Infection, Genetics and Evolution 70 (2019) 140–157

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Vaccinomics strategy for developing a unique multi-epitope monovalent vaccine against *Marburg marburgvirus*

Mahmudul Hasan[a,b,*], Kazi Faizul Azim[a,c], Aklima Begum[a], Noushin Anika Khan[a], Tasfia Saiyara Shammi[a], Abdus Shukur Imran[a], Ishtiak Malique Chowdhury[a], Shah Rucksana Akhter Urme[a]

[a] Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet 3100, Bangladesh
[b] Department of Pharmaceuticals and Industrial Biotechnology, Sylhet Agricultural University, Sylhet 3100, Bangladesh
[c] Department of Microbial Biotechnology, Sylhet Agricultural University, Sylhet 3100, Bangladesh

ARTICLE INFO

Keywords:
Reverse vaccinology
TLR8 receptor
T-cell immunity
Ravn Virus
Molecular docking
Simulation dynamics

ABSTRACT

Marburg virus is known to cause a severe hemorrhagic fever (MHF) in both humans and non-human primates with high degree of infectivity and lethality. To date no approved treatment is available for Marburg virus infection. A study was employed to design a novel chimeric subunit vaccine against Marburg virus by adopting reverse vaccinology approach. The entire viral proteome was retrieved from UniprotKB and assessed to design highly antigenic epitopes by antigenicity screening, transmembrane topology screening, allergenicity and toxicity assessment, population coverage analysis and molecular docking approach. Envelope glycoprotein (GP) and matrix protein (VP40) were identified as most antigenic viral proteins which generated a plethora of epitopes. The final vaccine was constructed by the combination of highly immunogenic epitopes along with suitable adjuvant and linkers. Physicochemical and secondary structure of the designed vaccine was assessed to ensure its thermostability, hydrophilicity, theoretical PI and structural behaviors. Disulfide engineering, molecular dynamic simulation and codon adaptation were further employed to develop a unique multi-epitope monovalent vaccine. Docking analysis of the refined vaccine structure with different MHC molecules and human immune TLR8 receptor present on lymphocyte cells demonstrated higher interaction. Moreover, disulfide engineering served to lessen the high mobility region of the designed vaccine in order to extend its stability. Complexed structure of the modeled vaccine and TLR8 showed minimal deformability at molecular level. Finally, translational potency and microbial expression of the modeled vaccine was analyzed with pET28a(+) vector for *E. coli* strain K12. However, further in vitro and in vivo investigation could be implemented for the acceptance and validation of the predicted vaccine against Marburg virus.

1. Introduction

Marburg virus is the first filovirus ever detected in human. In both humans and nonhuman primates, it causes a severe hemorrhagic fever, known as Marburg hemorrhagic fever (MHF) (Mehedi et al., 2011). The average case lethality rate of Marburg virus disease (MARV) since its first recognition in 1967 is 80%. To date 452 cases and 368 documented deaths have been reported as a result of this zoonotic disease. But, the literature reports are suggesting that the actual numbers might be higher (Brauburger et al., 2012; Olival and Hayman, 2014; Martines et al., 2015). Both Marburg Virus (MARV) and Ravn Virus (RAVV) are responsible for Marburg virus disease (MVD). However, previous studies revealed that most of the outbreaks so far caused by the former one with the association of maximum number of human deaths (Supplementary Table 1).

Marburg virus belongs to the same virus family to which Ravn Virus and Ebola virus belong named filoviridae. The genus *Marburgvirus* includes a single species, *Marburg marburgvirus* first made its appearance in 1967 in Europe causing severe and fatal hemorrhagic fever among the laboratory workers in Marburg and Frankfurt (Kuhn et al., 2011;

Abbreviations: MHF, Marburg hemorrhagic fever; MARV, Marburg virus disease; NCBI, National center for Biotechnology information; IEDB, Immune epitope database and analysis resource; CTLs, Cytotoxic T lymphocytes; MGL, Molecular graphics laboratory; MHC, Major histocompatibility complex

* Corresponding author at: Department of Pharmaceuticals and Industrial Biotechnology, Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet 3100, Bangladesh.
E-mail address: mhasan.pib@sau.ac.bd (M. Hasan).

https://doi.org/10.1016/j.meegid.2019.03.003
Received 2 January 2019; Received in revised form 9 February 2019; Accepted 4 March 2019
Available online 05 March 2019
1567-1348/ © 2019 Elsevier B.V. All rights reserved.

Adams and Carstens, 2012). About 4 weeks later further cases were observed in Belgrade. Marburg virus reemerged in two large outbreaks in Congo (DRC) in 1998–2000 and then, for the first time in West African country Angola in 2004–2005 (Bausch et al., 2006; Towner et al., 2006a). At that time, a total 406 cases were observed with a high fatality rate of 83% and 90% in Congo and Angola respectively. Those incidences revealed MARV as major threat for public health (Towner et al., 2006b; MacNeil et al., 2011; Grard et al., 2011). Three Marburg virus disease outbreaks have been reported in Uganda with first one documented in 2007. In 2012, MHF was responsible for with 15 deaths among 26 cases observed in multiple districts (Knust et al., 2015; Towner et al., 2009).

Close interaction between people and animals such as non-human primates, bats, and livestock was attributed as a probable cause of those outbreaks. Egyptian fruit bat (*Rousettus aegyptiacus*) is currently known as a reservoir of Marburg viruses (Towner et al., 2009) and cases have been linked to exposure in caves or mines residing by this organism (Towner et al., 2009; Timen, 2009). The pathological symptoms of the disease involve a number of systemic dysfunctions including edema, hemorrhages, shock, multi organ failure, often resulting in death (Towner et al., 2008).

Due to its high infectivity and lethality, handling of Marburg virus is restricted to high containment biosafety Level 4 laboratories (U.S. Department of Health and Human Services-CDC, 2009). It has also been classified as category A priority pathogen by the National Institute of Allergy and Infectious Diseases (NIAID), select agent by the Centers for Disease Control and Prevention (CDC) and risk group 4 agent by WHO. Currently, there are no vaccines or drugs approved for human to protect against Marburg virus (Brauburger eet al.,2012). Supportive care (fluids, antimicrobials, blood transfusion) has been the primary treatment of patients during the outbreaks (Olival and Hayman, 2014; Martines et al., 2015).

The conventional approach for vaccine designing is time consuming and only allows the identification of abundant antigens that are cultivable under laboratory conditions. Initial approaches using inactivated virus to develop a vaccine against MARV were unsuccessful or had contradictory results (Rollin, 2009). But, the reverse approach to vaccine development overcomes these problems taking benefit of the genome sequence of the pathogen. The strategy aims to combine immunogenomics and immunogenetics with bioinformatics for the development of novel vaccine target (Poland and Ovsyannikova, 2009). Subunit vaccines consist of only the antigenic part of the pathogen with the potential to induce a protective immune response inside host while overcoming the problems caused by live attenuated vaccines (Ahluwalia et al., 2017).

Marburg virus possesses a single surface protein named envelope glycoprotein (GP) that mediates attachment to target cells and virus entry (Will et al., 1993). Besides its function in entry, GP is also associated with immune evasion. Another protein of this virus matrix protein (VP40) plays a major role in the formation of virions and recruiting GP (Dolnik et al., 2008). Both the proteins hold potential to be an effective target for designing a vaccine against Marburgvirus. The present study was conducted to design a unique, non-allergic and immunogenic chimeric vaccine against Marburgvirus utilizing the vaccinomics approach while the wet lab researchers are anticipated to authenticate our prediction.

2. Materials and methods

In the present study, an in silico approach was employed to design vaccine candidates against Marburgvirus to prevent MHF. A flow chart showing the protocol over vaccinomics approach for developing an epitope-based chimeric vaccine has been illustrated in Fig. 1.

2.1. Viral strain selection

The National Center for Biotechnology Information (NCBI) was used for the selection of Marburgvirus *(Marburg marburgvirus)* strain Musoke-80 (https://www.ncbi.nlm.nih.gov/genome/genomes/5). The server provides access to biomedical and genomic information over numerous organisms. Study of other associated information including the genus, family, host, transmission, disease, genome and proteome analysis were performed by using ViralZone, a web-resource of Swiss Institute of Bioinformatics (https://viralzone.expasy.org/194).

2.2. Protein sequence retrieval

UniProt is database containing a large amount of information about the biological information of protein. The entire viral proteome of Marburgvirus *(Marburg marburgvirus)* strain Musoke-80 was retrieved from UniProtKB (https://www.uniprot.org/uniprot/?query+database) consisting seven proteins.

2.3. Antigenic protein screening and structure analysis

Antigenicity refers to the capacity of antigens to be recognized by the immune system. To investigate protein antigenicity and determine the most potent antigenic proteins, the VaxiJen v2.0 server (http://www.ddg-pharmfac.net/vaxijen/) was utilized (Doytchinova and Flower, 2007). From the seven viral proteins, the two structural proteins (envelope glycoprotein and matrix protein VP40) were selected based on their antigenic score. Different physicochemical properties of the proteins were predicted using ProtParam (Gasteiger et al., 2003), one of the ExPASy's servers for primary structure prediction of proteins (Das et al., 2015).

2.4. T-cell epitope prediction

The IEDB offers easy searching of experimental data characterizing antibody and T cell epitopes studied in human and other non-human primates. From this Immune Epitope Database, MHC-I prediction tool (http://tools.iedb.org/mhci/) and MHC-II prediction tool (http://tools.iedb.org/mhcii/) were used to predict the MHC-I binding and MHC-II binding respectively (Buus et al., 2003). Both MHC-I restricted CD8+ cytotoxic T lymphocytes (CTLs) and MHC-II restricted CD4+ cytotoxic T lymphocytes play a pivotal role in controlling viral infections. Hence, identification of T cell epitopes is crucial for understanding the mechanism of T cell activation and epitope driven vaccine design. The protein sequences, envelope glycoprotein and matrix protein VP40 were added to the query box at different time for both type of analysis. MHC class I and MHC class II alleles were selected.

2.5. Transmembrane topology prediction and antigenicity analysis of epitopes

The TMHMM server (http://www.cbs.dtu.dk/services/TMHMM/) predicted the transmembrane helices in proteins. The topology was determined according to the position of the transmembrane helices separated by 'i' if the loop is on the inside or 'o' if it is on the outside (Krogh et al., 2001). Again, VaxiJen v2.0 server (http://www.ddg-pharmfac.net/vaxijen/) was used to determine the epitope antigenicity (Doytchinova and Flower, 2007). The most potent antigenic epitopes were selected for further investigation.

2.6. Allergenicity assessment and toxicity analysis of T-cell epitopes

The prediction of allergens has been explored widely using bioinformatics, with many tools being developed in the last decade. Four servers named AllerTOP (http://www.ddg-pharmfac.net/AllerTop/) (Dimitrov et al., 2013), AllergenFP (http://www.ddg-pharmfac.net/

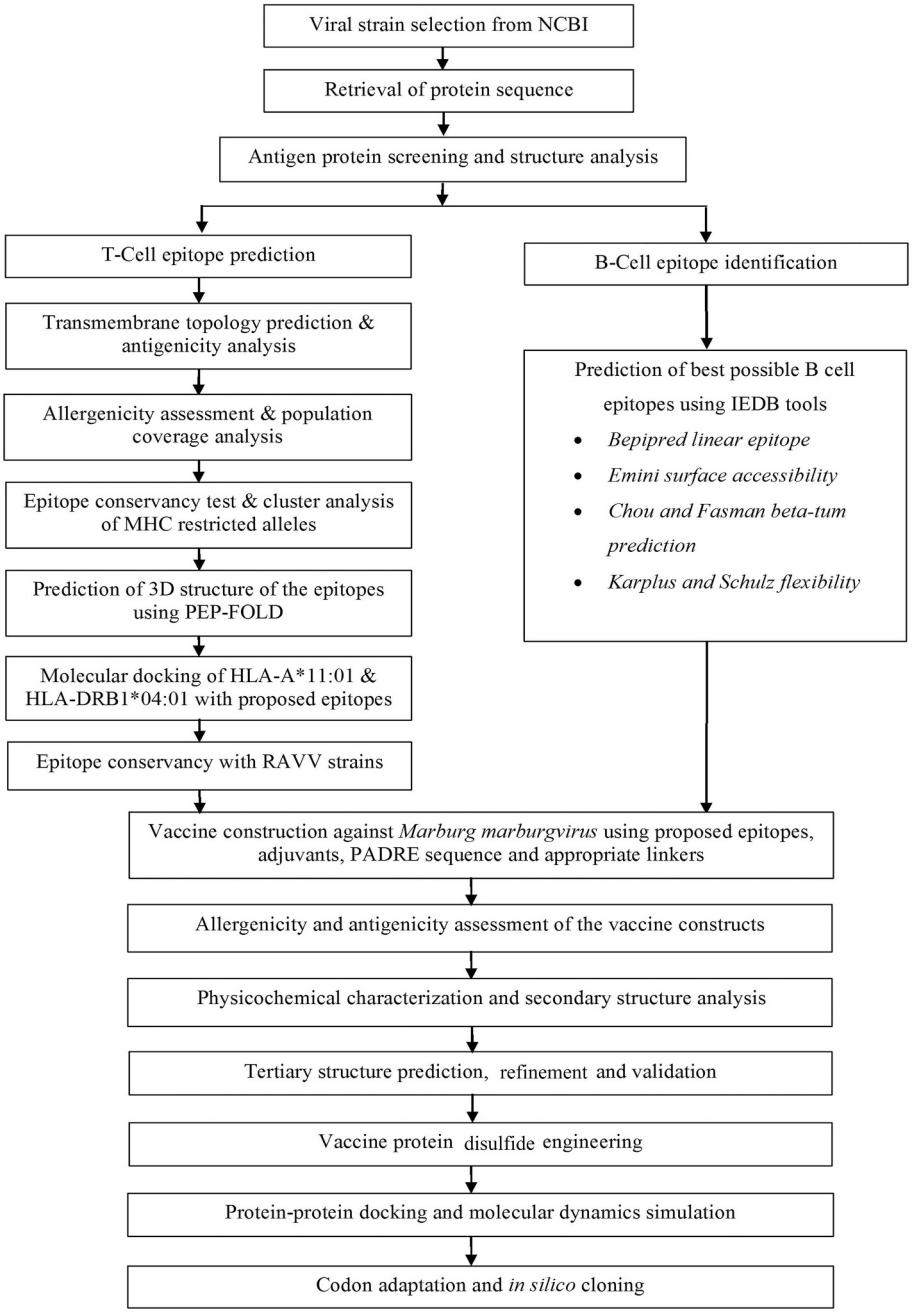

Fig. 1. Flow chart summarizing the protocols for the prediction of epitope based vaccine candidate by in silico reverse vaccinology technique.

AllergenFP/) (Dimitrov et al., 2014), PA^{3}P (http://www.Ipa.saogabriel.unipampa.edu.br:8080/pa3p/) (Chrysostomou and Seker, 2014) and Allermatch (http:// www.allermatch.org/allermatch.py/form) (Fiers et al., 2004) were used to predict the allergenicity of our proposed epitopes for vaccine development. Only the non-allergenic epitopes were allowed to demonstrate the toxicity level by ToxinPred server (http://crdd.osdd.net/raghava/toxinpred/).

2.7. Population coverage analysis

HLA distribution varies among different ethnic groups and geographic regions around the world. So, population coverage must be taken into account when designing an effective vaccine to cover as much as possible populations. In this study, population coverage for each individual epitope was analyzed by the IEDB population coverage calculation tool analysis resource (http://tools.iedb.org/population/).

2.8. Conservancy analysis

Epitope conservancy is a vital step in the immunoinformatic approach as it determines the extent of desired epitope distributions in the homologous protein set. IEDB's epitope conservancy analysis tool (http://tools.iedb.org/conservancy/) was selected for the analysis of conservancy level by concentrating on the identities of the selected proteins.

2.9. Cluster analysis of the MHC restricted alleles

Structure based clustering methods have been proven efficient to identify the super-families of MHC proteins with similar binding specificities. Due to the vast polymorphic nature among species, the distinct specificity of the MHC alleles remains uncharacterized in most cases. In this study, a tool from MHCcluster v2.0 (Thomsen et al., 2013)

server was used to produce pictorial tree-based visualizations and highly instinctive heat-map of the functional alliance between MHC variants.

2.10. Designing three-dimensional (3D) epitope structure

The top ranked epitopes were subjected for the docking study after analyzing through different bioinformatics tools. PEP-FOLD is a de novo approach aimed at predicting peptide structures from amino acid sequences (Maupetit et al., 2010). By Folding peptides on a user specified patch of a protein, it comes with the possibility to generate candidate conformations of peptide-protein complexes (Wang et al., 2011). The server was used to design and retrieve the 3D structure of most potent selected epitopes for further analysis.

2.11. Molecular docking analysis

MGLTools is a software developed for the visualization and analysis of molecular structures (Michel, 1999). It includes AutoDock (an automated docking software) designed to predict the interactions between small molecules (i.e. substrates or drug candidates) and receptor of known 3D structure (Morris et al., 2009). All the operations were performed at 1.00°A space keeping the exhaustiveness parameter at 8.00. The numbers of outputs were set at 10. The docking was conducted using AutoDOCK Vina program based on the above-mentioned parameters. OpenBabel (version 2.3.1) was used to convert the output PDBQT files in PDB format. The best output was selected on the basis of higher binding energy. The docking interaction was visualized with the PyMOL molecular graphics system, version 1.5.0.4 (https://www.pymol.org/).

2.12. Epitope conservancy analysis with RAVV strains

To ensure the broad spectrum efficacy of designed vaccine models we further demonstrated the conservancy pattern of the top epitopes with RAVV strains. Homologous sequences of the selected antigenic proteins (envelope glycoprotein and matrix protein VP40) of RAVV were retrieved from the NCBI database by using BLASTp tool. The epitope conservancy analysis tool (http://tools.iedb.org/conservancy/) at the IEDB was selected for analysis of conservancy pattern. The conservancy levels were determined by focusing on the identities of envelope glycoprotein and matrix protein VP40.

2.13. B-cell epitope identification

The objective of B cell epitope prediction was to find the potential antigen that would interact efficiently with B lymphocytes and initiate an immune response. From the experimental confirmation, it was confirmed that the flexibility of the peptide is associated to its antigenicity. For a B cell to be potential epitope, it must have proper surface accessibility as well. B cell epitope prediction tools from IEDB were used to identify the B cell antigenicity based on six different algorithms which include Kolaskar and Tongaonkar antigenicity scale (Kolaskar and Tongaonkar, 1990), Emini surface accessibility prediction (Emini et al., 1985), Karplus and Schulz flexibility prediction (Karplus and Schulz, 1985), Bepipred linear epitope prediction analysis (Jespersen et al., 2017), Chou and Fasman beta turn prediction (Chou and Fasman, 1978) and Parker hydrophilicity prediction (Parker et al., 1986).

2.14. Vaccine construction

Subunit vaccines consist of antigenic parts of a pathogen to stimulate an immunogenic reaction in the host. The predicted T-cell and B-cell epitopes were conjugated in a sequential manner to design the final vaccine construct. All three vaccine proteins started with an adjuvant followed by the top CTL epitopes for both capsid protein VP1 and protein VP2, then by top HTL epitopes and BCL epitopes respectively, in the similar fashion. Three vaccine sequence were constructed named V1, V2 and V3, each associated with different adjuvants including beta defensin (a 45 mer peptide), L7/L12 ribosomal protein and HABA protein (*M. tuberculosis*, accession number: AGV15514.1). Interactions of adjuvants with toll like receptors (TLRs) polarize CTL responses and induce robust immunoreactions [Rana and Akhter, 2016]. Beta defensin adjuvant can act as an agonist to TLR1, TLR2 and TLR4 receptor. On the contrary, L7/L12 ribosomal protein and HBHA protein are agonists to TLR4 only. To overcome the problems caused by highly polymorphic HLA alleles, PADRE sequence was also incorporated along with the adjuvant peptides. EAAAK linkers were used to join the adjuvant and CTL epitopes. Similarly, GGGS, GPGPG and KK linkers were used to conjugate the CTL, HTL and BCL epitopes respectively. Utilized linkers ensured the effective separation of individual epitopes in vivo (Hajighahramani et al., 2017; Pandey et al., 2016).

2.15. Allergenicity and antigenicity prediction of different vaccine constructs

AlgPred v.2.0 (Chrysostomou and Seker, 2014) sever was used to predict the non-allergic nature of the constructed vaccines. The server developed an algorithm by considering the auto cross covariance transformation of proteins into uniform vectors of similar length. The accuracy of results ranges from 70% to 89% depending on species. We further used VaxiJen v2.0 server (Doytchinova and Flower, 2007) to evaluate the probable antigenicity of the vaccine constructs in order to suggest the superior vaccine candidate. The server analyzed the immunogenic potential of the given proteins through an alignment-independent algorithm.

2.16. Physicochemical characterization and secondary structure analysis of vaccine protein

The final vaccine construct was characterized on the basis of physical and chemical properties. ProtParam, a tool provided by Expasy server (Gasteiger et al., 2003) was used to functionally characterize the vaccine constructs according to molecular weight, aliphatic index, isoelectric pH, hydropathicity, instability index, GRAVY values, estimated half-life and various physicochemical properties. By comparing the pK values of different amino acids, the server computes these parameters of a given protein sequence. Aliphaticindex is the volume occupied by the aliphatic side chains of protein. Grand average of hydropathicity was computed by summing the hydropathicity of all amino acid residues present in the protein sequence and then by dividing it by total number of amino acid residues. The PSIPRED v3.3 (Kosciolek and Jones, 2014) and NetTurnP 1.0 program (Petersen et al., 2010; Thaysen-Andersen and Packer, 2012) was used to predict the alpha helix, beta sheet and coil structure of the vaccine constructs.

2.17. Vaccine tertiary structure prediction, refinement and validation

The RaptorX server performed 3D modeling of the designed vaccines depending on the degree of similarity between target protein and available template structure from PDB (Kallberg et al., 2014; Peng and Xu, 2011; Hasan et al., 2015a). Refinement was conducted using ModRefiner (Xy and Zhang, 2011) followed by FG-MD refinement server (Zhang et al., 2011) to improve the accuracy of the predicted 3D modeled structure. ModRefiner drew the initial model closer to its native state based on hydrogen bonds, side-chain positioning and backbone topology, thus resulting in significant improvement in the physical quality of the local structure. FG-MD is another molecular dynamics based algorithm for structure refinement at atomic level. The refined protein structure was further validated by Ramachandran plot assessment at RAMPAGE (Lovell et al., 2002; Al-Hakim et al., 2015).

Table 1
ProtParam analysis of retrieved viral proteins.

Proteins	Accession ID	Molecular weight	Instability index	Half-life	Theoretical pI	No. of Amino acids	Total No. of Atoms	Estinction co-efficient
Envelope glycoprotein	P35253	74,376.19	45.52	30 h	5.88	681	10,350	80,620
Matrix protein VP40	P35260	33,793.72	25.46	30 h	9.40	303	4757	28,880

2.18. Vaccine protein disulfide engineering

Disulfide bonds enhance the geometric conformation of proteins and provide significant stability. DbD2, an online tool was used to design such bonds for the constructed vaccine protein (Craig and Dombkowski, 2013). The server detects and provides a list of residue pairs with proper geometry which have the capacity to form disulfide bond when individual amino acids are mutated to cysteine.

2.19. Protein-protein docking

Molecular docking aims to determine the binding affinity between a receptor molecule and ligand (Solanki and Tiwari, 2018). Inflammations caused by single stranded RNA virus are involved with immune receptors, mainly by TLR-7 and TLR-8 present over the immune cells (Heil et al., 2004; Cros et al., 2010). An approach for protein-protein docking was employed to determine the binding affinity of designed subunit vaccines with different HLA alleles and TLR-8 immune receptor by using ClusPro 2.0. (Comeau et al., 2004), hdoc (Macalino et al., 2018; Kangueane and Nilofer, 2018) and PatchDock server (Schneidman-Duhovny et al., 2005). The 3D structure of different MHC molecules and human TLR-8 receptor was retrieved from RCSB protein data bank. The above mentioned servers were used to obtain the desirable complexes in terms of better electrostatic interaction and free binding energy. PatchDock generated a number of solutions which were again subjected to the FireDock server to refine the complexes.

2.20. Molecular dynamics simulation

Molecular dynamics study is important to strengthen any in silico prediction and demonstrate the stability of protein-protein complex. Stability can be determined by comparing the essential dynamics of proteins to their normal modes (Aalten et al., 1997; Wuthrich et al., 1980). This powerful tool is an alternative to the costly atomistic simulation (Tama and Brooks, 2006; Cui and Bahar, 2007) iMODS server explains the collective motion of proteins by analyzing the normal modes (NMA) in internal coordinates (Lopez-Blanco et al., 2014). The structural dynamics of protein complex 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; Prabhakar et al., 2016). It predicted 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 the ability to deform at each of its residues for a given molecule. The eigenvalue related to each normal mode describes the motion stiffness. This value is directly linked to the energy required to deform the structure. Deformation is much easier if the eigenvalue is low and vice versa (Lopez-Blanco et al., 2011).

2.21. Codon adaptation and in silico cloning

Codon adaptation tools are used for adapting the codon usage to the well characterized prokaryotic organisms to accelerate the expression rate in them. *E. coli* strain K12 was selected as host for cloning purpose of the designed vaccine construct. Due to the lack of similarities between the codon usage of human and *E. coli*, the approach was adopted to achieve higher expression of vaccine protein V1 in the selected host. Rho independent transcription termination, prokaryote ribosome-binding site and cleavage sites of several restriction enzymes (i.e. *Bgl*II and Apa1) were avoided during the operation performed by JCAT server (Grote et al., 2005). The optimized sequence of vaccine protein V1 was reversed and then conjugated with BglII and Apa1 restriction site at the N-terminal and C-terminal sites respectively. SnapGene (Solanki and Tiwari, 2018) restriction cloning module was used to insert the adapted sequence between BglII (401) and *Apa*I (1334) of pET28a(+) vector.

3. Results

3.1. Protein sequence retrieval

The entire viral proteome of Marburgvirus *(Marburg marburgvirus)* Musoke-80 strain was extrated from UniProtKB consisting seven proteins named envelope glycoprotein, matrix protein VP40, RNA directed RNA polymerase L, nucloeoprotein, polymerase cofactor VP35, membrane-associated protein VP24 and minor nucleoprotein VP30. All the sequences were from Kenya.

3.2. Antigenic protein prediction and structure analysis

Envelope glycoprotein (Accession ID: P35253) and matrix protein VP40 (Accession ID: P35260) were selected as the most potent antigenic protein with total prediction score of 0.5474 and 0.4107 respectively and allowed for further analysis. Various physiochemical parameters of the proteins were analyzed by ProtParam tools as listed in Table 1.

3.3. T-cell epitope prediction

Numerous immunogenic epitopes from envelope glycoprotein and matrix protein VP40 were identified to be potent T cell epitopes using both MHC-I and MHC- II binding predictions of IEDB. Approximately 18,117 immunogenic epitopes of envelope glycoprotein and 9342 immunogenic epitopes of matrix protein VP40 were generated to be T cell epitopes that can bind a large number of HLA-A and HLA-B alleles with a very high binding affinity using the MHC-I binding predictions of the IEDB with recommended methods. 10,800 and 7722 immunogenic epitopes were also generated using MHC-II binding prediction tool of IEDB for envelope glycoprotein and matrix protein VP40 respectively. Epitopes that bind to the maximum number of HLA cells with high binding affinity were selected (Table 2 and Table 3).

3.4. Transmembrane topology prediction and antigenicity analysis

Top epitopes from both proteins were selected as putative T cell epitope candidates based on their transmembrane topology screening (Table 2) and antigenic scoring (Table 3). Epitopes with a high immunogenicity exhibited potential to elicit effective T-cell response. Predicted epitopes showed highly antigenicity scoring range (Envelop glycoprotein, 1.37 to 2.80; Matrix protein VP40, 0.98 to 2.22) where the threshold value of antigenic potential was 0.4 for VaxiJen server.

3.5. Allergenicity assessment and toxicity analysis of T-cell epitopes

Based on the allergenicity assessment by four servers (i.e. AllerTOP, AllergenFP, PA^{3}P, Allermatch), epitopes that were found to be non-

Table 2
Predicted T-cell epitopes (MHC-I peptides) of envelope glycoprotein and matrix protein VP40.

Envelope glycoprotein						Matrix protein VP40					
Epitope	Start	End	Topology	No. of HLAs binding epitope	AS	Epitope	Start	End	Topology	No. of HLAs binding epitope	AS
APIDFDPVP	34	42	outside	54	2.80	VPAWLPLGI	80	88	Outside	81	2.22
NAPIDFDPV	33	41	outside	81	2.47	AWLPLGIMS	82	90	Outside	27	1.93
PIDFDPVPN	35	43	outside	27	2.25	PAWLPLGIM	81	89	Outside	54	1.79
PEIKLTSTP	30	38	outside	54	2.11	YVGDLNLDDQ	43	52	Outside	27	1.57
CCIGIEDLS	49	57	outside	54	2.10	QGITPNYVG	37	45	Outside	27	1.55
IDFDPVPNT	36	44	outside	54	2.05	AWLPLGIMSN	82	91	Outside	27	1.52
DFDPVPNTK	37	45	outside	81	1.98	GDLNLDDQF	45	53	Outside	54	1.50
GTGWGLGGK	5	13	outside	54	1.97	VGDLNLDDQ	44	52	Outside	27	1.49
PPKNVEYTE	7	15	outside	54	1.78	YVGDLNLDD	43	51	Outside	27	1.48
GVPPKNVEY	5	13	outside	81	1.76	ADMFSVKEG	253	261	Outside	54	1.33

Table 3
Predicted T-cell epitopes (MHC-II peptides) of envelope glycoprotein and matrix protein VP40.

Envelope glycoprotein						Matrix protein VP40					
Epitope	Start	End	Topology	No. of HLAs binding Epitope	AS	Epitope	Start	End	Topology	No. of HLAs cells binding Epitope	AS
NAPIDFDPVPNTKTI	453	467	outside	27	1.73	VPAWLPLGIMSNFEY	80	94	Outside	27	1.38
APIDFDPVPNTKTIF	454	468	outside	27	1.63	GITPNYVGDLNLDDQ	38	52	Outside	27	1.16
APIDFDPVPNTKTIF	154	168	outside	27	1.63	QGITPNYVGDLNLDD	37	51	Outside	27	1.14
INAPIDFDPVPNTKT	452	466	outside	27	1.59	TVKGVPAWLPLGIMS	76	90	Outside	27	1.10
PIDFDPVPNTKTIFD	450	469	outside	27	1.57	TMHPAVSIHPNLPPI	190	204	Outside	27	1.07
RWAFRTGVPPKNVEY	69	83	outside	27	1.5	HPAVSIHPNLPPIVL	192	206	Outside	27	1.04
FRTGVPPKNVEYTEG	72	86	outside	27	1.45	PAWLPLGIMSNFEYP	81	95	Outside	27	1.01
AFRTGVPPKNVEYTE	71	85	outside	27	1.4	MHPAVSIHPNLPPIV	191	205	Outside	27	1.00
VQEDDLAAGLSWIPF	521	535	outside	27	1.38	AWLPLGIMSNFEYPL	82	96	Outside	27	0.99
LINAPIDFDPVPNTK	451	465	outside	27	1.37	HPNLPPIVLPTVKKQ	198	212	Outside	27	0.98

Table 4
Allergenicity assessment, toxicity test and conservancy analysis of the predicted epitopes generated from envelope glycoprotein and matrix protein VP40.

Envelope glycoprotein				Matrix protein VP40			
Epitope	Allergenicity	Toxicity	Conservancy	Epitope	Allergenicity	Toxicity	Conservancy
APIDFDPVP	Non Allergen	Non Toxin	72.9%	VPAWLPLGI	Non Allergen	Non Toxin	100%
CCIGIEDLS	Non Allergen	Toxin	91.6%	YVGDLNLDD	Non Allergen	Non Toxin	95.6%
GVPPKNVEY	Non Allergen	Non Toxin	95.8%	QGITPNYVG	Non Allergen	Non Toxin	52.1%
PPKNVEYTE	Non Allergen	Non Toxin	95.8%	VGDLNLDDQ	Non Allergen	Non Toxin	95.6%
PEIKLTSTP	Non Allergen	Non Toxin	4.17%	ADMFSVKEG	Non Allergen	Non Toxin	95.6%
NAPIDFDPVPNTKTI	Non Allergen	Non Toxin	70.8%	VPAWLPLGIMSNFEY	Non Allergen	Non Toxin	100.%
APIDFDPVPNTKTIF	Non Allergen	Non Toxin	70.8%	GITPNYVGDLNLDDQ	Non Allergen	Non Toxin	95.6%
RWAFRTGVPPKNVEY	Non Allergen	Non Toxin	89.5%	QGITPNYVGDLNLDD	Non Allergen	Non Toxin	52.1%
VQEDDLAAGLSWIPF	Non Allergen	Non Toxin	89.5%	TMHPAVSIHPNLPPI	Non Allergen	Toxin	39.1%
FRTGVPPKNVEYTEG	Non Allergen	Non Toxin	89.5%	HPAVSIHPNLPPIVL	Non Allergen	Non Toxin	39.1%

allergen for human were identified. Epitopes those were indicated as allergen for human and classified as toxic or undefined were removed from the predicted list of epitopes (Table 4).

3.6. Population coverage analysis

All indicated alleles in supplementary data were identified as optimum binders with the predicted epitopes and were used to determine the population coverage. Result showed that envelope glycoprotein covers around 90% and 80% of entire world population coverage for MHC class I and MHC class II respectively where the percentage was 90 and 60 for matrix protein 1. Overall Population coverage result for envelope glycoprotein and matrix protein VP40 are shown in Fig. 2A and Fig. 2B.

3.7. Conservancy analysis

Putative epitopes generated from envelope glycoprotein and matrix protein VP40 were found to be highly conserved with maximum conservancy level of 95.8% and 100% respectively (Table 4). Only the epitopes showing conservancy at a satisfactory level were allowed for further docking study and used to construct the final vaccine candidates to ensure a broad spectrum immune response.

3.8. Cluster analysis of the MHC restricted alleles

Clustering of both class I and class II HLA molecules that were detected to interact with the predicted epitopes was generated through conventional phylogenetic method based on sequence data available for different HLA-A and HLA-B alleles. Fig. 3 is illustrating the function based clustering of HLA alleles (heat map) where red zones indicating strong correlation and yellow zone showing weaker interaction.

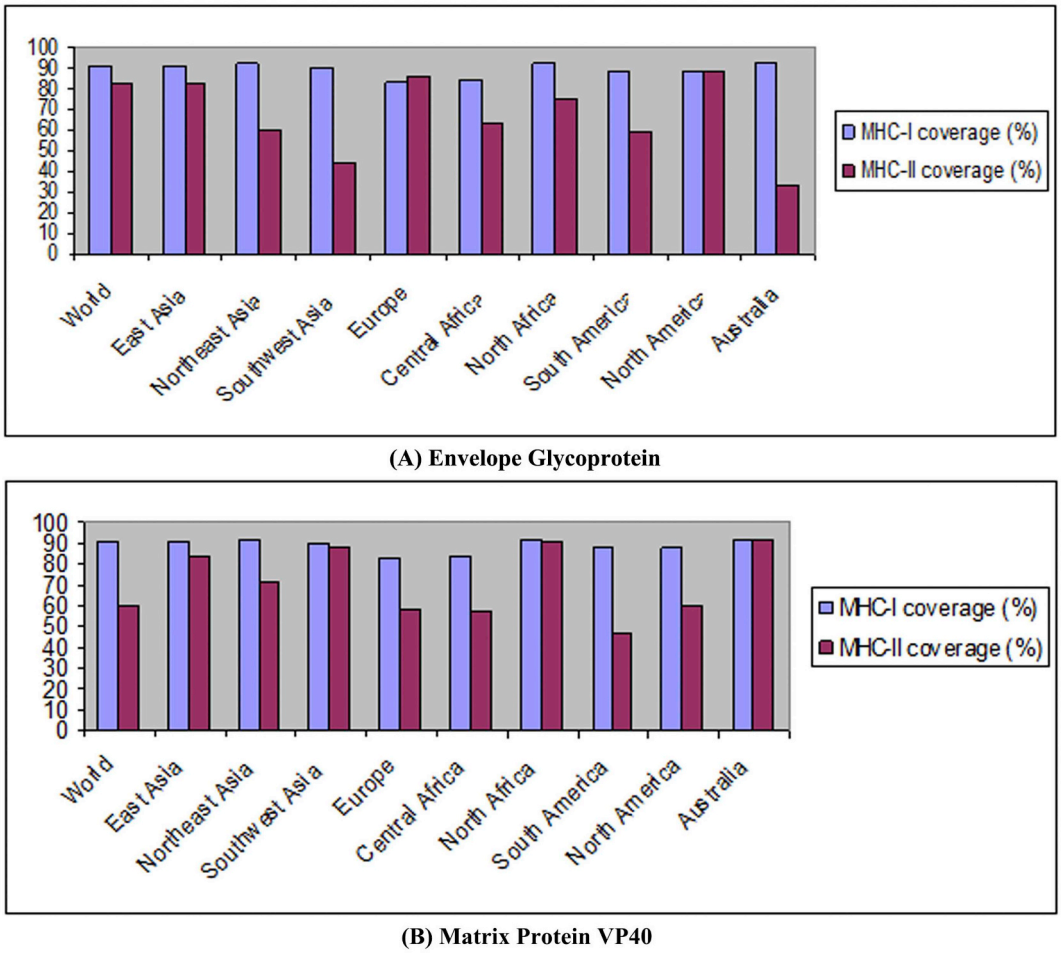

Fig. 2. Population coverage analysis of envelope glycoprotein (A) and matrix protein VP40 (B).

3.9. Molecular docking analysis and HLA allele interaction

12 T-cell epitopes (six from envelope glycoprotein and six from matrix protein VP40) were subjected to PEP-FOLD3 web-based server for 3D structure conversion in order to analyze their interactions with HLA molecule. From five structures, modeled by the server for each individual epitopes, the best one was identified for docking study. On the basis of available Protein Data Bank (PDB) structures deposited in the database, HLA-A*11:01 and HLA-DRB1*04:01 was selected for docking analysis with MHC class I and class II binding epitopes respectively.

All the selected epitopes were allowed for docking analysis and the

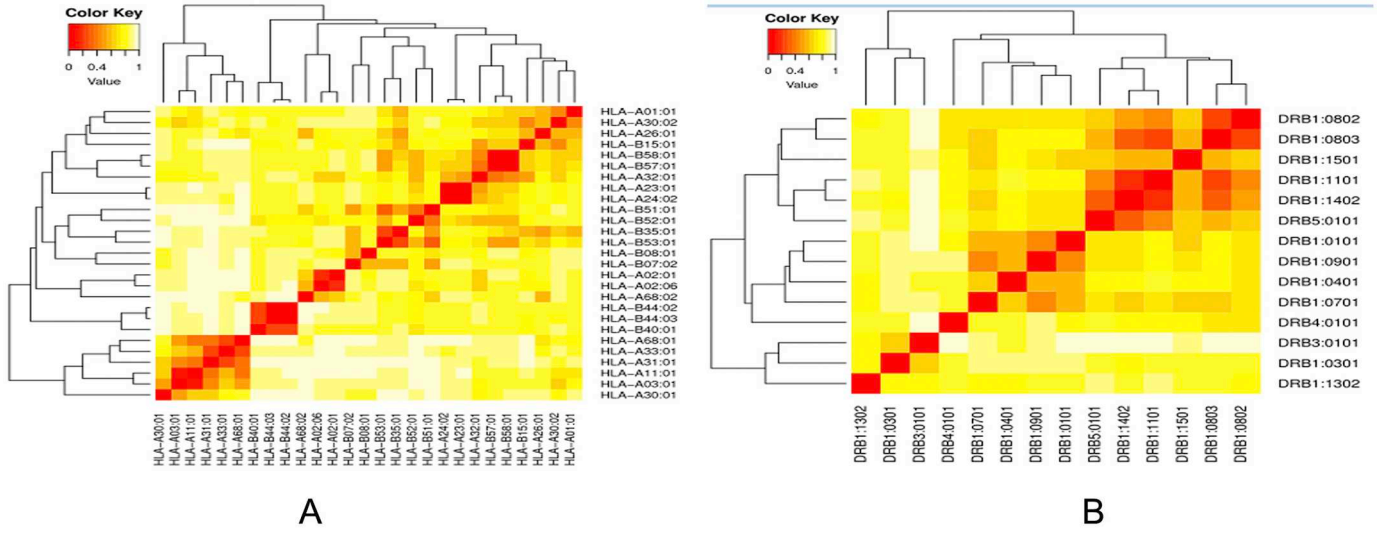

Fig. 3. Cluster analysis of the HLA alleles: (A: MHC-I molecules, B: MHC-II molecules (red color in the heat map indicating strong interaction, while the yellow zone indicating the weaker interaction). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
Binding energy of suggested T-cell epitopes with selected class I and class II MHC molecules generated from molecular docking analysis.

Epitopes	MHC allele	Binding Energy (kcal/mol)	Epitopes	MHC allele	Binding energy (kcal/mol)
APIDFDPVP		−9.5	NAPIDFDPVPNTKTI		−7.6
GVPPKNVEY		−8.8	RWAFRTGVPPKNVEY		−7.5
PPKNVEYTE		−7.8	VQEDDLAAGLSWIPF		−7.8
VPAWLPLGI	MHC-A*11:01	−9.1	VPAWLPLGIMSNFEY	HLA-DRB1*04:01	−7.0
YVGDLNLDD		−8.6	GITPNYVGDLNLDDQ		−6.5
QGITPNYVG		−7.9	QGITPNYVGDLNLDD		−6.2

Table 6
Epitope conservancy with RAVV strains.

Protein	Epitopes	Conservancy	Epitopes	Conservancy
Envelope Glycoprotein (GP)	GVPPKNVEY	80%	RWAFRTGVPPKNVEY	80%
	PPKNVEYTE	80%	VQEDDLAAGLSWIPF	80%
Matrix Protein VP40	VPAWLPLGI	100%	VPAWLPLGIMSNFEY	100%
	YVGDLNLDD	100%	GITPNYVGDLNLDDQ	100%

binding energies were analyzed (Table 5). Results showed that 'VQEDDLAAGLSWIPF' epitope of envelope glycoprotein (GP) bound in the groove of the HLA-DRB1*04:01 with an energy of −7.8 kcal/mol. The demonstrated energy was −7.0 kcal/mol for epitope 'VPAWLPLGIMSNFEY' contained the 9-mer core 'VPAWLPLGI' of matrix protein VP40. On the contrary, VP1-epitope 'APIDFDPVP' was found to be superior in terms free binding energy while interacted with HLA-A*11:01 (−9.5 kcal/mol).

3.10. Epitope conservancy analysis with RAVV strains

Two sets of homologous protein sequences for envelope glycoprotein and matrix protein VP40 of RAVV were retrieved for the input of IEDB conservancy analysis server. Results showed that 8 of the final 12 T-Cell epitopes used to construct the vaccines were conserved within (80–100)% of the RAVV strains (Table 6).

3.11. B-cell epitope identification

For envelope glycoprotein, Bepipred prediction method predicted the peptide sequences from 466 to 489 and 612–658 amino acids as potential B cell epitopes that could induce the preferred immune responses (Fig. 4A) Emini-surface accessibility prediction was also conducted which indicated 293–31 and 626–634 amino acid residues to be more accessible (Fig. 4B). Chou and Fasman beta-turn prediction method displayed regions from 26 to 32 and 308–314 as potential Beta-turn regions (Fig. 4C). Karplus and Schulz flexibility prediction method found the region of 190–196 and 488–494 amino acid residues as most flexible regions (Fig. 4D). In contrast, Kolaskar and Tongaonkar antigenicity result confirmed the region from 424 to 432 and 601–614 as highly antigenic (Fig. 4E) while Parker hydrophilicity prediction indicated 407–413 and 606–612 amino acid residues to be more potent (Fig. 4F).

In case of matrix protein VP40, bepipred prediction method identified the peptide sequences from 29 to 40 and 214–223 amino acids with ability to induce preferred immunity (Fig. 5A). The regions from 177 to 184 and 208–222 amino acid residues were more accessible based on Emini surface accessibility prediction algorithm (Fig. 5B). Regions from 29 to 35 and 217–223 were potential beta-turn regions on the basis of Chou and Fasman beta-turn prediction (Fig. 5C). Karplus and Schulz flexibility prediction method showed the region of 178–184 and 216–222 amino acid residues as a most flexible region (Fig. 5D). On the contrary, Kolaskar & Tongaonkar antigenicity result confirmed the region from 193 to 212 & 225–240 as highly antigenic (Fig. 5E) while Parker hydrophilicity prediction indicated 217–223 amino acid residues to be more potent (Fig. 5F). Allergenicity pattern of the predicted B-cell epitopes is shown in Table 7.

3.12. Vaccine construction

Each of the constructs consisted of a protein adjuvant followed by PADRE peptide sequence, while the rest was occupied by the T-cell and B-cell epitopes and their respective linkers. PADRE sequence was incorporated to maximize the efficacy and potency of the peptide vaccine. All three designed vaccines comprised 6 CTL epitopes, 6 HTL epitopes and 8 BCL epitopes combined together in a sequential manner. A total 3 vaccines of 403 (V1), 488 (V2) and 517 (V3) amino acid long were constructed (Table 8) and further analyzed to investigate their immunogenic potential.

3.13. Allergenicity and antigenicity prediction of different vaccine constructs

Results showed that all three constructs (V1, V2 and V3) were non-allergic in behavior. However, V1 was best in terms of safety and found superior as potential vaccine candidate with better antigenicity (0.772) and ability to stimulate preferred immune response (Table 8).

3.14. Physicochemical characterization of vaccine protein

The molecular weight of the vaccine construct V1 was 41.27 kDa which ensured its good antigenic potential. The theoretical pI 9.33 indicated that the protein will have net negative charge above the pI and vice versa. The extinction coefficient was 56,380, assuming all cysteine residues are reduced at 0.1% absorption. The estimated half-life of the constructed vaccine was expected to be 1 h in mammalian reticulocytes in vitro while > 10 h in *E. coli* in vivo. Thermostability and hydrophilic nature of the vaccine protein was represented by aliphatic index and GRAVY value which were 65.32 and − 0.631 respectively. The computed instability index of the protein was 38.317 which classified it as a stable one.

3.15. Secondary and tertiary structure prediction

Secondary structure of the construct V1 confirmed to have 21.11% alpha helix, 5.08% sheet and 73.81% coil structure (Fig. 6). RaptorX generated the tertiary structure of the designed construct V1 consisting single domain (Fig. 7A and B). Homology modeling was performed by detecting and using 1kj6A from protein data bank (PDB) as best suited template for Vaccine V1. The quality of the 3D model was defined by *P*

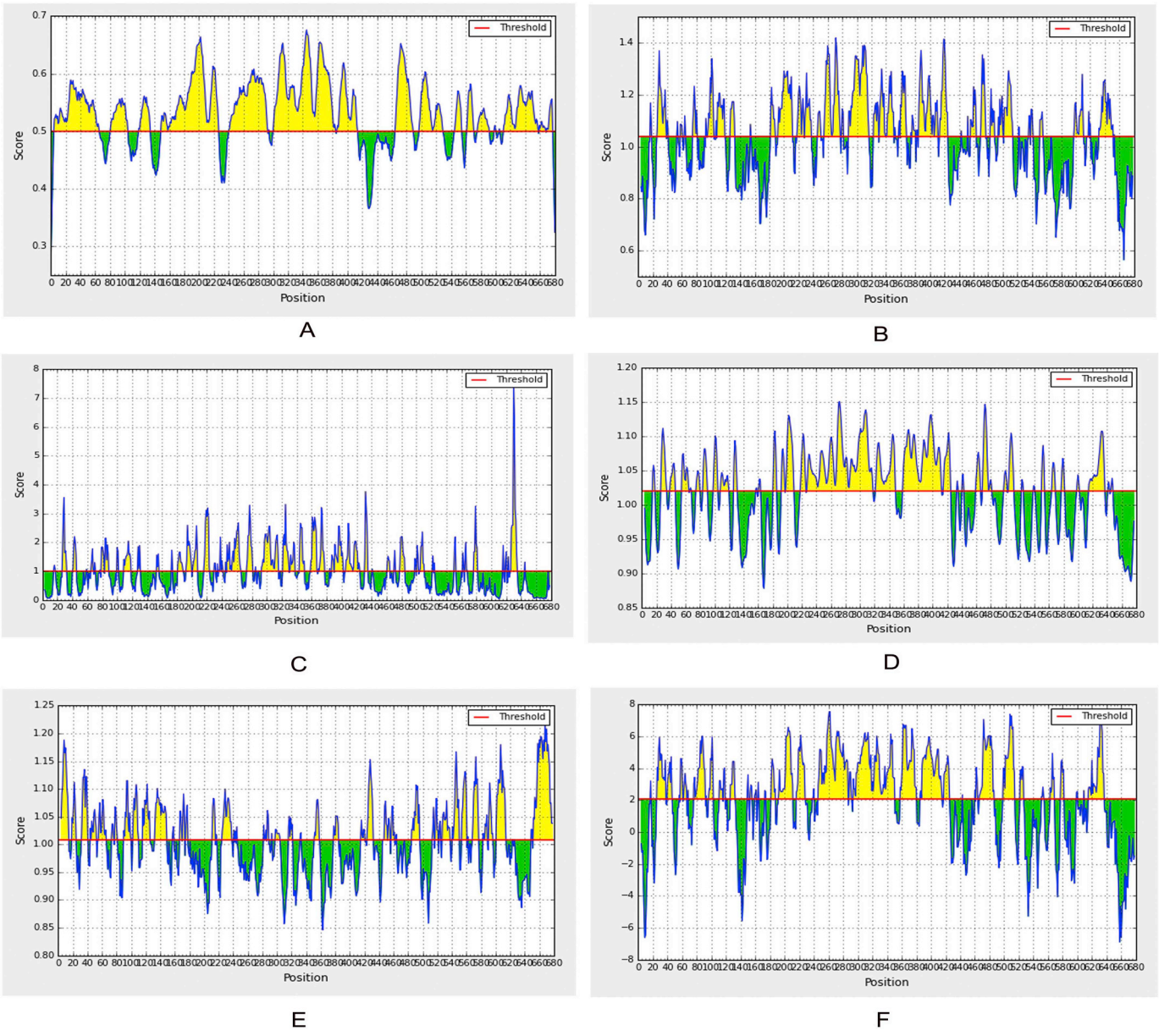

Fig. 4. Prediction of B cell linear epitope and intrinsic properties for envelope glycoprotein using different scales (A: Bepipred, B: Surface accessibility, C: Emini surface, D: Flexibility, E: Antigenicity, F: Hydrophilicity).
Notes: For each graph: x-axis and y-axis represent the position and score; residues that fall above the threshold value are shown in yellow color; the highest peak in yellow color identifies most favored position). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

value which was $2.63e^{-06}$ for the predicted vaccine protein. The low P value ensured better model quality of the predicted vaccine.

3.16. Tertiary structure refinement and validation

Refinement was performed to improve the quality of predicted 3D modeled structure beyond the accuracy. Before refinement Ramachandran plot analysis revealed that 88.9% residues were in the favored, 9.5% residues in the allowed and 1.6% residues in the outlier region. However, after refinement 97.5% and 2.5% residues were in the favored and allowed region respectively, but no residues were found in the outlier region. (Fig. 7C). Modeled tertiary structure of vaccine construct V2 and V3 have been shown in Fig. 8.

3.17. Vaccine protein disulfide engineering

Disulfide engineering was performed by mutating the residues in the highly mobile region of the protein sequence with cystein (Fig. 9). A total 17 pairs of amino acid residues were identified with the capacity to form disulfide bond by DbD2 server. When the residue pairs were evaluate in terms of energy, chi3 and B-factor parameters, only 2 pair satisfied the disulfide bond formation. All 4 residues were replaced with ALA 4 - VAL 118 and LYS 104 - SER 123 cysteine residue. The value of chi3 considered for the residue screening lay between −87 to +97 while the energy value was < 2.5.

3.18. Protein-protein docking

Again, docking analysis was employed between the vaccine constructs and different HLA alleles (Table 9). Construct V1 exhibited

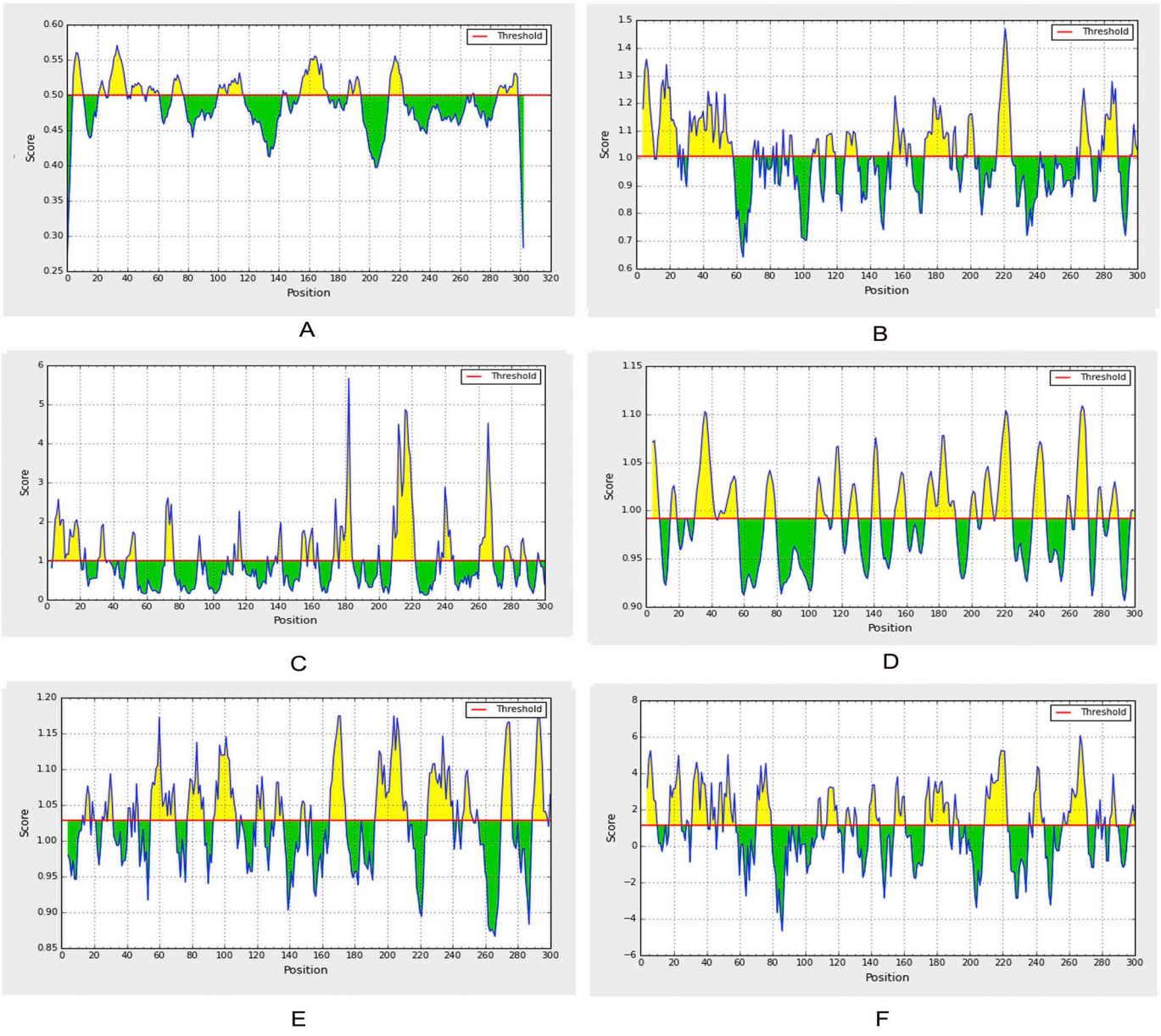

Fig. 5. Prediction of B cell linear epitope and intrinsic properties for matrix protein VP40 using different scales (A: Bepipred, B: Surface accessibility, C: Emini surface, D: Flexibility, E: Antigenicity, F: Hydrophilicity).
Notes: For each graph: x-axis and y-axis represent the position and score; residues that fall above the threshold value are shown in yellow color; the highest peak in yellow color identifies most favored position). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
Allergenicity pattern of the predicted B-cell epitopes generated from envelope glycoprotein and matrix protein VP40.

Protein	Start	End	Length	Peptide	Allergenicity
Envelope glycoprotein	466	489	24	TIFDESSSSGASAEEDQHASPNIS	Non Allergen
	562	571	10	LANQTAKSLE	Allergen
	216	223	8	QEYNSTKN	Non Allergen
	626	634	9	QIKKDEQKE	Non Allergen
	4	14	11	TCFLISLILIQ	Non Allergen
	601	614	14	TCKVLGPDCCIGIE	Allergen
Matrix protein VP40	177	184	8	DAWRPSKD	Non Allergen
	208	222	15	TVKKQAYRQHKNPNN	Allergen
	214	223	10	YRQHKNPNNG	Non Allergen
	29	40	12	PADQLSNQQGIT	Non Allergen
	55	69	15	GNVCHAFTLEAIIDI	Allergen
	165	174	10	TNLVLSVQKL	Non Allergen

Table 8
Allergenicity and antigenicity analysis of the constructed vaccines.

Vaccine Constructs	Composition	Complete Sequence of Vaccine Constructs	Allergenicity	VaxiJen score (Threshold 0.4)
V1	Predicted CTL, HTL & BCL epitopes of envelope glycoprotein and matrix protein VP40 with β defensin adjuvant & PADRE sequence	**EAAAK**GIINTLQKYYCRVRGGRCAVLSCL PKEEQIGKCSTRGRKCCRRKK**EAAAK**AKFVA AWTLKAAA**GGGS**APIDFDPVP**GGGS**GVPPKNV EY**GGGS**PPKNVEYTE**GGGS**VPAWLPLGI**GGGS**YV GDLNLDD**GGGS**QGITPNYVG**GPGPG**NAPIDFDPVPNTK-TI**GPGPG**RWAFRTGVPPKNVEY**GPGPG**VQEDDLAAGLS-WIPF**GPGPG**VPAWLPLGIMSNFEY**GPGPG**GITP-NYVGDLNLDDQ**GPGPG**QGITPNYVGDLNLDD**KK**TIF-DESSSSGASAEEDQHASPNIS**KK**QEYNSTKN**KK**QIKK-DEQKE**KK**TCFLISLILIQ**KK**DAWRPSKD**K-K**YRQHKNPNNG**KK**PADQLSNQQGIT**KK**TNLVLSVQKL**K-K**AKFVAAWTLKAAA**GGGS**	Non Allergen	0.772
V2	Predicted CTL, HTL & BCL epitopes of envelope glycoprotein and matrix protein VP40 with L7/L12 ribosomal protein adjuvant & PADRE sequence	**EAAAK**MAKLSTDELLDAFKEMTLLELSDFVKK FEETFEVTAAAPVAVAAAGAAPAGAAVEAAEEQSEF DVILEAAGDKKIGVIKVVREIVSGLGLKEAKDLVDGAP KPLLEKVAKEAADEAKAKLEAAGATVTVK**EAAAK**AK FVAAWTLKAAA**GGGS**APIDFDPVP**GGGS**GVPPKNV EY**GGGS**PPKNVEYTE**GGGS**VPAWLPLGI**GGGS**YVGD LNLDD**GGGS**QGITPNYVG**GPGPG**NAPIDFDPVPNTK-TI**GPGPG**RWAFRTGVPPKNVEY**GPGPG**VQEDDLAAGLS-WIPF**GPGPG**VPAWLPLGIMSNFEY**GPGPG**GITP-NYVGDLNLDDQ**GPGPG**QGITPNYVGDLNLDD**KK**TIF-DESSSSGASAEEDQHASPNIS**KK**QEYNSTKN**KK**QIKK-DEQKE**KK**TCFLISLILIQ**KK**DAWRPSKD**K-K**YRQHKNPNNG**KK**PADQLSNQQGIT**KK**TNLVLSVQKL**K-K**AKFVAAWTLKAAA**GGGS**	Non Allergen	0.692
V3	Predicted CTL, HTL & BCL epitopes of envelope glycoprotein and matrix protein VP40 with HABA adjuvant & PADRE sequence	**EAAAK**MAENPNIDDLPAPLLAALGAADLALAT VNDLIANLRERAEETRAETRTRVEERRARLTKFQEDLPE QFIELRDKFTTEELRKAAEGYLEAATNRYNELVERGEAAL QRLRSQTAFEDASARAEGYVDQAVELTQEALGTVASQTR AVGERAAKLVGIEL**EAAAK**AKFVAAWTLKAAA**GGGS**A PIDFDPVP**GGGS**GVPPKNVEY**GGGS**PPKNVEYTE**GG GS**VPAWLPLGI**GGGS**YVGDLNLDD**GGGS**QGITP-NYVG**GPGPG**NAPIDFDPVPNTKTI**GPGPG**R-WAFRTGVPPKNVEY**GPGPG**VQEDDLAAGLS-WIPF**GPGPG**VPAWLPLGIMSNFEY**GPGPG**GITP-NYVGDLNLDDQ**GPGPG**QGITPNYVGDLNLDD**KK**TIF-DESSSSGASAEEDQHASPNIS**KK**QEYNSTKN**KK**QIKK-DEQKE**KK**TCFLISLILIQ**KK**DAWRPSKD**K-K**YRQHKNPNNG**KK**PADQLSNQQGIT**KK**TNLVLSVQKL**K-K**AKFVAAWTLKAAA**GGGS**	Non Allergen	0.707

biologically significant results and found to be superior in terms of free binding energy. In addition, docking was also undertaken to determine the binding affinity of the designed vaccine construct with human immune TLR8 receptor using ClusPro, HDOC and PatchDock servers (Fig. 10). The ClusPro server generated 30 protein-ligand complexes as output along with respective free binding energy. The lowest energy of −1249.8 was achieved for the complex 9. The hdoc server hypothesized the binding energy for the protein-protein complex was −278.32, while FireDock output refinement of PatchDock server showed the lowest global energy of −29.00 for solution 9.

3.19. Molecular dynamics simulation

Normal mode analysis (NMA) was performed to investigate the stability of proteins and their mobility at large scale by considering the internal coordinates of the docked complex (Fig. 11A). Vaccine protein V1 and TLR8 were directed towards each other and the direction of each residue in the 3D model was given by arrows. The degree of mobility was indicated by the length of the line. The deformabilty of the complex depends on the individual distortion of each residues, represented by hinges in the chain (Fig. 11D). The eigenvalue found for the complex was $1.0162e^{-04}$ (Fig. 11B). There was an inverse relationship between eigenvalue and the variance associated to each normal mode (Kovacs et al., 2004). The B-factor values derived from NMA, was identical to RMS (Fig. 11C). Coupling between pairs of residues was indicated by the covariance matrix where different pairs showed correlated, anti-correlated or uncorrelated motions, represented by red, blue and white colors respectively (Fig. 11E). An elastic network model was also generated which detected the pairs of atoms connected via springs (Fig. 11F). In the diagram each dot was equivalent to one spring between the corresponding pair of atoms and colored according to the level of stiffness. The darker the grays, the stiffer the springs was.

3.20. Codon adaptation and in silico cloning

Codon adaptation was performed considering the expression system of the host. Construct V1 was reverse-transcribed. In the adapted codons, codon adaptation index (CAI) was 0.980 ensuring the higher proportion of most abundant codons. The GC content of the optimized codons (52.671%) was significant as well. The construct did not contain restriction sites for *Bgl*II and *Apa*I and thus indicating its safety for cloning purpose. Finally, the optimized codons were introduced into pET28a(+) vector along with BglII and ApaI restriction sites. A clone of 5625 base pair was produced comprising 1189 bp desired sequence (shown in red color in between the sequence of pET28a(+) vector) and the rest belonging to the vector only (Fig. 12).

4. Discussion

Marburg virus is known to cause severe hemorrhagic fever in both humans and nonhuman primates with high degree of infectivity and

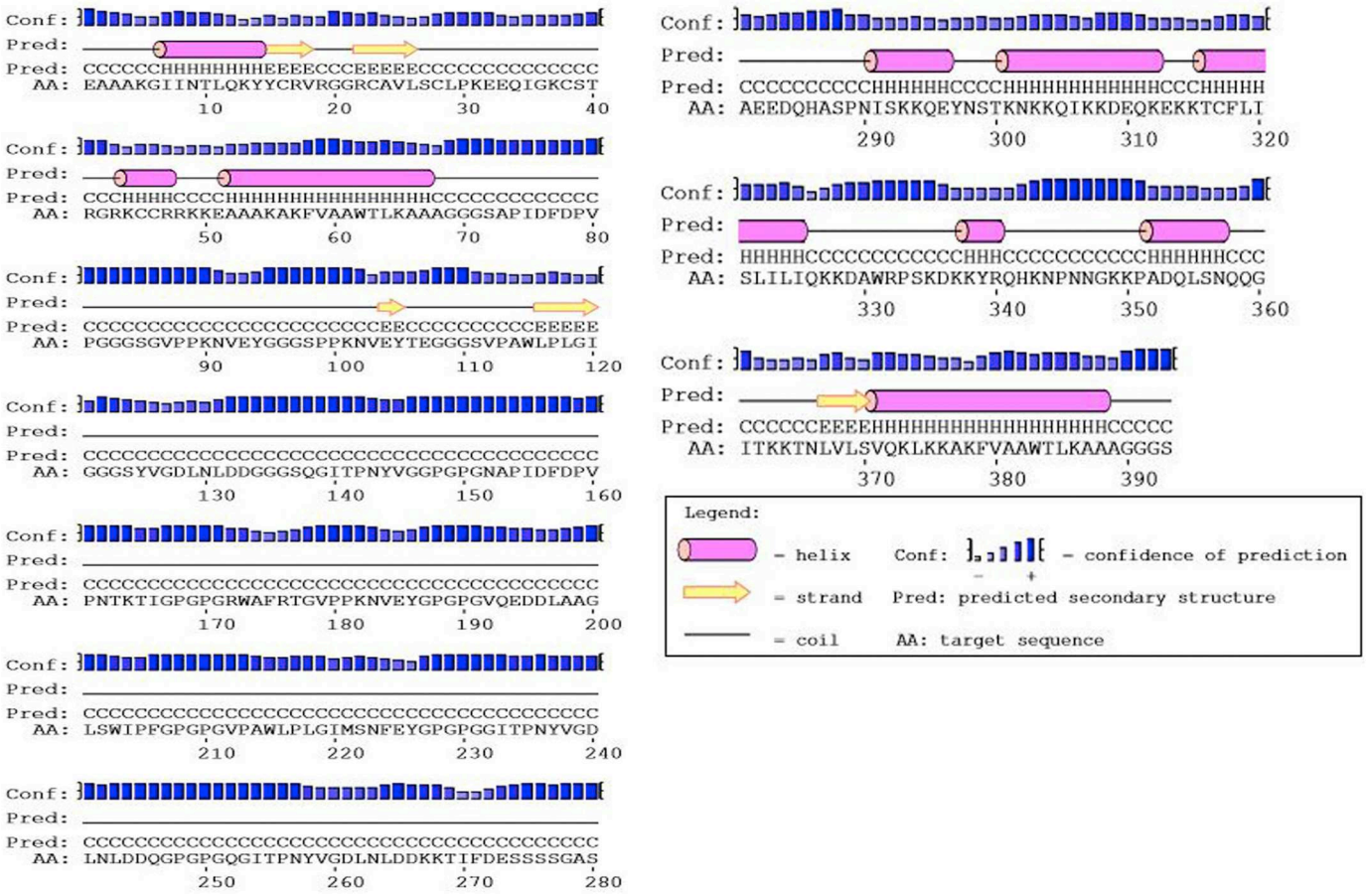

Fig. 6. Secondary structure prediction of designed vaccine V1 using PESIPRED server.

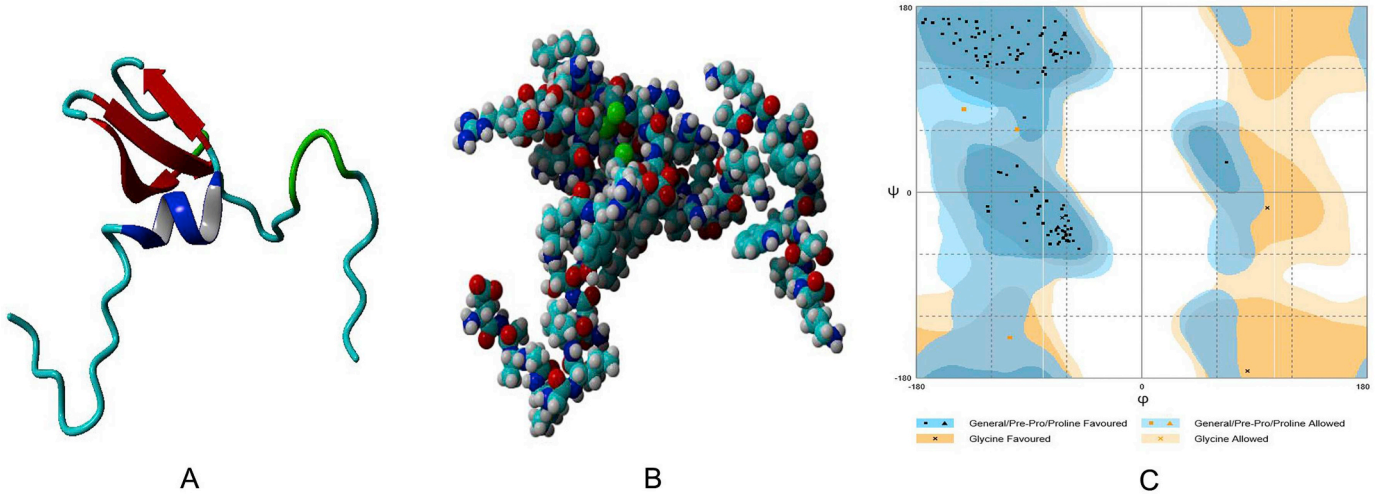

Fig. 7. Tertiary structure prediction and validation of vaccine protein V1, A: Cartoon format, B: Ball structure, C: Validation of the 3D structure of vaccine protein V1 by Ramachandran plot analysis.

lethality (Mehedi et al., 2011). To date no approved treatment is available for Marburg virus infection (Brauburger et al., 2012; Martines et al., 2015). Therefore, it is essential to take preventive measures against it.

Conventional vaccines comprise attenuated or killed agents which depend on adequate antigen expression from in vitro culture models and may take 15 years or more to develop. Microorganisms are often difficult to cultivate and in some instances to attenuate resulting in undesirable or adverse immune responses demonstrating that these approaches are not always feasible (Poland et al., 2009). Again, majority of the methods so far utilized to acquire and purge the target antigen were unsuccessful, resulting in less suitable vaccine candidates (Flower, 2008). For this reason vaccine development is now being shifted to the most efficient and less time consuming pre-screening program (Purcell et al., 2007; Rappuoli, 2000). Recombinant subunit vaccine comprises the antigenic parts only which does not carry any

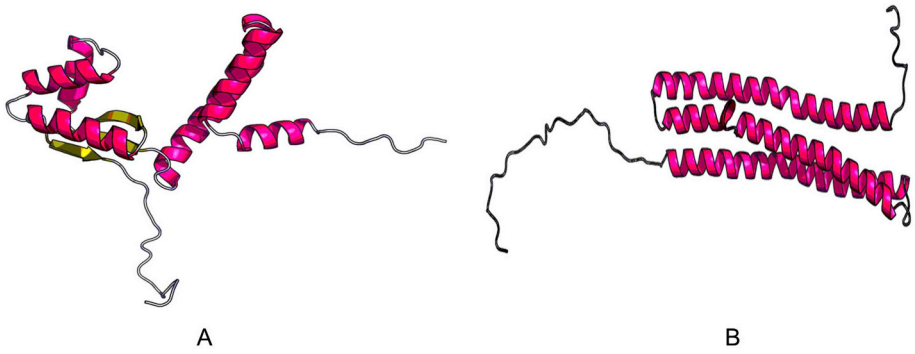

Fig. 8. 3D modeled structure of vaccine protein V2 (A) and V3 (B) generated via RaptorX server.

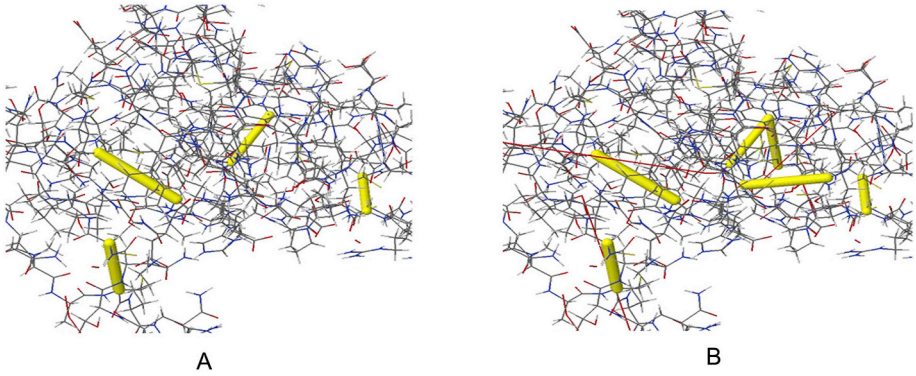

Fig. 9. Disulfide engineering of vaccine protein V1; A: Initial form, B: Mutant form.

Table 9
Binding energy of predicted epitopes with selected MHC class I and MHC class II molecules generated from molecular docking by AutoDock.

Vaccine construct	HLA alleles PDB ID's	Global energy	Hydrogen bond energy	ACE	Score	Area
V1	1A6A	−32.33	−2.94	7.31	13,596	1786.20
	1H15	−40.65	−3.36	−2.26	14,082	1871.30
	2SEB	−48.96	−5.59	7.45	14,252	1658.40
	2Q6W	−22.89	−1.03	3.27	13,676	1952.00
	2FSE	−41.93	−4.28	11.39	15,708	2290.20
	3C5J	−4.70	0.00	0.06	13,934	2145.90
V2	1A6A	−3.62	0.00	−1.58	14,980	2474.00
	1H15	−0.30	−3.32	2.62	16,412	2312.80
	2Q6W	−23.83	−3.61	0.37	15,488	2330.10
	2SEB	−21.05	−1.23	−5.29	16,412	2312.80
	2FSE	−22.44	−4.86	7.72	15,048	1949.90
	3C5J	−36.32	−1.06	4.70	16,002	2337.30
V3	1A6A	−6.51	−1.78	2.33	15,318	2173.20
	1H15	−27.87	−1.55	5.16	15,820	2031.10
	2Q6W	−21.64	−5.43	10.99	14,716	2020.90
	2SEB	−21.05	−1.23	−5.29	16,552	3415.20
	2FSE	−24.57	−2.56	10.28	16,180	1919.90
	3C5J	0.46	−2.12	7.65	20,458	2901.00

risk of emergence of virulence property. Reverse vaccinology, a novel approach to combine immunogenomics and immunogenetics with bioinformatics that has been used tremendously to introduce new vaccines (Mora et al., 2003). This rapid in silico based method has acquired great acceptance with the recent development in the genome and protein sequence databases (Rappuoli et al., 2016; Hasan et al., 2015b) which was employed here to develop a novel multi-epitope monovalent vaccine against Marburg Virus Disease (MVD).

The entire viral proteome of *Marburg marburgvirus* Musoke-80 strain was retrieved from UniProtKB and the physiochemical properties of the proteins were analyzed using ProtParam server. VaxiJen server was used to assess the antigenicity of all the retrieved protein sequences in order to find out the most potent antigenic protein. Among the seven viral proteins, envelope glycoprotein and matrix protein VP40 were identified as the best antigenic protein candidates based on their ability to confer immunity and allowed for further analysis. Vaccine induces production of antibodies that are synthesized by B cells and mediates effector functions by binding specifically to a toxin or a pathogen (Cooper and Nemerow, 1984). Monovalent vaccines are directed against a specific pathogen or organism that have the ability to trigger both B cell and T cell response (Cooper and Nemerow, 1984). Cytotoxic CD8 + T lymphocytes (CTL) play a vital role by recognizing and killing infected cells or secreting specific antiviral cytokines, thus restricting the spread of infectious agents in the body (Hasan et al., 2019; Garcia et al., 1999). Thus, T cell epitope-based vaccination is a unique process of eliciting strong immune response against infectious agents such as viruses (Shrestha and Diamond, 2004).

Top epitopes, bound with the highest number of HLA alleles were selected as putative T cell epitope candidates based on their protein transmembrane topology screening and VaxiJen score (Doytchinova and Flower, 2007). Today, most vaccines stimulate the immune system into an allergic reaction (McKeever et al., 2004). According to the WHO/FAO, if a sequence has an identity of at least six contiguous amino acids over a window of 80 amino acids (0.35% sequence identity) to a known allergens), it is considered to be potentially allergenic. In this study, only the epitopes with non-allergic behavior were allowed for further analysis from both proteins. The result showed that > 90% population of the world can be covered by the predicted T-cell epitopes. As the MHC super-families play a vital role in vaccine design and drug development, MHC cluster analysis was also performed to determine the functional relationship between MHC variants.

To ensure effective binding between HLA molecules and predicted

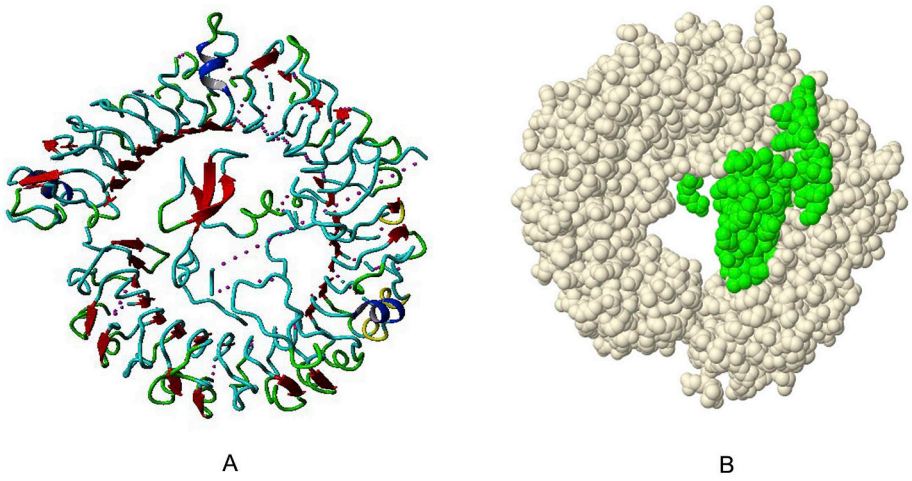

Fig. 10. Docked complex of vaccine construct V1 with human TLR8; A: Cartoon format and B: Ball structure.

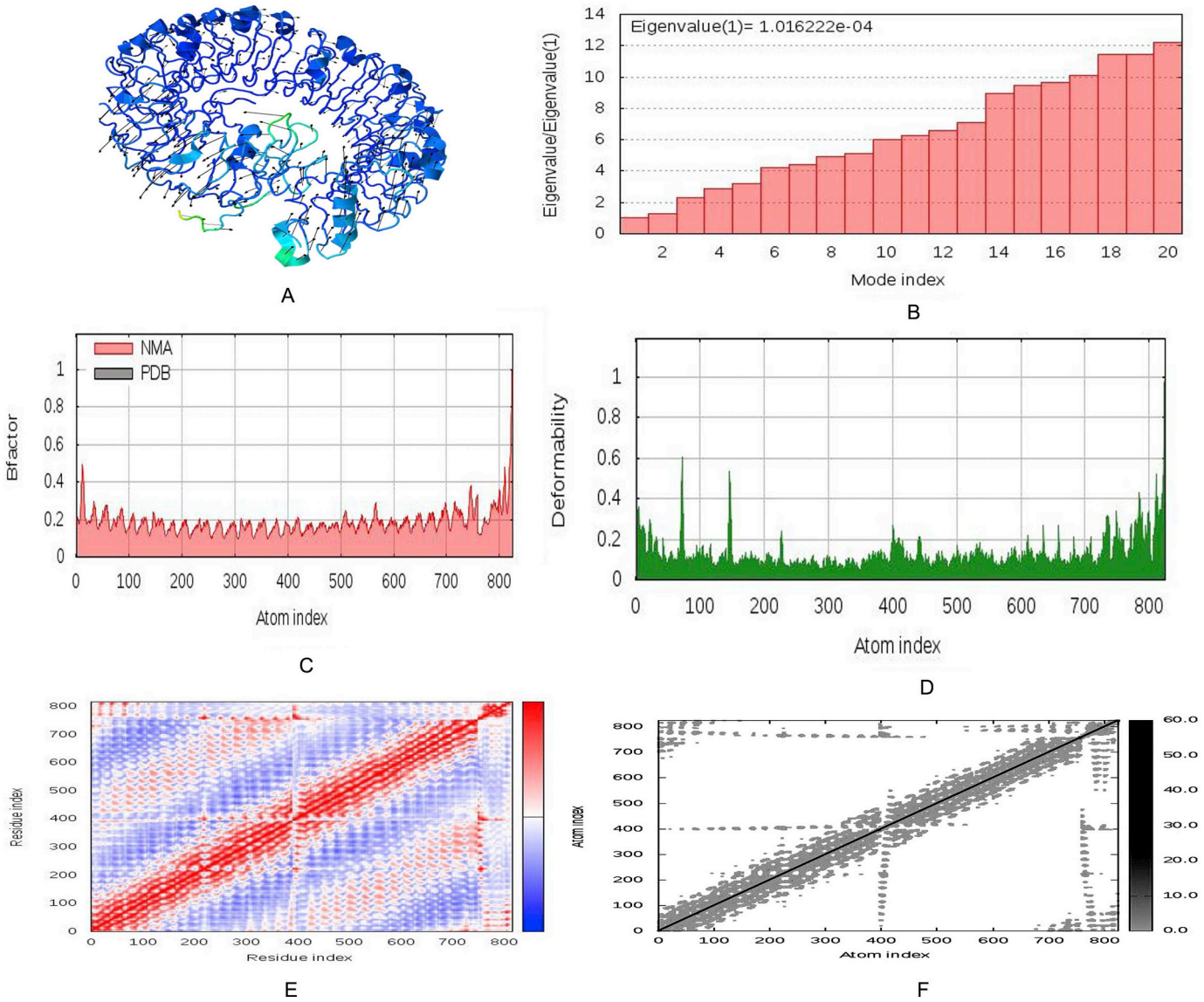

Fig. 11. Molecular dynamics simulation of vaccine protein V1-TLR8 complex. Stability of the protein-protein complex was investigated through mobility (A), eigenvalue (B), B-factor (C), deformability (D), covariance (E) and elastic network (F) analysis.

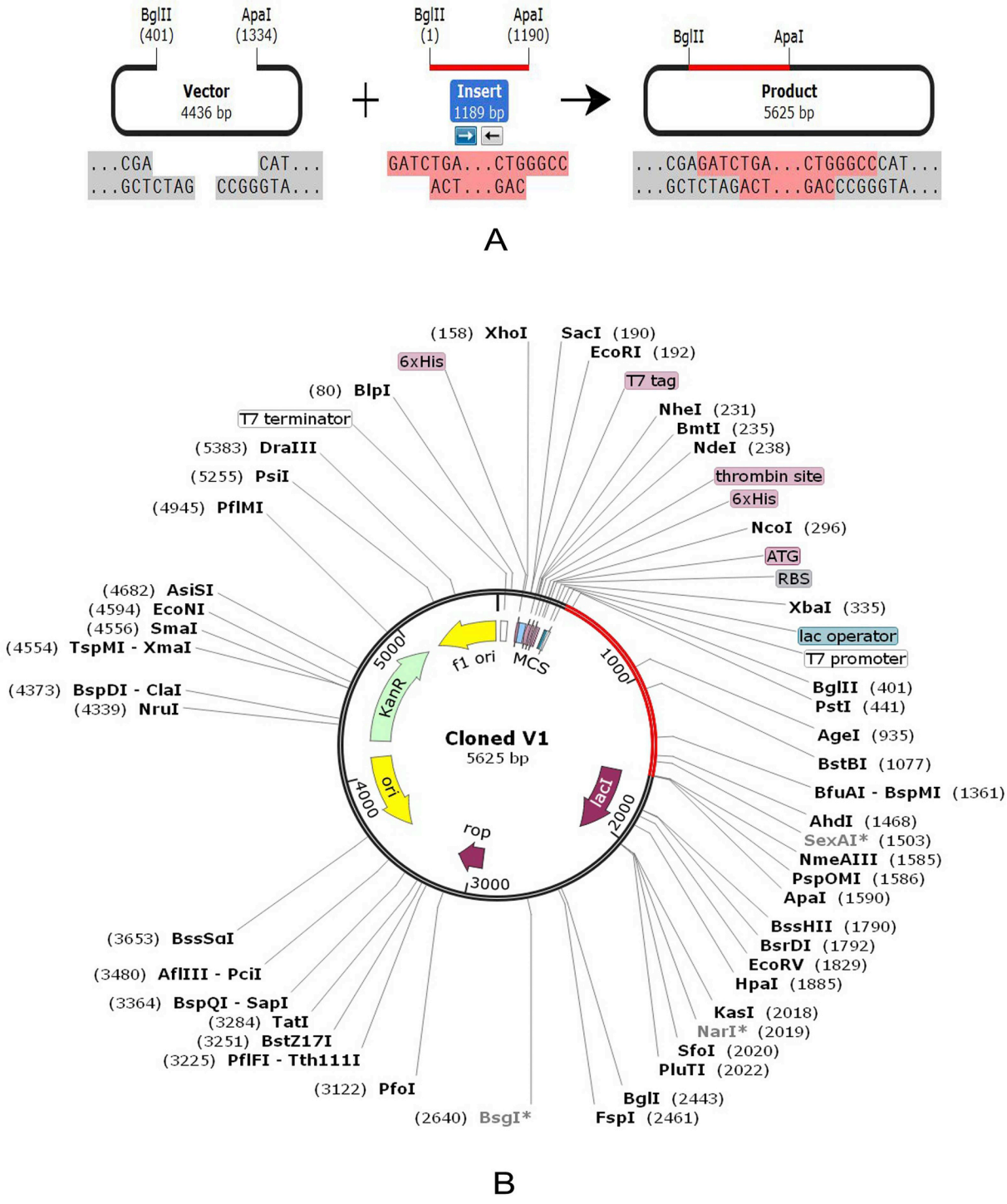

Fig. 12. Restriction digestion (A) and in silico cloning (B) of the gene sequence of final vaccine construct V1 into pET28a(+) expression vector. Target sequence was inserted between *Bgl*II (401) and *Apa*I (1334).

epitopes, a docking study was performed. The suggested 12 T-cell epitopes from both proteins were subjected to PEP-FOLD3 web-based server for 3D structure conversion. HLA-A*11:01 and HLA-DRB1*04:01 was selected for docking analysis with MHC class I and class II binding epitopes respectively. Epitope 'VQEDDLAAGLSWIPF' from envelope glycoprotein was found to be best as it bound in the groove of the HLA-DRB1*04:01 with an energy of −7.8 kcal/mol with significant numbers of hydrogen bonding. Again, VP40-epitope 'VPAWLPLGIMSNFEY' containing the 9-mer core 'VPAWLPLGI' was found best in terms of binding energy (−7.0 Kcal/mol). However, in this study we developed a multi-epitope subunit vaccine to ensure better immune protection. All the finalized epitopes showed a lower binding energy which was biologically significant.

For B-cell epitope prediction, we predicted amino acid scale-based methods for the identification of potential B-cell epitopes using four algorithms from IEDB server. The most potent B cell epitopes for envelope glycoprotein and matrix protein VP40 were identified as vaccine candidates against Marburgvirus. The final vaccine proteins were constructed using the promiscuous epitopes and protein adjuvants along with PADRE sequence. Individual epitopes were linked together via

suitable linker to ensure effective immune response. Literature studies revealed that PADRE containing vaccine construct showed better CTL responses than the vaccines lacked it (Wu et al., 2010).

The constructed vaccines were further checked for their non-allergic behavior and immunogenic potential. Construct V1 was superior in terms of antigenicity and Vaxijen score. The physicochemical properties and secondary structure of V1 was also analyzed before tertiary structure prediction and refinement of 3D model. To strengthen our prediction, we checked the interaction between our vaccine construct with different HLA molecules (DRB1*0101, DRB3*0202, DRB5*0101, DRB3*0101, DRB1*0401, and DRB1*0301). Again construct V1 was found to be best considering the free binding energy. Moreover, docking analysis was also performed to explore the binding affinity of vaccine protein V1 and human TLR8 receptor to evaluate the efficacy of used adjuvant. Molecular dynamics study was conducted to determine the complex stability as well. Structural dynamics had been investigated previously using subsets of atoms and covariance analysis (Aalten et al., 1997). Literature studies linked the stability of macromolecules with correlated fluctuations of atoms (Clarage et al., 1995; Caspar, 1995). Essential dynamics was compared to the normal modes of proteins to determine its stability through iMODS server. The analysis revealed negligible chance of deformability for each individual residues, as location of hinges in the chain was not significant and thereby strengthening our prediction. Finally, the designed vaccine construct V1 was reverse transcribed and adapted for *E. coli* strain K12 prior to insertion within pET28a(+) vector for its heterologous cloning and expression.

Whole-genome analysis of all Marburg viruses in 2006 revealed the existence of five distinct genetic lineages (Kuhn et al., 2010). Four of those lineages differed from each other by only 0–7.8% where the fifth one differed by up to 21% and reclassified as Ravn virus (RAVV) (Kuhn et al., 2010). Genomes of MARV and RAVV diverge from the prototype Marburg virus variant Musoke only by < 10% at the nucleotide level (Towner et al., 2007). To check the broad spectrum efficacy of our designed vaccine models, the conservancy pattern of the top epitopes with RAVV strains was analyzed. The study revealed that 8 of the final 12 T-Cell epitopes of envelop glycoprotein and matrix protein VP40 were conserved within 80 to 100% of the RAVV respective homologous sequence sets. That is why the proposed vaccine candidate is expected to confer immunity against RAVV as well.

The genomic-based technologies continue to transform the field of vaccinology through aiding selection of potential vaccine candidates and facilitating the optimization of the chosen immunogens. In case of conventional vaccines, the virulent form of attenuated vaccine may reverse or inactivation process may fail to suppress the virulence property (Tameris et al., 2013; Merten, 2002). Some recent clinical trials raised safety concerns while emphasizing on conventional approach for vaccine development (Stratton et al., 2002; Hasson et al., 2015; Kaufmann et al., 2014). But subunit vaccine could overcome these undesirable consequences. The predicted in silico vaccine candidates were based on comprehensive analysis of entire viral proteome by using different immune databases and bioinformatics tools which could be explored in further wet lab based analysis using model animals for the experimental the validation.

5. Conclusion

Prevention of newly emerging Marburg virus infections and control during outbreaks is both mandatory and challenging. In silico studies can guide the experimental work for finding the desired solutions with fewer trials and error repeats and thus saving both time and costs for the researchers. In this study, we developed a novel chimeric subunit vaccine (V1) against Marburgvirus using the highly immunogenic epitopes and also suggested the expression system which definitely can confer a good message to the scientific community to carry out further in vivo trials. The study prompts future vaccine development against Marburgvirus as well as other infectious diseases.

Funding

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

Conflict of interest

Authors declare no conflict of interests.

Author contributions

Mahmudul Hasan: Conceptualization, Supervision, Project administration and reviewing.

Kazi Faizul Azim: Experiment Design, Data Handling, Data Analysis, Manuscript writing and draft Preparation.

Aklima Begum, Noushin Anika Khan, Tasfia Saiyara Shammi: Data Handling and Data analysis.

Md. Abdus Shukur Imran, Ishtiak Malique Chowdhury: Data Handling.

Shah Rucksana Akhter Urme: Data Analysis.

Acknowledgments

Authors like to acknowledge the authority of the Bioinformatics Laboratory of Sylhet Agricultural University for the technical support of the project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.03.003.

References

Aalten, D.M.F., Groot, B.L., Findlay, J.B.C., Berendsen, H.J.C., Amadei, A., 1997. A comparison of techniques for calculating protein essential dynamics. J. Comput. Chem. 18, 169–181.

Adams, M.J., Carstens, E.B., 2012. Ratification vote on taxonomic proposals to the international committee on taxonomy of viruses. Arch. Virol. 157, 1411–1422.

Ahluwalia, P.K., Pandey, R.K., Sehajpal, P.K., Prajapati, V.K., 2017. Perturbed microRNA expression by mycobacterium tuberculosis promotes macrophage polarizationleading to pro-survival foam cell. Front. Immunol. 8, 107.

Al-Hakim, M. Hasan, Ali, R., Rabbee, M.F., Marufatuzzahan, Z.F. Joy, 2015. In-silico characterization and homology modeling of catechol 1,2 dioxygenase involved in processing of catechol- an intermediate of aromatic compound degradation pathway. Glob. J. Sci. Front. Res. G Bio-Tech Genet. 15, 1–13.

Awan, F.M., Obaid, A., Ikram, A., Janjua, H.A., 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, 139.

Bausch, D.G., Nichol, S.T., Muyembe-Tamfum, J.J., Borchert, M., Rollin, P.E., Sleurs, H., Campbell, P., Tshioko, F.K., Roth, C., Colebunders, R., et al., 2006. Marburg hemorrhagic fever associated with multiple genetic lineages of virus. N. Engl. J. Med. 355, 909–919.

Biosafety in, 2009. Microbiological and Biomedical Laboratories, 5th ed. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute of Health Washington DC, National Academic Press, USA.

Brauburger, K., Hume, A.J., Muhlberger, E., Olejnik, J., 2012. Forty-five years of Marburg virus research. Viruses 4, 1878–1927.

Buus, S., Lauemoller, S., Worning, P., Kesmir, C., Frimurer, T., Corbet, S., et al., 2003. Sensitive quantitative predictions of peptide-MHC binding by a 'query by Committee'artificial neural network approach. HLA 62, 378–384.

Caspar, D.L.D., 1995. Problems in simulating macromolecular movements. Structure 3, 327–329.

Chou, P., Fasman, G., 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47, 45–148.

Chrysostomou, C., Seker, H., 2014. Prediction of protein allergenicity based on signal-processing bioinformatics approach. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 808–811.

Clarage, J.B., Romo, T., Andrews, B.K., Pettitt, B.M., Phillips, G.N., 1995. A sampling problem in molecular dynamics simulations of macromolecules. Proc. Natl. Acad. Sci. U. S. A. 92, 3288–3292.

Comeau, S.R., Gatchell, D.W., Vajda, S., Camacho, C.J., 2004. ClusPro: a fully automated

algorithm for protein-protein docking. Nucl. Acids Res. 32, 96–99.
Cooper, N.R., Nemerow, G.R., 1984. The role of antibody and complement in the control of viral infections. J. Investig. Dermatol. 83, 121–127.
Craig, D.B., Dombkowski, A.A., 2013. Disulfide by design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinform. 14, 346.
Cros, J., Cagnard, N., Woollard, K., Patey, N., Zhang, S.Y., Senechal, B., Puel, A., Biswas, S.K., Moshous, D., Picard, C., Jais, J.P., 2010. Human CD14 dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. Immunity. 33, 375–386.
Cui, Q., Bahar, I., 2007. Normal mode analysis theoretical and applications to biological and chemical systems. Briefing Bioinform. 8, 378–379.
Das, K., Chakraborty, S., Hasan, M., Rahman, A.M., 2015. In silico analysis to elect superior bacterial alkaline protease for detergent and leather industries. J. Adv. Biotechnol. 5, 685–698.
Dimitrov, I., Bangov, I., Flower, D.R., Doytchinova1, I., 2013. AllerTOP v.2- a server for insilico prediction of allergens. J. Mol. Model. 20, 2278.
Dimitrov, I., Naneva, L., Doytchinova1, I., Bangov, I., 2014. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 30, 846–851.
Dolnik, O., Kolesnikova, L., Becker, S., 2008. Filoviruses: interactions with the host cell. Cell. Mol. Life Sci. 65, 756–776.
Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform. 8, 4.
Emini, E.A., Hughes, J.V., Perlow, D., Boger, J., 1985. Induction of hepatitis a virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55, 836–839.
Fiers, M.W., Kleter, G.A., Nijland, H., Peijnenburg, A.A., Nap, J.P., van Ham, R.C., 2004. Allermatch™, a web tool for the prediction of potential allergenicity according to current FAO/WHO codex alimentarius guidelines. BMC Bioinform. 5, 133.
Flower, D.R., 2008. Bioinformatics for Vaccinology. John Wiley & Sons (1st edition 302).
Garcia, K.C., Teyton, L., Wilson, I.A., 1999. Structural basis of T cell recognition. Annu. Rev. Immunol. 17, 369–397.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., 2003. Protein identification and analysis tools on the ExPASy server. Nucleic Acids Res. 31, 3784–3788.
Grard, G., Biek, R., MuyembeTamfum, J.J., Fair, J., Wolfe, N., Formenty, P., Paweska, J., Leroy, E., 2011. Emergence of divergent Zaire ebola virus strains in Democratic Republic of the Congo in 2007 and 2008. J. Infect. Dis. 204, 776–784.
Grote, A., Hiller, K., Scheer, M., Munch, R., Nortemann, B., Hempel, D.C., Jahn, D., 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 33, W526–W531.
Hajighahramani, N., Nezafat, N., Eslami, M., Negahdaripour, M., Rahmatabadi, S.S., Ghasemi, Y., 2017. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus. Infect. Genet. Evol. 48, 83–94.
Hasan, M., Hakim, A., Iqbal, A., Bhuiyan, F.R., Begum, M.K., Sharmin, S., Abir, R.A., 2015a. Computational study and homology modeling of phenol hydroxylase: Key enzyme for phenol degradation. Int. J. Comput. Bioinfom. 4, 691–698.
Hasan, M., Joy, Z.F., Bhuiyan, E.H., Islam, M.S., 2015b. In Silico characterization and motif election of neurotoxins from Snake venom. Am. J. Biochem. Biotechnol. 11, 84.
Hasan, M., Ghosh, P., Azim, K., Mukta, S., Abir, R.A., Nahar, J., Khan, M.M.H., 2019. Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus. Microb. Pathog. doi.org/10.1016/j.micpath.2019.02.023.
Hasson, S.S., Al-Busaidi, J.K., Sallam, T.A., 2015. The past, current and future trends in DNA vaccine immunizations. Asian Pac. J. Trop. Biomed. 5, 344–353.
Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., Bauer, S., 2004. Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. Science 303, 1526–1529.
Jespersen, M.C., Peters, B., Nielsen, M., Marcatili, P., 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 45, 24–29.
Kallberg, M., Margaryan, G., Wang, S., Ma, J., Xu, J., 2014. RaptorX server: A resource for template-based protein structure modeling. In: Protein Structure Prediction. vol. 1137. pp. 17–27.
Kangueane, P., Nilofer, C., 2018. Protein-Protein Docking: Methods and Tools: Protein-Protein and Domain-Domain Interactions. pp. 161–168.
Karplus, P., Schulz, G., 1985. Prediction of chain flexibility in proteins. Naturwissenschaften 72, 212–213.
Kaufmann, S.H., McElrath, M.J., Lewis, D.J., Giudice, G., 2014. Challenges and responses in human vaccine development. Curr. Opin. Immunol. 28, 18–26.
Knust, B., Schafer, I.J., Wamala, J., Nyakarahuka, L., Okot, C., Shoemaker, T., et al., 2015. Multidistrict outbreak of Marburg virus disease-Uganda, 2012. J. Infect. Dis. 212, 119–128.
Kolaskar, A., Tongaonkar, P.C., 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 276, 172–174.
Kosciolek, T., Jones, D.T., 2014. De novo structure prediction of globular proteins aided by sequence variation-derived contacts. PLoS One 9, e92197.
Kovacs, J., Chacon, P., Abagyan, R., 2004. Predictions of protein flexibility: first order measures. Proteins 56, 661–668.
Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.
Kuhn, J.H., Becker, S., Ebihara, H., Geisbert, T.W., Johnson, K.M., Kawaoka, Y., Lipkin, W.I., Negredo, A.I., Netesov, S.V., Nichol, S.T., Palacios, G., Peters, C.J., Tenorio, A., Volchkov, V.E., Jahrling, P.B., 2010. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. Arch. Virol. 155, 2083–2103.
Kuhn, J.H., Becker, S., Ebihara, H., Geisbert, T.W., Jahrling, P.B., Kawaoka, Y., Netesov, S.V., Nichol, S.T., Peters, C.J., Volchkov, V.E., Ksiazek, T.G., 2011. Virus taxonomy: classification and nomenclature of viruses. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Filoviridae. Elsevier Academic Press, Amsterdam, Netherlands, pp. 665–671.
Lopez-Blanco, J.R., Garzon, J.I., Chacon, P., 2011. iMod: multipurpose normal mode analysis in internal coordinates. Bioinformatics 27, 2843–2850.
Lopez-Blanco, J.R., Aliaga, J.I., Quintana-Orti, E.S., Chacon, P., 2014. iMODS: internal coordinates normal mode analysis server. Nucleic Acids Res. 42, W271–W276.
Lovell, S.C., Davis, I.W., Arendall, W.B., Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2002. Structure validation by Calpha geometry: phi, psi and beta deviation. Proteins 50, 437–450.
Macalino, S., Basith, S., Clavio, N., Chang, H., Kang, S., Choi, S., 2018. Evolution of in Silico strategies for protein-protein interaction drug discovery. Molecules 23, 1963.
MacNeil, A., Farnon, E.C., Morgan, O.W., Gould, P., Boehmer, T.K., Blaney, D.D., Wiersma, P., Tappero, J.W., Nichol, S.T., Ksiazek, T.G., Rollin, P.E., 2011. Filovirus outbreak detection and surveillance: lessons from Bundibugyo. J. Infect. Dis. 204, 761–767.
Martines, R.B., Ng, D.L., Greer, P.W., Rollin, P.E., Zaki, S.R., 2015. Tissue and cellular tropism, pathology and pathogenesis of Ebola and Marburg viruses. J. Pathol. 232, 153–174.
Maupetit, J., Derreumaux, P., Tuffery, P., 2010. A fast method for large scale De novo peptide and miniprotein structure prediction. J. Comput. Chem. 31, 726–738.
McKeever, T.M., Lewis, S.E., Smith, C., Hubbard, R., 2004. Vaccination and allergic disease: a birth cohort study. Am. J. Public Health 94, 985–989.
Mehedi, M., Groseth, A., Feldmann, H., Ebihara, H., 2011. Clinical aspects of Marburg hemorrhagic fever. Future Virol. 6, 1091–1106.
Merten, O.W., 2002. Virus contaminations of cell cultures–a biotechnological view. Cytotechnology 39, 91–116.
Michel, F., 1999. Python: a programming language for software integration and development. J. Mol. Graph. Model. 15, 57–61.
Mora, M., Veggi, D., Santini, L., Pizza, M., Rappuoli, R., 2003. Reverse vaccinology. Drug Discov. Today 8, 459–464.
Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., et al., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791.
Olival, K.J., Hayman, D.T., 2014. Filoviruses in bats: current knowledge and future directions. Viruses 6, 1759–1788.
Pandey, R.K., Sundar, S., Prajapati, Prajapati V.K., 2016. Differential expression of miRNA regulates T cell differentiation and plasticity during visceral leishmaniasis infection. Front. Microbiol. 7, 206.
Parker, J., Guo, D., Hodges, R., 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 25, 5425–5432.
Peng, J., Xu, J., 2011. RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins 79, 161–171.
Petersen, B., Lundegaar, C., Petersen, T.N., 2010. NetTurnP–neural network prediction of beta-turns by use of evolutionary information and predicted protein sequence features. PLoS One 5, e15079.
Poland, G.A., Ovsyannikova, I.G., Jacobson, R.M., 2009. Application of pharmacogenomics to vaccines. Pharmacogenomics 10, 837–852.
Prabhakar, P.K., Srivastava, A., Rao, K.K., Balaji, P.V., 2016. Monomerization alters the dynamics of the lid region in campylobacter jejuni CstII: an MD simulation study. J. Biomol. Struct. Dyn. 34, 778–779.
Purcell, A.W., McCluskey, J., Rossjohn, J., 2007. More than one reason to rethink the use of peptides in vaccine design. Nat. Rev. Drug Discov. 6, 404–414.
Rana, A., Akhter, Y., 2016. Multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. Immunobiology 221, 544–557.
Rappuoli, R., 2000. Reverse vaccinology. Curr. Opin. Microbiol. 3, 445–450.
Rappuoli, R., Bottomley, M.J., D'Oro, U., Finco, O., Gregorio, E., 2016. Reverse vaccinology 2.0: human immunology instructs vaccine antigen design. J. Exp. Med. 213, 469.
Rollin, P.E., 2009. Filoviruses: a compendium of 40 years of epidemiological, clinical, and laboratory studies. Emerg. Infect. Dis. 15, 2079.
Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl. Acids Res. 33, W363–W367.
Shrestha, B., Diamond, M.S., 2004. Role of CD8+ T cells in control of West Nile virus infection. J. Virol. 12, 8312–8321.
Solanki, V., Tiwari, V., 2018. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against Acinetobacter baumannii. Sci. Rep. 8, 9044.
Stratton, K., Almario, D.A., McCormick, M.C. (Eds.), 2002. Institute of Medicine (US) immunization safety review committee. National Academies Press, Washington (DC).
Tama, F., Brooks, C.L., 2006. Symmetry, form, and shape: guiding principles for robustness in macromolecular machines. Annu. Rev. Biophys. Biomol. Struct. 35, 115–133.
Tameris, M.D., Hatherill, M., Landry, B.S., et al., 2013. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. Lancet 381, 1021–1028.
Thaysen-Andersen, M., Packer, N.H., 2012. Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching. Glycobiology 22, 1440–1452.
Thomsen, M., Lundegaard, C., Buus, S., Lund, O., Nielsen, M., 2013. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 65, 655–665.

Timen, A., 2009. Response to imported case of Marburg Hemorrhagic fever, the Netherlands. Emerg. Infect. Dis. 15, 1171–1175.
Towner, J.S., Khristova, M.L., Sealy, T.K., Vincent, M.J., Erickson, B.R., Bawiec, D.A., Hartman, A.L., Comer, J.A., Zaki, S.R., Stroher, U., et al., 2006a. Marburgvirus genomics and association with a large hemorrhagic fever outbreak in Angola. J. Virol. 80, 6497–6516.
Towner, J.S., Khristova, M.L., Sealy, T.K., Vincent, M.J., Erickson, B.R., Bawiec, D.A., et al., 2006b. Marburgvirus genomics and association with a large Hemorrhagic fever outbreak in Angola. J. Virol. 80, 6497–6516.
Towner, J.S., Pourrut, X., Albariño, C.G., Nkogue, C.N., Bird, B.H., Grard, G., Ksiazek, T.G., Gonzalez, J.P., Nichol, S.T., Leroy, E.M., 2007. Marburg virus infection detected in a common African bat. PLoS One 2, e764.
Towner, J.S., Sealy, T.K., Khristova, M.L., Albariño, C.G., Conlan, S., Reeder, S.A., Quan, P.L., Lipkin, W.I., Downing, R., Tappero, J.W., Okware, S., 2008. Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. PLoS Pathog. 4, e1000212.
Towner, J., Amman, B., Sealy, T., et al., 2009. Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. PLoS Pathog. 5, e1000536.
Wang, Z., Eickholt, J., Cheng, J., 2011. APOLLO: a quality assessment service for single and multiple protein models. Bioinformatics 27, 1715–1716.
Will, C., Muhlberger, E., Linder, D., Slenczka, W., Klenk, H.D., Feldmann, H., 1993. Marburg virus gene 4 encodes the virion membrane protein, a type I transmembrane glycoprotein. J. Virol. 67, 1203–1210.
Wu, C.Y., Monie, A., Pang, X., Hung, C.F., Wu, T.C., 2010. Improving therapeutic HPV peptide-based vaccine potency by enhancing CD4+T help and dendritic cell activation. J. Biomed. Sci. 17, 88.
Wuthrich, K., Wagner, G., Rene, R., Werner, B., 1980. Correlations between internal mobility and stability of globular proteins. Biophys. J. 32, 549–558.
Xy, D., Zhang, Y., 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophys. J. 101, 2525–2534.
Zhang, J., Liang, Y., Zhang, Y., 2011. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. Structure 19, 1784–1795.