicity, and efficacy of the ML29 reassortant vaccine for Lassa fever in small non-human primates. *Vaccine.* 2008;26:5246–5254

M. Hebditch, M.A. Carballo-Amador, S. Charonis, R. Curtis, J. Warwicker, Protein-Sol: a web tool for predicting protein solubility from sequence, *Bioinformatics* 33 (2017) 3098–3100.

Maiztegui JI, McKee KT Jr, Barrera Oro JG, Harrison LH, Gibbs PH, Feuillade MR, Enria DA, Briggiler AM, Levis SC, Ambrosio AM, Halsey NA, Peters CJ. 1998. Protective efficacy of a live attenuated vaccine against Argentine hemorrhagic fever. AHF Study Group. *J Infect Dis* 177:277–283.

Manziona ND, Salas RA, Paredes H, Godoy O, Rojas L, Araoz F., Peters, C. J. (1998). Venezuelan hemorrhagic fever: clinical and epidemiological studies of 165 cases. *Clinical infectious diseases*, 26(2), 308-313.

Maupetit J, Derreumaux P, Tuffery P (2010) A fast method for large scale De Novo peptide and miniprotein structure prediction. *Journal of computational chemistry* 31:726-38

Mendenhall M, Russell A, Juelich T, Messina EL, Smee DF, Freiberg AN, et al. T-705 (Favipiravir) inhibition of arenavirus replication in cell culture. *Antimicrob Agents Chemother*. 2011;55:782–787.

Moreno H, Gallego I, Sevilla N, de la Torre JC, Domingo E, Martin V. 2011. Ribavirin can be mutagenic for arenaviruses. *J Virol* 85:7246–7255.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS et al (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30:2785-91.

Nadezhda E. Yun and David H. Walker. (2012). "Pathogenesis of Lassa Fever". *Viruses*. 4 (12): 2031–2048.

Nosrati M, Behbahani M, Mohabatkar H. Towards the first multi-epitope recombinant vaccine against Crimean-Congo hemorrhagic fever virus: A computer-aided vaccine design approach. *Journal of biomedical informatics*. 2019 May 1;93:103160.

Ogbu O, Ajuluchukwu E, Uneke CJ (2007). "Lassa fever in West African sub-region: an overview". *Journal of Vector Borne Diseases*. 44 (1): 1–11.

Olschlager S, Flatz L. Vaccination strategies against highly pathogenic arenaviruses: the next steps toward clinical trials. *PLoS Pathog*. 2013;9:e1003212

Paweska JT. Lujo virus hemorrhagic fever. In: *Clinical Case Study of Emerging Infectious Diseases*. O Ergonul, F Can, M Akova, L Madoff, editors. Elsevier/Academic Press London. 2014:95–110.

Paweska, J. et al. (2009). "Nosocomial outbreak of novel arenavirus infection: Southern Africa". *Emerg Infect Dis*. 15 (10): 1598–1602.

Ponomarenko JV, Bui H, Li W, Fusseder N, Bourne PE, Sette A, Peters B. 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* 9:514.

Rana A, Akhter Y (2016) A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. *Immunobiology* 221:544–557

Rodriguez-Carreno MP, Nelson MS, Botten J, Smith-Nixon K, Buchmeier MJ, Whitton JL. Evaluating the immunogenicity and protective efficacy of a DNA vaccine encoding Lassa virus nucleoprotein. *Virology*. 2005 Apr 25;335(1):87-98.

Saha S, Raghava GP (2006) AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research* 34(suppl_2):W202-9

Salas R, Pacheco ME, Ramos B, Taibo ME, Jaimes E, Vasquez C, Araoz F. (1991). Venezuelan haemorrhagic fever. *The Lancet*, 338(8774), 1033-1036.

Salvato MS, editor. *The arenaviridae*. Springer Science & Business Media. 2012 Dec 6.

Schildknecht A, Welti S, Geuking MB, et al. Absence of CTL responses to early viral antigens facilitates viral persistence. *J Immunol*. 2008;180(5):3113–21.

Sewlall NH, Paweska JT; Lujó virus: current concepts; Dovepress 2017;9:41-47.

Shao J, Liang Y, Ly H. Human hemorrhagic fever causing arenaviruses: molecular mechanisms contributing to virus virulence and disease pathogenesis. *Pathogens*. 2015 Jun;4(2):283-306.

Shoemaker T, Choi M. CDC yellow book 2020, Health Information for international travel; Viral Hemorrhagic Fevers, travel-related-infectious-diseases (Chapter 4); last updated 2019.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*. 2011 Jan 1;7(1).

Solanki V, Tiwari V (2018) Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against *Acinetobacterbaumannii*. *Scientific reports* 8(1):9044

Spiropoulou CF, Kunz S, Rollin PE, Campbell KP. Old stone New World arenaviruses clade C, but not clade A and B viruses, utilizes a-dystroglycan as its major receptor *J. Virol.* 76 (2002) 5140-5146

Stratton K, Almario DA, Wizemann TM, McCormick MC (2003) Immunization safety review: vaccinations and sudden unexpected death in infancy. Institute of Medicine (US) Immunization Safety Review Committee, National Academies Press (US)

Tama F, Brooks III CL. Symmetry, form, and shape: guiding principles for robustness in macromolecular machines. *Annu. Rev. Biophys. Biomol. Struct.*. 2006 Jun 9;35:115-33.

ter Meulen J, Badusche M, Kuhnt K, et al. : Characterization of human CD4 + T-Cell clones recognizing conserved and variable epitopes of the lassa virus nucleoprotein. *J Virol.* 2000;74(5):2186–92.

Ter Meulen J, Lukashevich I, Sidibe K, Inapogui A, Marx, M, Dorlemann A, Schmitz H. (1996). Hunting of peridomestic rodents and consumption of their meat as possible risk factors for rodent-to-human transmission of Lassa virus in the Republic of Guinea. *The American journal of tropical medicine and hygiene*, 55(6), 661-666.

Tesh RB, Jahrling PB, Salas R, Shope RE. (1994). Description of Guanarito virus (Arenaviridae: Arenavirus), the etiologic agent of Venezuelan hemorrhagic fever. *The American journal of tropical medicine and hygiene*, 50(4), 452-459.

Trovato M, De Berardinis P. Novel antigen delivery systems. *World journal of virology.* 2015;4:156.

Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR et al (2014) The immune epitope database (IEDB) 3.0. *Nucleic acids research* 43:D405-D12

Wright R, Johnson D, Neumann M, Ksiazek TG, Rollin P, Keech RV, Bonthius DJ, et al. Congenital lymphocytic choriomeningitis virus syndrome: A disease that mimics congenital toxoplasmosis or cytomegalovirus infection. *Pediatrics.* 1997;100:1–6.

Xu D, Zhang Y (2011) Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-step Atomic-level Energy Minimization. *Biophysical Journal* 101:2525-2534

Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC bioinformatics.* 2008;9(1):40.

Figure Legends

Fig. 1: Schematic presentation of the procedures used for multi-epitope vaccine development against Arenaviridae family.

Fig. 2: Population coverage analysis of predicted T-cell epitopes (MHC-I and MHC-II peptides).

Fig. 3: Solubility prediction of designed vaccine protein V1 using via Protein-sol server.

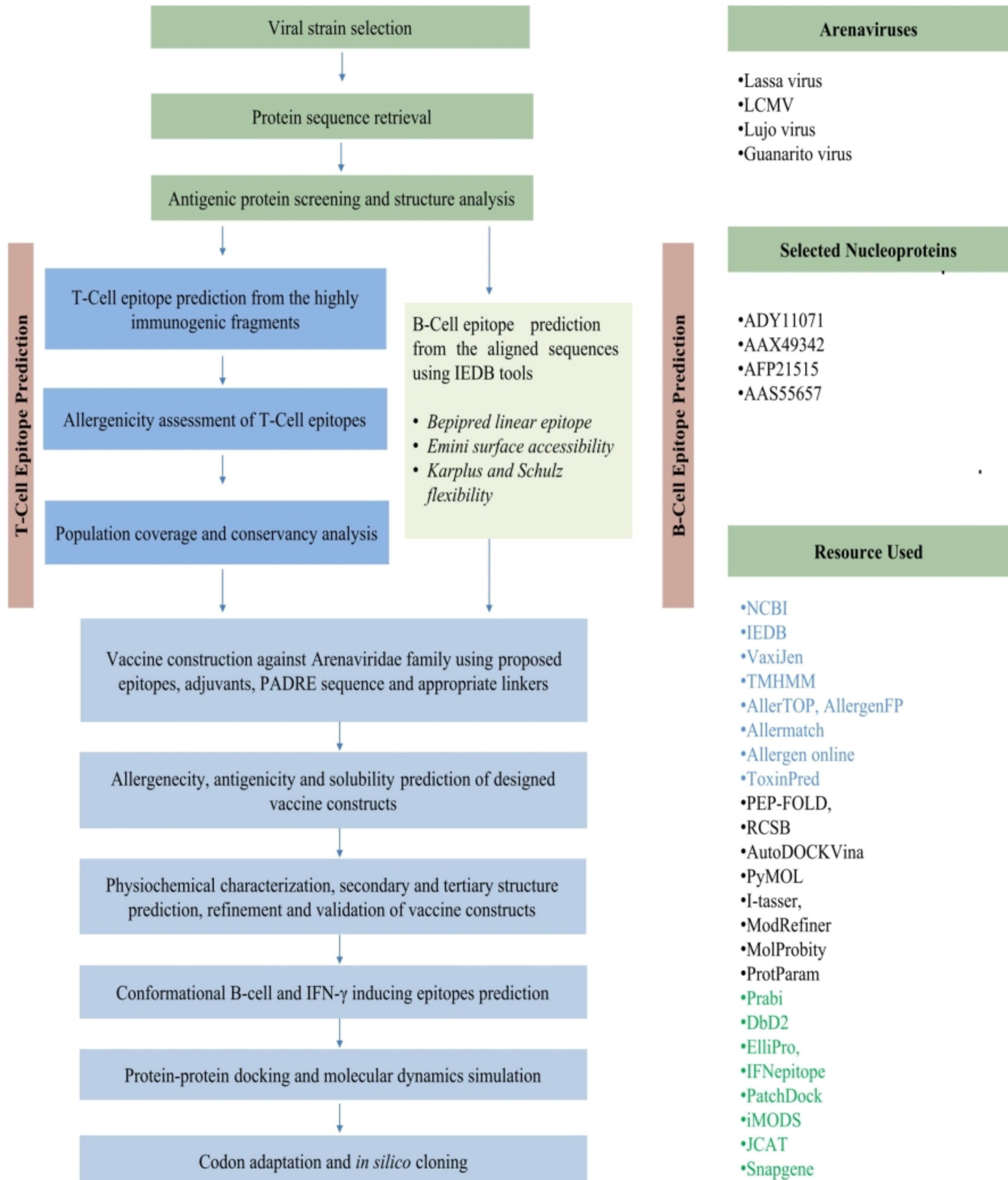
Fig. 4: Secondary structure analysis of the designed construct V1.

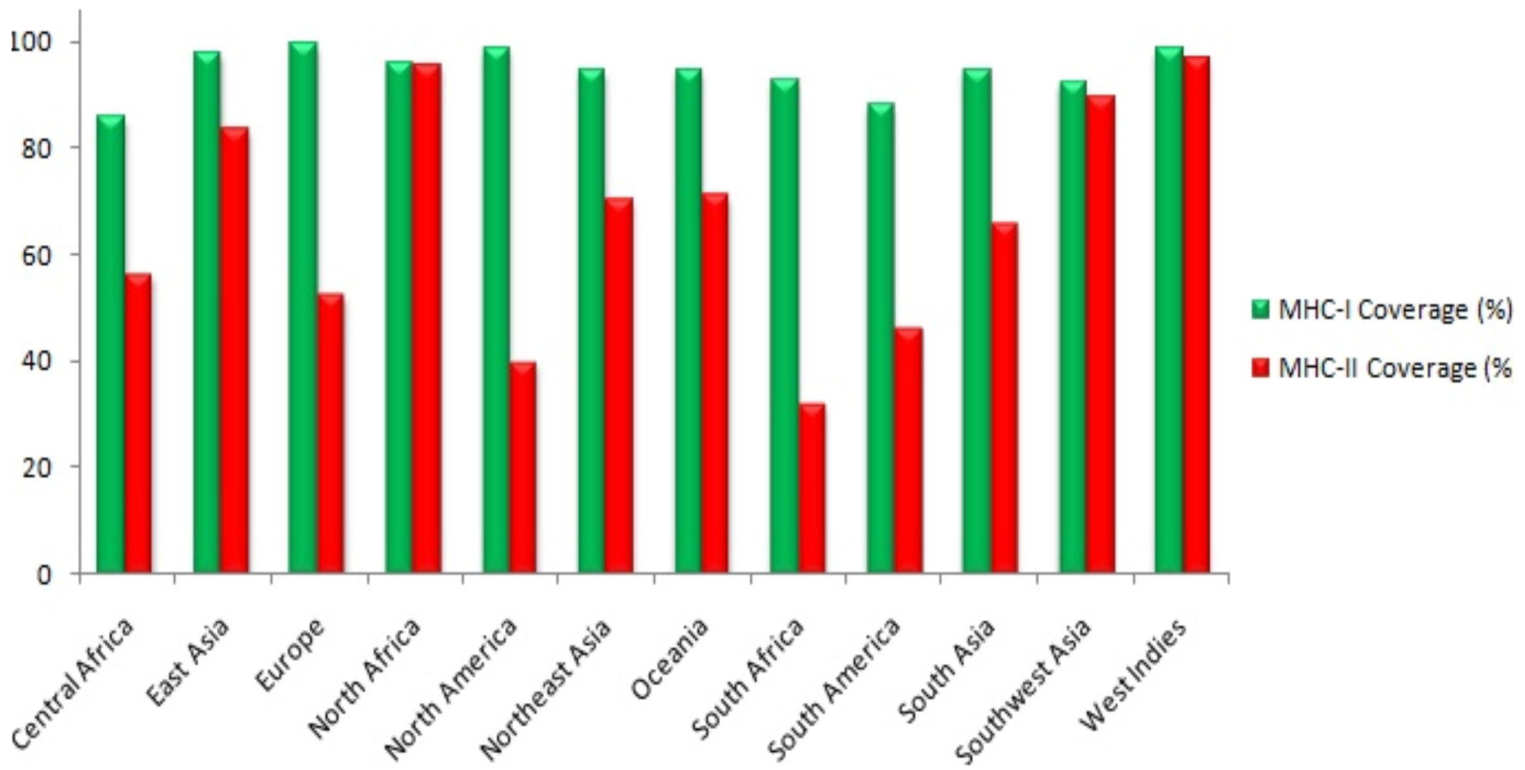
Fig. 5: Homology modeling of vaccine protein V1 via I-TASSER (A) and validation of the 3D smodel via by Ramachandran plot analysis (B).

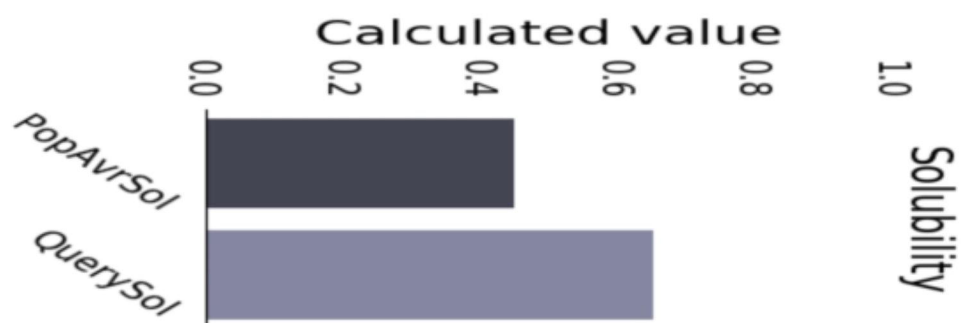
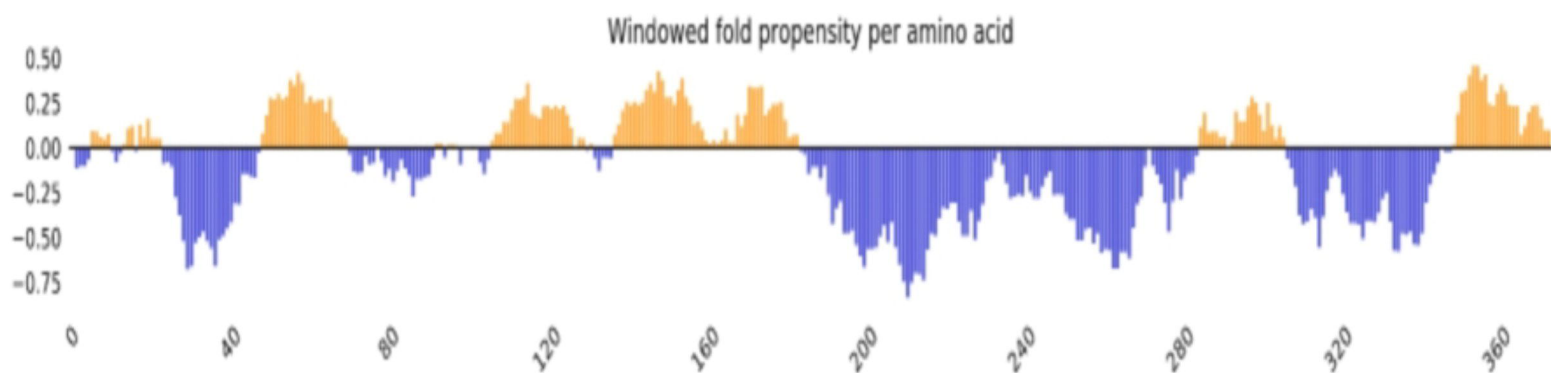
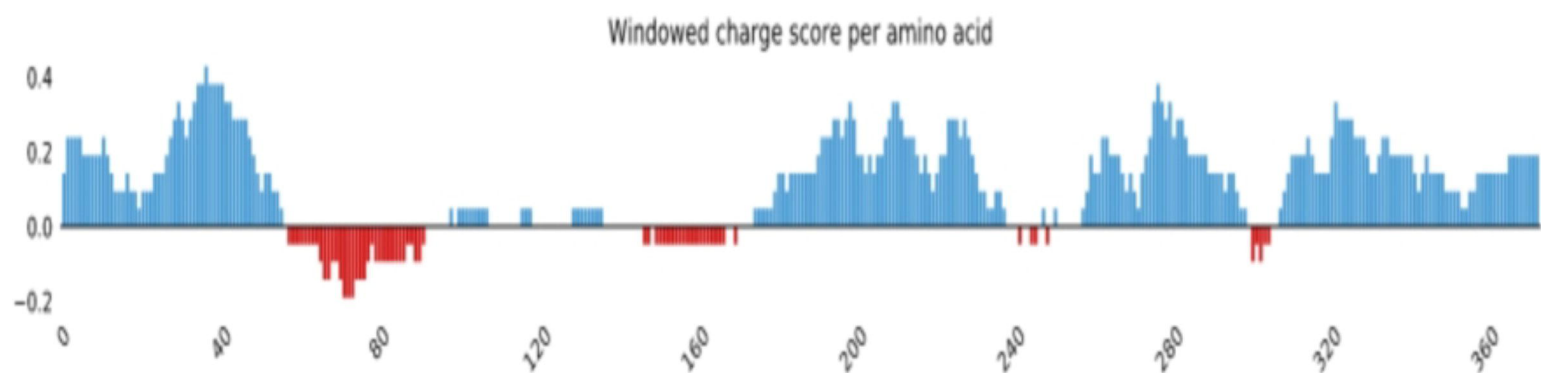
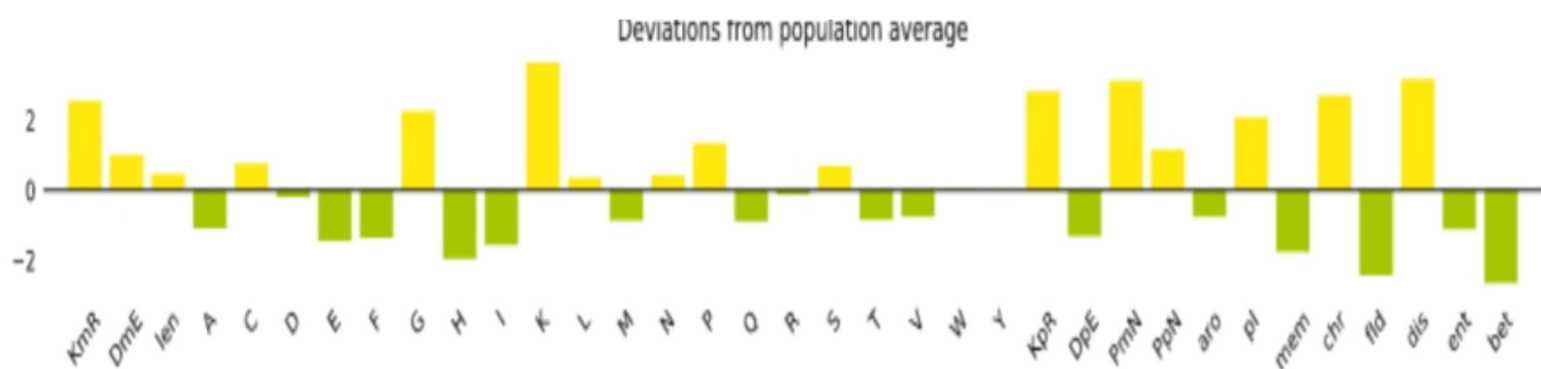
Fig. 6: 3D modelled structure of vaccine protein V2 (A) and V3 (B).

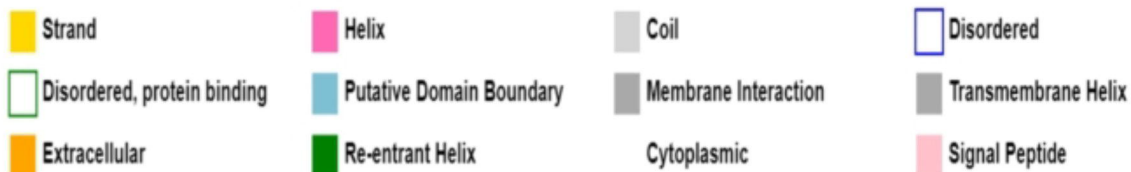
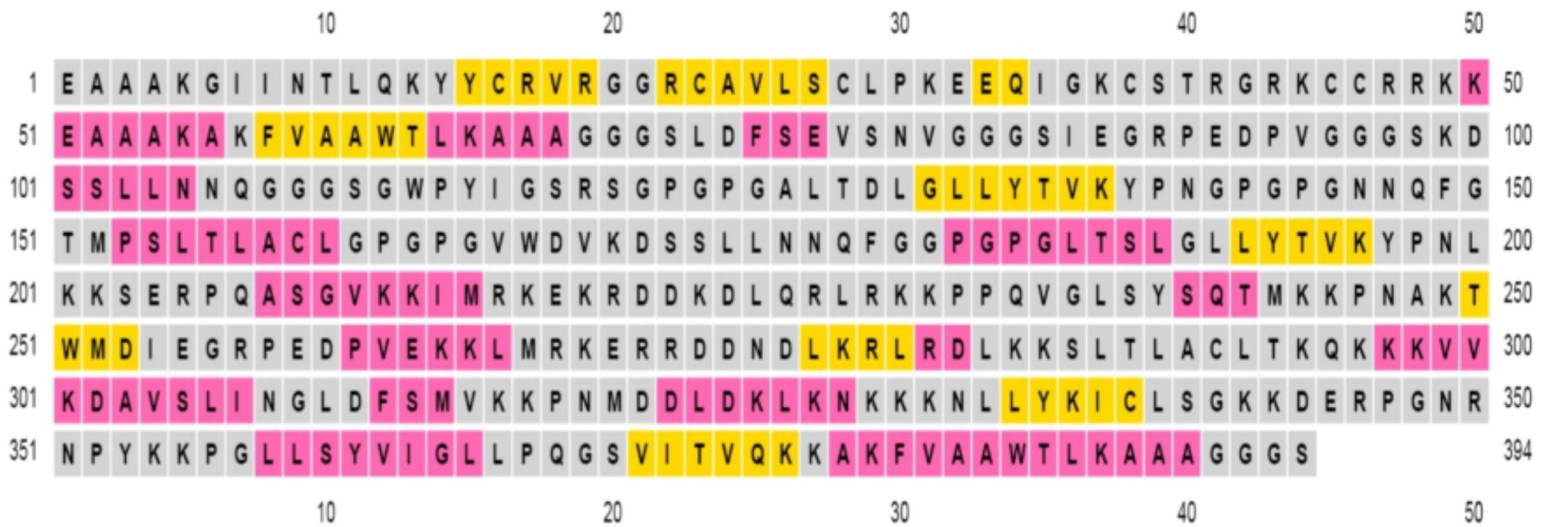
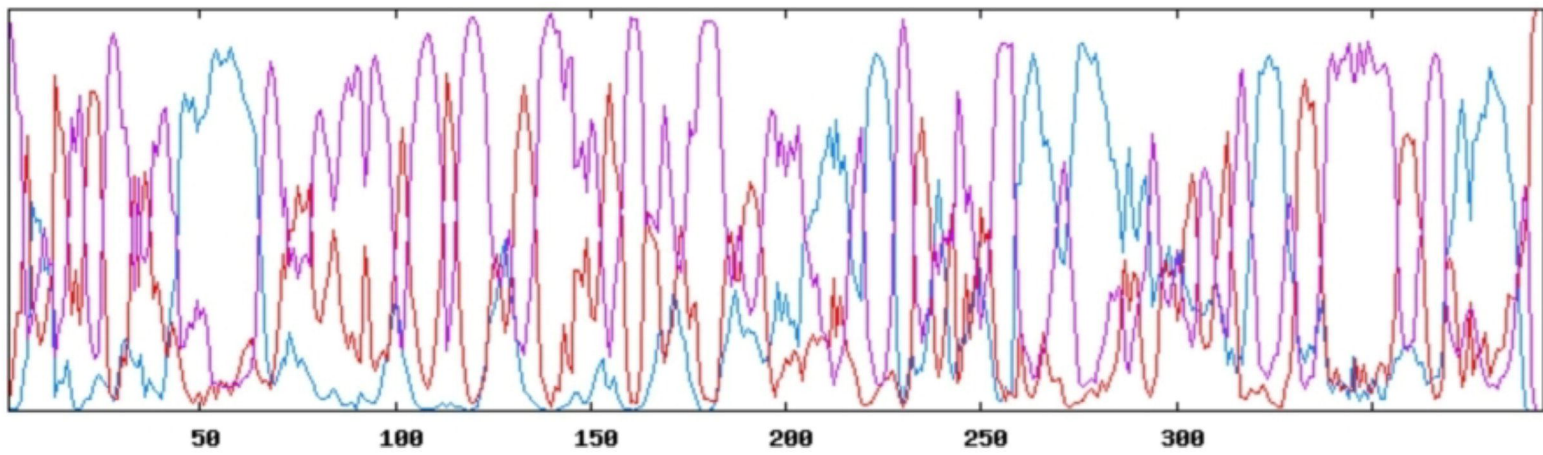
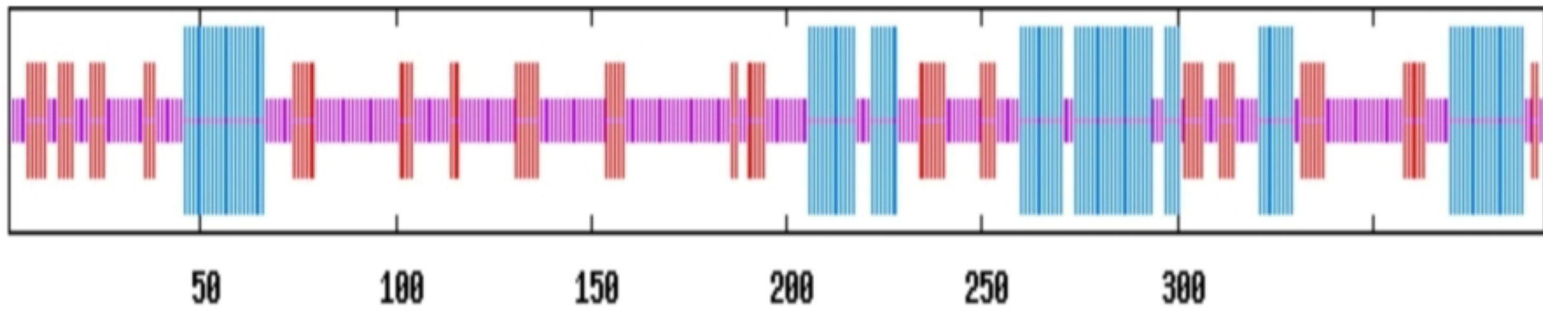
Fig. 7: Molecular dynamics simulation of vaccine protein V1 and TLR-3 complex. Stability of the protein-protein complex was investigated through deformability (A), B-factor (B), eigenvalue (C), variance (D), covariance (E) and elastic network (F) analysis.

Fig. 8: Restriction digestion (A) and *in silico* cloning (B) of the gene sequence of final construct V1 into pET28a(+) expression vector. Target sequence was inserted between BglII (401) and BglII (2187) indicated in violate color.



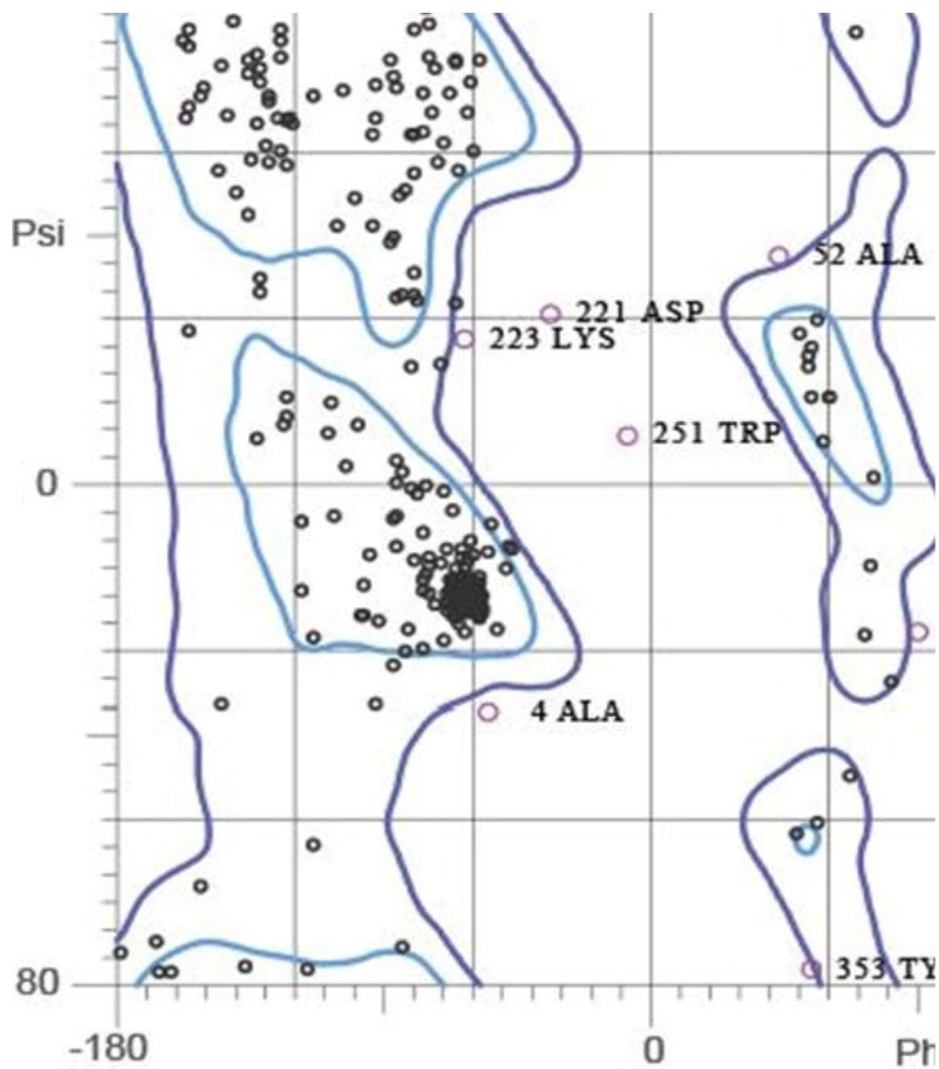




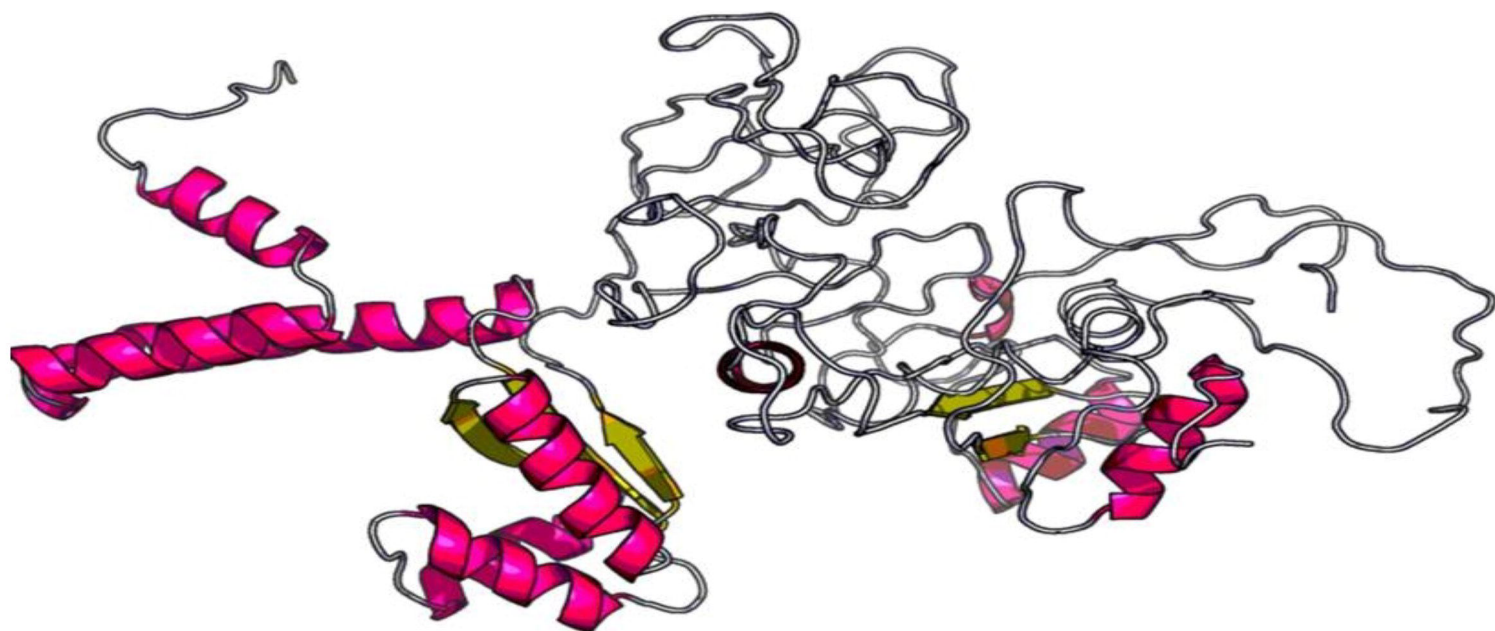




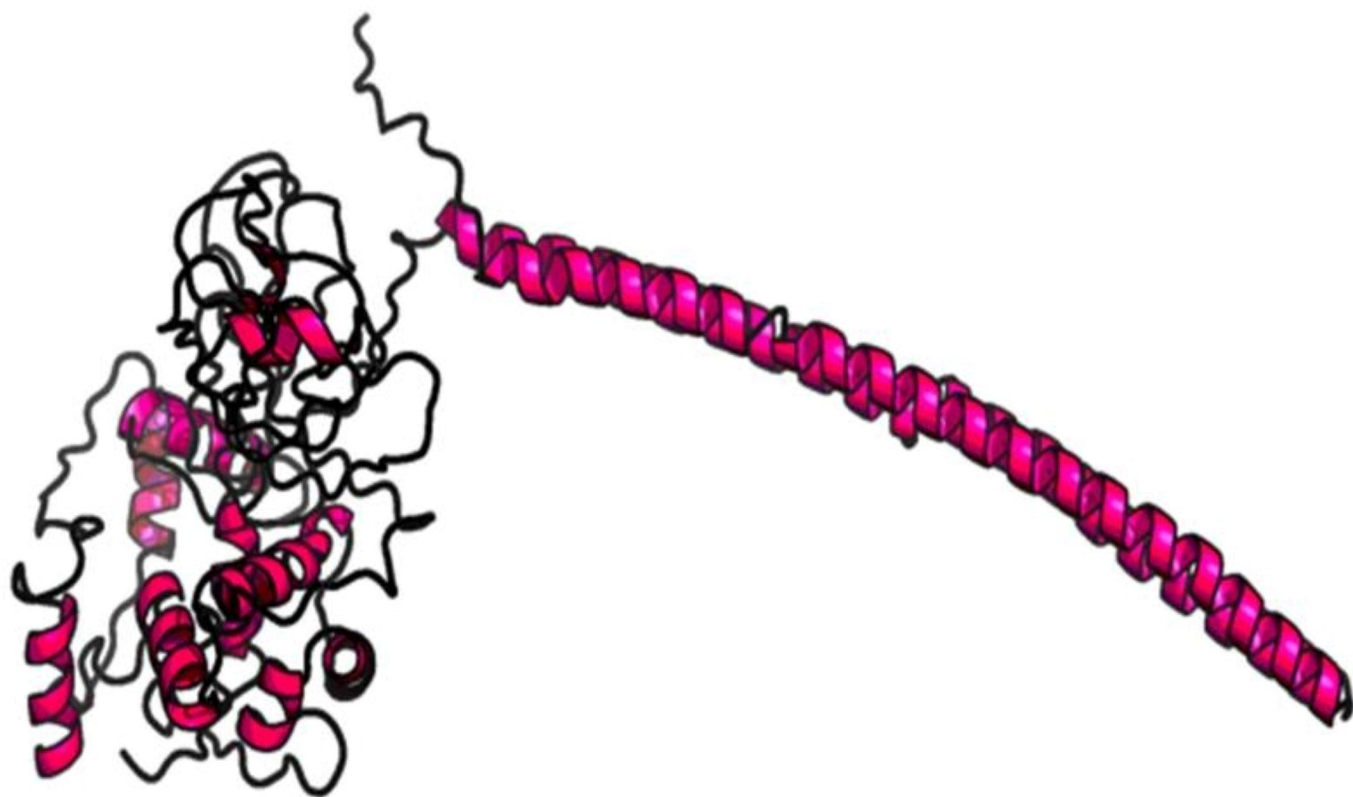
A



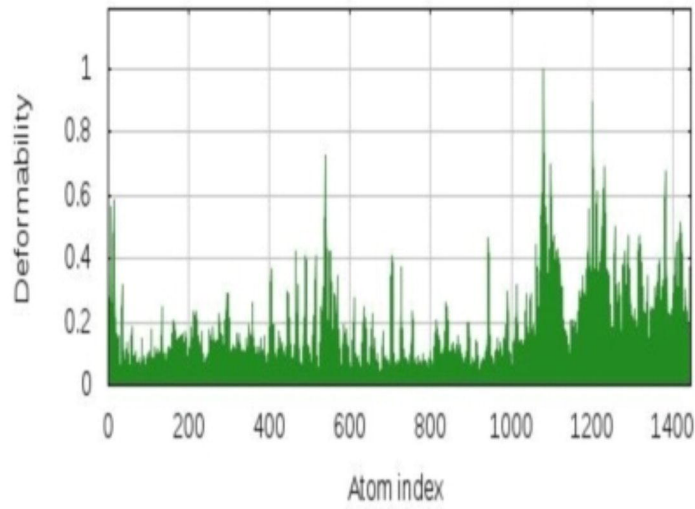
B



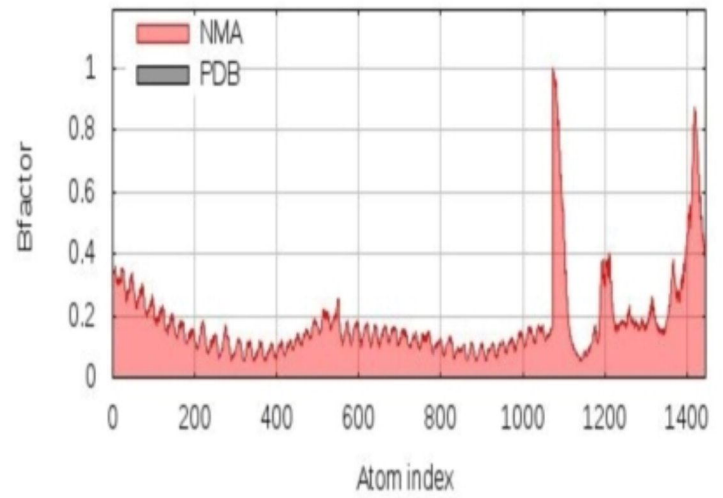
A



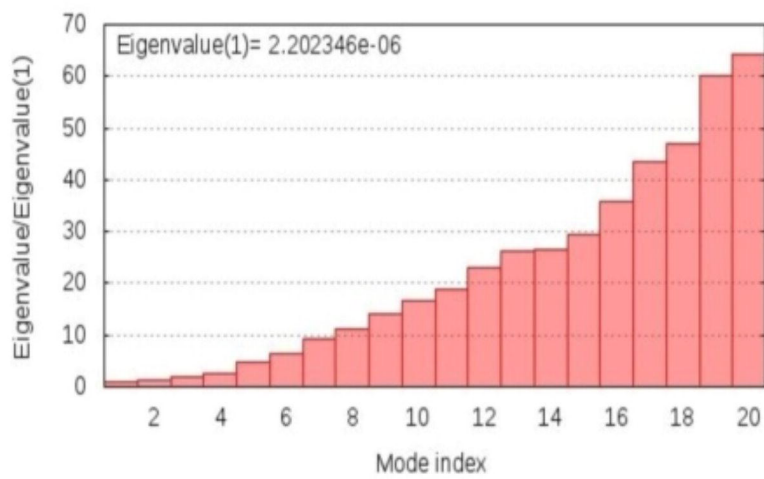
B



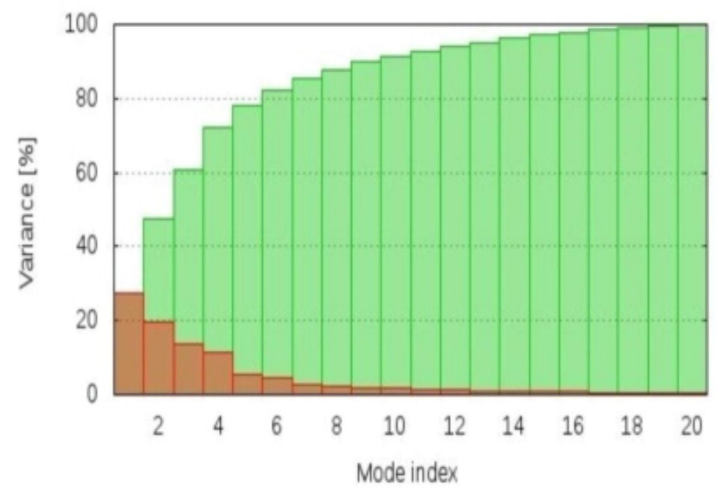
A



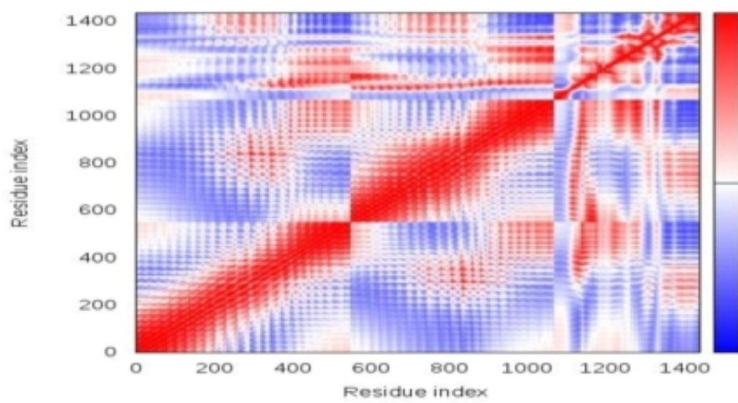
B



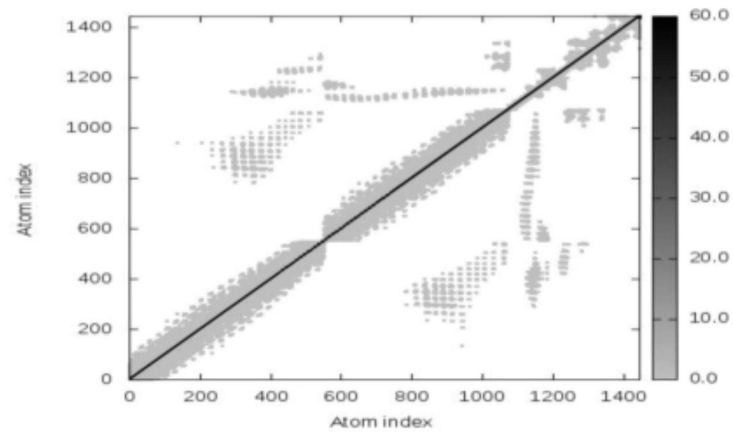
C



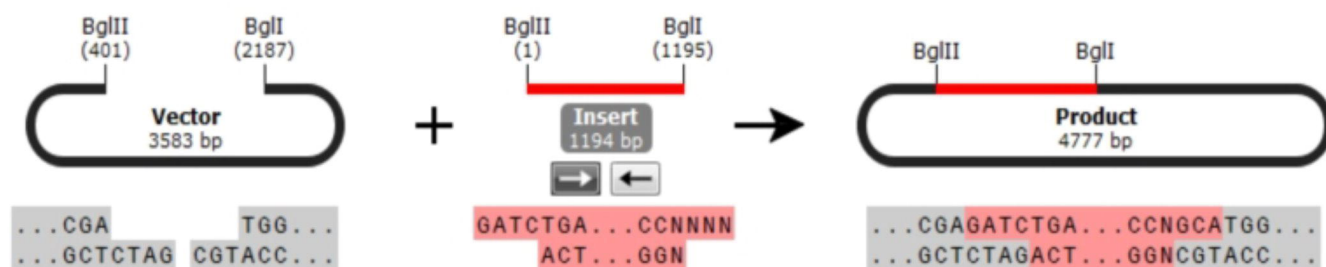
D



E



F



A



B