



Exploiting the reverse vaccinology approach to design novel subunit vaccines against Ebola virus



Md. Asad Ullah, Bishajit Sarkar*, Syed Sajidul Islam

Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Jahangirnagar University, Savar, Dhaka, Bangladesh

ARTICLE INFO

Keywords:
Ebola virus
Vaccine
Immunoinformatics
Bioinformatics
Epitopes

ABSTRACT

Ebola virus is a highly pathogenic RNA virus that causes the Ebola haemorrhagic fever in human. This virus is considered as one of the dangerous viruses in the world with very high mortality rate. To date, no epitope-based subunit vaccine has yet been discovered to fight against Ebola although the outbreaks of this deadly virus took many lives in the past. In this study, approaches of reverse vaccinology were utilized in combination with different tools of immunoinformatics to design subunit vaccines against Ebola virus strain Mayinga-76. Three potential antigenic proteins of this virus i.e., matrix protein VP40, envelope glycoprotein and nucleoprotein were selected to construct the subunit vaccine. The MHC class-I, MHC class-II and B-cell epitopes were determined initially and after some robust analysis i.e., antigenicity, allergenicity, toxicity, conservancy and molecular docking study, EV-1, EV-2 and EV-3 were constructed as three potential vaccine constructs. These vaccine constructs are also expected to be effective on few other strains of Ebola virus since the highly conserved epitopes were used for vaccine construction. Thereafter, molecular docking study was conducted on these vaccines and EV-1 emerged as the best vaccine construct. Afterward, molecular dynamics simulation study revealed the good performances and stability of the intended vaccine protein. Finally, codon adaptation and *in silico* cloning were carried out to design a possible plasmid (pET-19b plasmid vector was used) for large scale production of the EV-1 vaccine. However, further *in vitro* and *in vivo* studies might be required on the predicted vaccines for final validation.

1. Introduction

Highly pathogenic Ebola viruses are single stranded, negative sense RNA viruses and with high mortality rates, this group of viruses causes the Ebola haemorrhagic fever (Nabel et al., 2012; Sanchez et al., 1996; Klenk and Feldmann, 2004). The first recorded Ebola outbreaks were observed in Africa between June and November of 1976 which caused 150 deaths out of 284 infected victims with an astonishing mortality rate of 53 % (Pourrut et al., 2005; Leroy et al., 2005). Ebola virus is a member of the taxonomic viral family, filoviridae (Kiley et al., 1982). The virus contains a negative-sense RNA genome in their virions and the virions are cylindrical or tubular in shape and about 80 nm in diameter. The virions also contain 7–10 nm long glycoprotein spikes projecting from their surfaces. Matrix protein VP40 and nucleocapsid are two important proteins of the Ebola virus since they confer the structural integrity of the virus (Klenk and Feldmann, 2004; Feldmann et al., 1993). Ebola viruses are subdivided into four different subtypes: Zaire, Sudan, Cote d'Ivoire and Reston. The latest discovery includes another African species of human-pathogenic Ebola virus is the

Bundibugyo Ebola virus strain. These subtypes contain many pathogenic strains i.e., Zaire Mayinga-76, Gabon-96, Kikwit-95, Eckron-76 etc. Ebola hemorrhagic viral fever is a severe and often deadly disease of humans and nonhuman primates caused by Ebola virus (Wamala et al., 2010). After the attack of Ebola virus, the immune system produces a systemic inflammatory response that causes the impairment of the vascular, coagulation, asthenia and arthralgia systems, which lead to multi-organ failure resembling septic shock (Feldmann and Geisbert, 2011; Volchkov et al., 1997; Leroy et al., 2000). In 2014, World Health Organization (WHO) organized an emergency meeting attended by 60 researchers and public-health administrators to assess and produce safe and effective Ebola vaccines as soon as possible (World Health Organization, 2014). A recent study shows that a trial vaccine called recombinant vesicular stomatitis virus-Zaire Ebola virus (rVSV-ZEBOV) has been shown to be safe and protective against the Zaire strain and is recommended for use in Ebola outbreaks. The rVSV-ZEBOV is a live attenuated vaccine where the gene of the native envelope glycoprotein is replaced with the envelope glycoprotein of another strain; Kikwit-95 of Ebola virus. However, one potential drawback of the live attenuated

* Corresponding author.

E-mail address: sarkarbishajit@gmail.com (B. Sarkar).

vaccines is their possible reversion to their virulent form which may cause serious problems. The subunit vaccines could be a solution to this problem with no whole viral agent and such subunit vaccines can be declared as safer than the live attenuated vaccines (Regules et al., 2017; Martínez-Romero and García-Sastre, 2015; Choi et al., 2015; Sarkar et al., 2019).

Recently, a few anti-viral therapies have been developed against the Ebola virus i.e., REGN-EB3 and mAb114 based anti-viral drugs; KZ52 monoclonal antibody-based therapy etc. and many of these anti-viral therapies are currently in different stages of pre-clinical and clinical trials (Dyer, 2019; Saphire et al., 2018). Nowadays, researchers are also working to develop cocktails of monoclonal antibodies (MAb) to fight multiple filoviruses. Such cocktails are called pan-Ebola virus antibody preparations which are developed with the target antigen(s) that remain conserved in multiple filovirus species. Scientists have already developed a pan-Ebola virus cocktail preparation which exhibited protection against the Ebola virus and Sudan virus. The preparation neutralizes these two filovirus species with a single antibody (Holtsberg et al., 2016). One of the main limitations of developing vaccines against the Ebola virus is the mucin-like domain of the glycoprotein of the virus, which is responsible for the viral entry into the cells. Moreover, it also obstructs the access of antibodies like anti-MHC class-I and β -integrin antibodies to the epitopes of the glycoprotein as well as other surface proteins. Thus, the mucin-like domain sometimes causes problems in successful vaccination. Researches are currently going on to develop vaccines or antibody therapies targeting the envelope glycoprotein of the Mayinga-76 strain, avoiding the mucin-like domain. According to a study, the Mayinga-76 strain was also responsible for the recent Ebola outbreak in Congo which took many lives. For slowing down the current Ebola outbreaks, all efforts should be made to produce effective and safe vaccines against this deadly disease (Martinez et al., 2011; Kanapathipillai et al., 2014; Patel et al., 2019).

This study was designed to construct vaccines against the Mayinga-76 strain of Ebola virus using different tools of bioinformatics, immunoinformatics and reverse vaccinology (Fig. 1). Moreover, since the epitopes that were used for vaccine construction remain highly conserved across the other strains i.e., Gabon-94 and Kikwit-95, the predicted vaccines are expected to provide immunity towards the infections caused by these two strains also. Reverse vaccinology is an approach of vaccine development where the novel antigens are identified by analyzing the genetic makeup and genomic information of a pathogen or organism or a virus with the aid of different tools of *in silico* biology. This method is a quick, easy and cost-effective way to design vaccine and this technique along with immunoinformatics approaches are widely used for designing vaccines against many viruses i.e., Zika virus, Chikungunya virus etc. However, no commercial and licensed epitope-based vaccine is available till now in the market which has been developed using these bioinformatics approaches (Chong and Khan, 2019; María et al., 2017).

2. Materials and methods

2.1. Strain identification and selection

The strain of the Ebola virus strain Mayinga-76 was identified and selected from National Center for Biotechnology Information or NCBI database (<https://www.ncbi.nlm.nih.gov/>).

2.2. Retrieval of the protein sequences

The viral envelope glycoprotein (accession number: Q05320), matrix protein VP40 (accession number: Q05128) and nucleoprotein (accession number: P18272) were retrieved from the UniProt database (<https://www.uniprot.org/>).

2.3. Antigenicity prediction and physicochemical property analysis of the protein sequences

The antigenicity of the protein sequences was predicted by the online server for antigenicity analysis, VaxiJen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm>); keeping the prediction accuracy parameter threshold at 0.4. The tumor model was used during antigenicity prediction because this model generated excellent results (when compared to other models) in both the leave-on-out cross-validation (LOO-CV) and external validation. Moreover, tumor model is usually used when the vaccines are designed for intended use in human body. The sensitivity, accuracy and specificity of a prediction by the server depends on the prediction threshold which was kept at 0.4 in this experiment for improving the accuracy of the prediction. The server provides user-friendly interface and processes for predicting the antigenicity of the proteins (Doytchinova and Flower, 2007a,b; Doytchinova and Flower, 2008). Thereafter, the selected protein sequences were analyzed by ExPASy's online tool ProtParam (<https://web.expasy.org/protparam/>) to determine their different physicochemical properties i.e., the number of amino acids, theoretical pI, extinction coefficients, half-lives, instability indexes and grand average of hydropathicity (GRAVY) etc. (Gasteiger et al., 2005).

2.4. T-cell and B-cell epitope prediction

The T-cell and B-cell epitopes were predicted using Immune Epitope Database or IEDB (<https://www.iedb.org/>); an online epitope prediction server. The IEDB database has a collection of experimental data on T-cell epitopes as well as antibodies that are collected from different types of experiments conducted on human, non-human primates and other animals (Vita et al., 2019). The MHC class-I restricted CD8+ cytotoxic T-lymphocyte (CTL) epitopes of the selected sequences were predicted using NetMHCpan EL 4.0 prediction method for HLA-A*11-01 allele and the MHC class-II restricted CD4+ helper T-lymphocyte (HTL) epitopes were predicted for HLA DRB1*04-01 allele, where Sturniolo prediction method was used (Sturniolo et al., 1999). Ten of the top twenty MHC class-I and MHC class-II epitopes were randomly selected for each of the proteins based on their percentile scores and antigenicity scores (AS). Five random B-cell lymphocyte (BCL) epitopes for each of the proteins were selected based on their length (the sequences that have ten amino acids or above were selected) and predicted using BepiPred linear epitope prediction method (Larsen et al., 2006).

2.5. Transmembrane topology and antigenicity prediction of the selected epitopes

The transmembrane topology experiment of the epitopes selected in the previous step was carried out using the TMHMM v2.0 server (<http://www.cbs.dtu.dk/services/TMHMM/>). It is a fast and easy tool for determining the transmembrane topology of the epitopes (Möller et al., 2001). The antigenicity of the epitopes was again predicted by the online VaxiJen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm>) server, setting the threshold value to 0.4 which aids in improving the prediction accuracy for the tumor model. The tumor model was used for its best and excellent performances (compared to other models) in the LOO-CV and external validation and it was also selected because the vaccines were designed for use in human (Doytchinova and Flower, 2007a,b; Doytchinova and Flower, 2008).

2.6. Allergenicity and toxicity prediction of the epitopes

The allergenicity of the selected epitopes were predicted using AllerTOP v2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>) as well as AllergenFP v1.0 (<http://ddg-pharmfac.net/AllergenFP/>) servers.

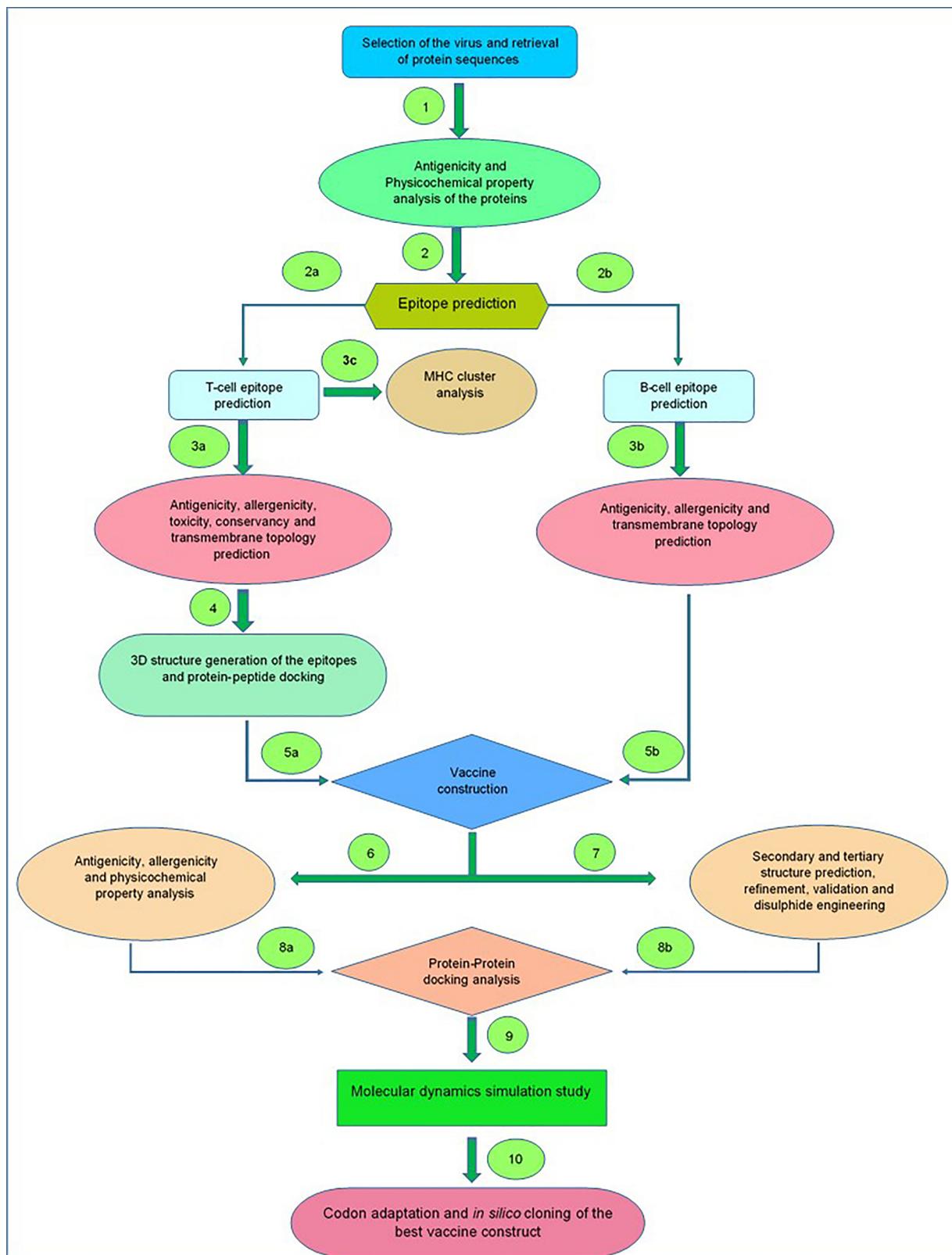


Fig. 1. The flowchart of the procedures carried out in conducting the vaccine designing experiment.

However, more priority was given to the results predicted by AllerTOP because the server has better accuracy of prediction (88.7 %) than AllergenFP server (87.9 %) (Dimitrov et al., 2014, 2013). The toxicity of the selected epitopes was predicted utilizing ToxinPred server (<http://crdd.osdd.net/raghava/toxinpred/>), where support-vector

machine (SVM) (Swiss-Prot) based method was used keeping all the parameters default. The SVM is a machine learning technique which is used by the server for discriminating the toxic and non-toxic peptides. Moreover, the server can also be used for other purposes i.e., peptide designing, protein scanning etc. (Gupta et al., 2013).

2.7. Conservancy analysis of the selected epitopes

IEDB server (<https://www.iedb.org/conservancy/>) was used to determine the conservancy of the selected epitopes setting the sequence identity threshold at ' $> = 50\%$ '. Conservancy of epitopes can be defined as the fraction of protein sequences that contain the epitopes at or above a certain level of identity (Vita et al., 2019; Bui et al., 2007). For the conservancy analysis of the selected epitopes of the Ebola virus strain Mayinga-76, the matrix protein VP40, envelope glycoprotein and nucleoprotein of Ebola virus strains Gabon-94 and Kikwit-95 (UniProt accession numbers: Q05128, Q2PDK5, Q77DJ6, Q05320, O11457, P87666, Q9QCE9, P18272 and O72142) were used for comparison. Considering the results of the antigenicity, allergenicity, toxicity and conservancy tests, the best epitopes were selected for vaccine construction. The T-cell epitopes that were found to be antigenic, non-allergenic, non-toxic and highly (more than 90 %) conserved and had more than 50 % minimum identity, were considered as the best selected epitopes. And the B-cell epitopes that were found to be antigenic as well as non-allergenic were then selected as the best B-cell epitopes for further analysis.

2.8. Cluster analysis of the MHC alleles

The cluster analysis experiment of the MHC alleles is usually carried out to identify the alleles of the MHC class-I and class-II molecules having similar binding specificities. The cluster analysis was performed by an online tool MHCCluster 2.0 (<http://www.cbs.dtu.dk/services/MHCCluster/>) (Thomsen et al., 2013). During the analysis, the number of peptides to be included was kept 50,000 (by default) and the number of bootstrap calculations was set to 100 (by default). All the HLA super-type representatives (MHC class-I) and HLA-DR representatives (MHC class-II) were also selected. The server generated results in the forms of MHC specificity tree and MHC specificity heat-map.

2.9. Generation of the 3D structures of the selected epitopes

The 3D structures of the best selected epitopes were then generated by an online 3D structure generating server, PEP-FOLD3 (<http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/>). Only the best selected epitopes were used for the 3D structure generation (Lamiable et al., 2016; Shen et al., 2014; Thévenet et al., 2012).

2.10. Molecular docking of the best selected epitopes

Molecular docking analysis is one of the necessary steps in reverse vaccinology, which is performed for predicting the binding of epitopes with the antibodies or the MHC receptors (María et al., 2017).

The epitopes that followed the four selection criteria i.e., high antigenicity, non-allergenicity, non-toxicity and high conservancy in the previous steps were considered as the best epitopes which were then subjected to molecular docking experiments with the HLA-A*11-01 allele (PDB ID: 5WJL) and HLA DRB1*04-01 allele (PDB ID: 5JLZ).

The molecular docking was performed by the online docking tool PatchDock (<https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php>) which is a very efficient and effective tool for peptide-protein docking. The tool works on specific algorithms that divide the Connolly dot surface representations of the compounds into concave, convex and also, flat patches. Later, the complementary patches are matched for generating candidate transformations by the server and finally an RMSD (root mean square deviation) clustering score is applied to the candidate solutions to discard the redundant solutions. The refinement and re-scoring of the docking results generated in the previous step were carried out by the FireDock server (<http://bioinfo3d.cs.tau.ac.il/FireDock/php.php>). This tool generates global energies of the best solutions as standard protocol and the lowest global energy is always considered as the best value during the refinement process (Duhovny

et al., 2002; Schneidman-Duhovny et al., 2005; Gujula, 2008; Atapour et al., 2019). Thereafter, the best results were visualized using Discovery Studio Visualizer (Biovia et al., 2016).

2.11. Vaccine construction

Three possible vaccines were constructed against the selected Ebola virus strain Mayinga-76. The predicted CTL, HTL and BCL epitopes were joined together by linkers in order to construct complete vaccine in the order of following sequence: adjuvant, pan HLA-DR epitope (PADRE) sequence, CTL epitopes, HTL epitopes and BCL epitopes. Three different adjuvants i.e., beta defensin, L7/L12 ribosomal protein and HABA protein (*M. tuberculosis*, accession number: AGV15514.1) were used in the vaccine construction. Beta-defensin adjuvants act as agonists and stimulate the activation of the toll like receptors (TLRs) i.e., 1, 2 and 4. The L7/L12 ribosomal protein and HABA protein have the ability to activate TLR-4. EAAAK linkers were used to link the adjuvant and PADRE sequence and GGGG linkers were used to conjugate the CTL epitopes during vaccine construction. Again, GPGPG linkers were used to join the HTL epitopes with other HTL epitopes. Moreover, the KK linkers were used to connect the BCL epitopes (Rana and Akhter, 2016; Hancock et al., 2012; Toussi and Massari, 2014; Funderburg et al., 2007; Lee et al., 2014; Hajighahramani et al., 2017; Pandey et al., 2016). Studies have proved that the PADRE sequence enhances the CTL response of the vaccines that contain the sequence (Wu et al., 2010). Fig. 2 represents a schematic diagram of the three vaccines in their appropriate orientation that were constructed using different epitopes and linkers.

2.12. Antigenicity, allergenicity and physicochemical property analysis

The antigenicity of the constructed vaccines was again determined by the online server VaxiJen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm>) keeping the threshold at 0.4 (Doytchinova and Flower, 2007a,b; Doytchinova and Flower, 2008). Moreover, AlgPred (<http://crdd.osdd.net/raghava/algpred/>) and AllerTop v2.0 (<http://www.ddg-pharmfac.net/AllerTOP/>) servers were used for the allergenicity prediction of the vaccine constructs (Saha and Raghava, 2006). MEME/MAST motif prediction approach was used in the allergenicity prediction by AlgPred server. After that, another online server called ProtParam (<https://web.expasy.org/protparam/>) was used for determining different physicochemical properties of the constructed vaccines (Gasteiger et al., 2005).

2.13. Secondary and tertiary structure prediction of the vaccine constructs

The secondary structures of the vaccine constructs were generated using the online tool PRISPRed (<http://bioinf.cs.ucl.ac.uk/psipred/>). PRISPRed is a fast and straightforward secondary structure generating tool that also predicts the transmembrane topology, transmembrane helix, fold and domain recognition etc. efficiently (Buchan and Jones, 2019; Jones, 1999). The PRISPRed 4.0 prediction method was used to predict the secondary structures of the vaccine constructs. The β -sheet structures of the vaccines were determined by NetTurnP v1.0 (<http://www.cbs.dtu.dk/services/NetTurnP/>) online server (Petersen et al., 2010). Then, the tertiary or 3D structures of the vaccines were generated by RaptorX (<http://raptord.uchicago.edu/>) server (Källberg et al., 2012; Ma et al., 2013; Peng and Xu, 2011).

2.14. 3D structure refinement and validation

Online refinement tool 3Drefine (<http://sysbio.rnet.missouri.edu/3Drefine/>) was used for 3D structure refinement of the constructed vaccines. The protein structures may lack their true and native structure when they are predicted by computational methods. As a result, protein structure refinement is carried out to convert the low resolution

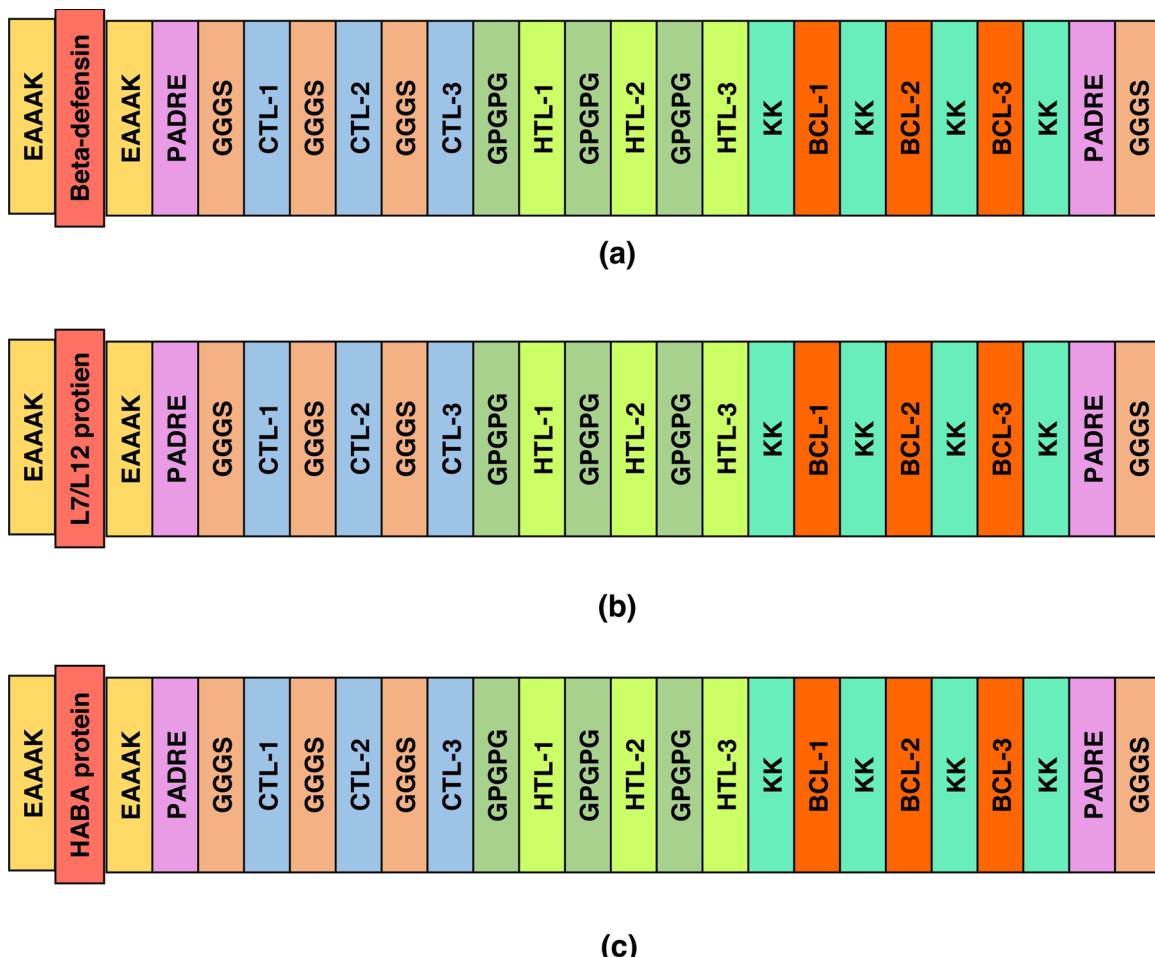


Fig. 2. A schematic representation of the three possible vaccine constructs with linkers (EAAAK, GGGS, GPGPG, KK), PADRE sequence, adjuvants (beta-defensin, L7/L12 protein, HABA protein) and epitopes (CTL, HTL, BCL) in sequential and appropriate manner. (a) is the first vaccine constructed using beta-defensin adjuvant, (b) is the second vaccine constructed using L7/L12 adjuvant protein and (c) is the third vaccine constructed using HABA protein as adjuvant. CTL; cytotoxic T lymphocytic epitope, HTL; helper T lymphocytic epitope, BCL; B cell lymphocytic epitope. The three vaccine constructs differ from each other only in their adjuvant sequences.

predicted model to high-resolution structure which should closely resemble the native structure of the specific protein with better quality. The 3Drefine is an effective and efficient tool that refines the protein structures based on the i3Drefine protocol which can perform the refinement of a 300 amino acids long protein in just 5 min and thus requires only 25 min to build 5 refined models (Bhattacharya et al., 2016; Bhattacharya and Cheng, 2013). The refined model 1 for each of the vaccine was downloaded for validation. Thereafter, the refined vaccine proteins were validated with the aid of Ramachandran plot which was generated by the PROCHECK (<https://services.mbi.ucla.edu/PROCHECK/>) tool (Laskowski et al., 2006; Morris et al., 1992).

2.15. Vaccine protein disulfide engineering

The disulfide connectivity is an important characteristic of proteins as it provides conformational stability to the folded proteins. Protein disulfide engineering was carried out in this experiment in order to confer the stability on the constructed vaccines. The Disulfide by Design 2 v12.2 (<http://cptweb.cpt.wayne.edu/DbD2/>) online server was used for disulfide engineering which is a simple, fast and user-friendly tool for vaccine protein disulfide engineering (Craig DB and Dombkowski, 2013). The intra-chain, inter-chain and C_β for glycine residue were selected during the disulfide engineering experiment. The χ_3 angle was kept -87° or +97° ± 5 because numerous putative disulfides were determined using the default angle (+97° ± 30° and -87° ± 30°) and the

mentioned χ_3 angle was used for identifying a fewer amount of disulfides with desired characteristics. Moreover, C_α-C_β-S_γ angle was kept at its default value of 114.6° ± 10, since the C_α-C_β-S_γ angle was estimated to reach a peak near 115° and covers a range from 105° to 125° in known disulfides (Petersen et al., 1999).

2.16. Protein-protein docking

The constructed Ebola virus vaccines were docked against various MHC alleles as well as TLRs and one best vaccine is usually selected for later experiments based on its performances in the docking experiment. The viral antigens are recognized by the MHC complex when infections occur. Different alleles encode the different segments of the MHC molecules. The vaccines should have good binding affinity with these MHC segments that are encoded by different alleles, so that they can stimulate potential immune responses (Stern and Calvo-Calle, 2009). In this experiment, the vaccines constructs were docked with DRB3*0101 (PDB ID: 2Q6W), DRB1*0401 (PDB ID: 2SEB), DRB3*0202 (PDB ID: 1A6A), DRB5*0101 (PDB ID: 1H15), DRB1*0301 (PDB ID: 3C5J) and DRB1*0101 (PDB ID: 2FSE). TLRs-8 are embedded on the surface of the immune cells which are responsible for generating the immune response against the RNA viruses and TLRs-3 of the immune cells generate immune response against the DNA viruses (Thompson and Iwasaki, 2008; Lester and Li, 2014). Since Ebola virus is an RNA virus, the vaccine constructs were also docked against TLR-8 (PDB ID: 3W3M)

(Regnery et al., 1980). The protein-protein docking was carried out by three different online servers in order to improve the accuracy and reliability of the docking. At first, the docking was carried out by ClusPro 2.0 (<https://cluspro.bu.edu/login.php>). Although the server ranks the clusters of docked complexes based on their lowest energy scores, these scores do not represent the actual binding affinities of the proteins with their targets (Vajda et al., 2017; Kozakov et al., 2017, 2013). The binding affinity (ΔG in kcal mol⁻¹) of docked complexes was then calculated by PRODIGY tool of HADDOCK server (<https://haddock.science.uu.nl/>), where, the lower the binding energy represents the higher the binding affinity (Vangone and Bonvin, 2015; Xue et al., 2016; Yuriev and Ramsland, 2013). Thereafter, the PatchDock (<https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php>) server was used for docking of the vaccines against their target proteins and then the docking results were refined and re-scored by FireDock server (<http://bioinfo3d.cs.tau.ac.il/FireDock/php.php>) (Duhovny et al., 2002; Schneidman-Duhovny et al., 2005; Gujjuila, 2008; Atapour et al., 2019). Again, the docking was carried out using HawkDock server (<http://cadd.zju.edu.cn/hawkdock/>) for the third time and thereafter the Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) score was also calculated using the same server which predicts the result in the context of the affinity score and the lowest score of prediction is considered as the better score (Weng et al., 2019; Hou et al., 2010; Feng et al., 2017; Sun et al., 2014). The HawkDock server generates several models of docked complex and for each of the vaccines and their respective targets, the score of model 1 was selected for analysis and the same model of every complex was also used in the MM-GBSA study. The docked structures were then visualized by PyMol tool (DeLano, 2002).

2.17. Molecular dynamics simulation

The molecular dynamics simulation study was conducted only for the complex (EV-1-TLR-8) containing best vaccine construct that was selected in the previous step. The vaccine construct that showed the best results in molecular docking study was considered as the best selected vaccines. The online server iMODS (<http://imods.chaconlab.org/>) was used for molecular dynamics simulation study which is a fast, online, user-friendly and effective molecular dynamics simulation server for determining and measuring the protein flexibility. This tool predicts different dynamic parameters i.e., deformability, B-factor (mobility profiles), eigenvalues, variance, co-variance map and elastic network of the protein complexes, quite efficiently. The deformability depends on the ability to deform each of total amino acids in a specific protein structure backbone. The eigenvalue represents the motion stiffness of the protein complex and the energy that is required to deform the structure and the lower eigenvalue represents easy deformability of the complex (Awan et al., 2017; Prabhakar et al., 2016; López-Blanco et al., 2014; López-Blanco et al., 2011; Kovacs et al., 2004).

2.18. Codon adaptation and *in silico* cloning

The best predicted vaccine was reverse translated to the possible DNA sequence that is expected to encode the vaccine protein, which was then adapted according to the desired organism so that, the cellular machinery of that particular organism could use the codons of the adapted DNA sequences efficiently for producing the desired vaccine. Codon adaptation is a necessary step of *in silico* cloning because an amino acid can be encoded by more than one codon in different organisms (codon bias). Moreover, the cellular mechanisms of an organism may be completely different from another organism and a codon for a specific amino acid may not work well in different organisms. For this reason, codon adaptation experiment is carried out for predicting the suitable codon that can encode a specific amino acid in a particular organism. The predicted protein sequence of the best selected vaccine was used for codon adaptation by the Java Codon Adaptation Tool or JCat server (<http://www.jcat.de/>). The codon adaptation by the

JCat server ensures maximal expression of a protein in a target organism (in this experiment *E. coli* strain K12 was used) (Grote et al., 2005; Khatoon et al., 2017; Angov, 2011). Prokaryotic *E. coli* strain K12 was selected at the server and at the same time rho-independent transcription terminators, prokaryotic ribosome binding sites and SgrA1 and SphI cleavage sites of restriction enzymes were avoided. The protein sequence was reverse translated to the optimized possible DNA sequence by the JCat server and then the optimized DNA sequence was taken and SgrA1 and SphI restriction sites were attached to the N-terminal and C-terminal sites respectively. Finally, the SnapGene restriction cloning module was used to insert the newly adapted DNA sequence between the SgrA1 and SphI restriction sites of pET-19b vector (Solanki and Tiwari, 2018).

3. Results

3.1. Identification, selection and retrieval of viral protein sequences

The Zaire Ebola virus, strain Mayinga-76 was identified and selected from the NCBI database. Three proteins from the viral structures i.e., envelope glycoprotein (accession number: Q05320), matrix protein VP40 (accession number: Q05128) and nucleoprotein (accession number: P18272) were selected for the possible vaccine construction. The protein sequences were then retrieved from the UniProt database (Supplementary Table S1).

3.2. Antigenicity prediction and physicochemical property analysis of the protein sequences

VaxiJen v2.0 servers were used for antigenicity analysis and physicochemical property prediction of the selected protein sequences. In the physicochemical property analysis, the number of amino acids, molecular weights, theoretical pI, extinction coefficients, half-lives, instability indexes, aliphatic indexes and GRAVY of the three proteins were predicted and analyzed. All the three selected protein sequences were predicted to have good level of antigenicity scores. The matrix protein VP40 had the lowest molecular weight and only envelope glycoprotein was found to be stable according to the prediction. All of the three proteins had a predicted half-life of 30 h in mammalian cells. The matrix protein VP40 also showed the highest predicted theoretical pI, aliphatic index and GRAVY values (Supplementary Table S2). However, further *in vitro* and *in vivo* studies might be required to finally confirm the accuracy of the predictions since all of these predictions are made by online tools.

3.3. T-cell and B-cell epitope prediction and topology determination of the epitopes

The T-cell epitopes of MHC class-I of the three proteins were determined by NetMHCpan EL 4.0 prediction method of the IEDB server using the Sturniolo prediction method (Sturniolo et al., 1999). For each protein, ten of the top twenty epitopes were selected randomly for further analysis since the top twenty epitopes yielded almost similar AS and percentile scores. Moreover, the B-cell epitopes of the proteins were selected using BepiPred linear epitope prediction method of the IEDB server and epitopes were selected based on their higher lengths (Supplementary Figure S1). The topology of the selected epitopes was determined by TMHMM v2.0 server. Supplementary Table S3 and Supplementary Table S4 list the potential T-cell epitopes of matrix protein VP40, Supplementary Table S5 and Supplementary Table S6 list the potential T-cell epitopes of envelope glycoprotein, Supplementary Table S7 and Supplementary Table S8 list the potential T-cell epitopes of nucleoprotein and Supplementary Table S9 list the potential B-cell epitopes with their respective topologies.

Table 1

List of the selected MHC class-I, MHC class-II and B-cell epitopes. For the MHC class-I and class-II epitopes, criteria: high antigenicity, non-allergenicity, non-toxicity, high conservancy. For B cell epitopes, criteria: high antigenicity, non-allergenicity and non-toxicity.

Name of the proteins	MHC class-I epitopes	MHC class-II epitopes	B-cell epitopes
Matrix protein VP40	MVNVISGPK	QKTYSFDSSTAA	RSNSTIARGGNSNTGFLTPESVNGDTPSNPLRPIADDTIDHASHTPGSV
	ILLPNKSGK	KTYSFDSTTAII	NKSGKKGNASDLTSPEK
	GISFHPKLR	WLPLGVADQKTY	KKVTSKNGQP
Envelope glycoprotein	STHINTPVYK	TPQFLLQLNETI	NGVATDVPSATKRWGFSGVPPKVNVYEAGEWAE
	ATQVEQHHR	PQFLLQLNETIY	KKPDGSECLPAAPDGIRG
	SGYYSTTIR	FLLQLNETIYTTS	NGAKNISGQSPARTSSDPGTNTTEDH
Nucleoprotein	TVLDHILQK	SFLMLCLHHAY	AMPEEETTEANA
	VVFSTSDGK	FLLMLCLHHAYQ	APLTNDRRNEPGSTSPRML
	QTNAMVTLR	VVFSTSDGKEYT	VVFSTSDGKEYTYPDSLEEYPPWLKEAMNE

3.4. Antigenicity, allergenicity, toxicity and conservancy analysis

In the antigenicity, allergenicity, toxicity and conservancy analysis, the T-cell epitopes that were predicted to be highly antigenic, non-allergenic, non-toxic, had minimum identity of over 50 % and conservancy of over 90 %, were selected for vaccine construction. Table 1 lists the selected MHC class-I, MHC class-II and B-cell epitopes of the viral proteins. The highly antigenic and non-allergenic sequences were considered as the best B-cell epitopes and these were then used for vaccine construction. Total nine B-cell epitopes (three for each viral protein) were selected for vaccine construction on the basis of selection criteria.

3.5. Cluster analysis of the MHC alleles

The cluster analysis of the possible MHC class-I and MHC class-II alleles that may interact with the predicted epitopes were performed with the help of online tool MHCcluster 2.0. The tool generates the clusters of the alleles in phylogenetic manner. Supplementary Figure S2 illustrates the result of the experiment where the red zone indicates strong interaction and the yellow zone corresponds to weaker interaction.

3.6. Generation of the 3D structures of the epitopes and peptide-protein docking

All the T-cell epitopes were subjected to 3D structure generation by the PEP-FOLD3 server and the 3D structures were then used in the peptide-protein docking experiment. The docking was performed to find out, whether all of the selected epitopes had the ability to bind with the MHC class-I and MHC class-II molecule or not. The selected epitopes were docked with the HLA-A*11–01 allele and HLA DRB1*04–01. Among the MHC class-I epitopes of matrix protein VP40, GISFHPKLR showed the best result with the lowest global energy of -31.68. And SGYYSTTIR generated the lowest and best global energy score of -42.87 among the MHC class-I epitopes of envelope glycoprotein. Moreover, QTNAMVTLR generated the best global energy score of

-45.40 among the MHC class-I epitopes of nucleoprotein. Among the MHC class-II epitopes of matrix protein VP40, KTYSFDSTTAII generated the best global energy score of -24.32 and TPQFLLQLNETI generated the lowest global energy of -2.80 and VVFSTSDGKEYT generated the lowest global energy of -12.06 among the MHC class-II epitopes of envelope glycoprotein and nucleoprotein respectively (Table 2). Supplementary Figure S3 illustrates the best poses between the docked epitopes with their respective targets.

3.7. Vaccine construction

After successful docking, three vaccines were constructed, that is expected to trigger immune response against Ebola virus strain Mayinga-76. Three different adjuvants i.e., beta defensin, L7/L12 ribosomal protein and HABA protein were used to construct three different vaccines. PADRE sequence is a pan HLA-DR epitope peptide that was also used for vaccine construction to enhance their potency with minimal toxicity. The PADRE sequence improves the CTL response as it has the capability to bind with different MHC class-II molecules with high binding affinity (Wu et al., 2010). Moreover, EAAAK, GGGG, GPGPG and KK linkers were used for conjugating the HTL, CTL and BCL epitopes during the vaccine construction. Three different vaccine constructs differed from each other only in their adjuvant sequences and the newly constructed vaccines were designated as: EV-1, EV-2 and EV-3 (Table 3).

3.8. Antigenicity, allergenicity and physicochemical property analysis of the vaccine constructs

The antigenicity prediction of the three vaccine constructs showed that all of the vaccine constructs were possible antigens. Moreover, since all of the vaccine constructs were predicted to be non-allergen, they are expected to be safe for intended use. Thereafter, the number of amino acids, molecular weight, extinction coefficient ($\text{in } \text{M}^{-1} \text{ cm}^{-1}$), theoretical pI, half-life, instability index, aliphatic index and GRAVY were determined in the physicochemical property analysis experiment. All the vaccines were reported to have similar theoretical pI and

Table 2

Results of molecular docking analysis of the selected epitopes.

Name of the protein	Epitope	MHC allele	Global energy	Hydrogen bond energy	Epitope	MHC allele	Global energy	Hydrogen bond energy
Matrix protein VP40	MVNVISGPK	HLA-A*11–01 allele	-27.34	-0.63	QKTYSFDSSTAA	HLA DRB1*04–01	-3.45	0.00
	ILLPNKSGK	(PDB ID: 5WJL)	-31.35	-0.94	KTYSFDSTTAII	(PDB ID: 5JLZ)	-24.32	-3.49
	GISFHPKLR		-31.68	-3.76	WLPLGVADQKTY		-17.07	-0.78
Envelope glycoprotein	STHINTPVYK		-37.49	-3.52	TPQFLLQLNETI		-2.80	-4.87
	ATQVEQHHR		-30.49	-3.87	PQFLLQLNETIY		-0.39	0.00
	SGYYSTTIR		-42.87	-4.99	FLLQLNETIYTTS		-1.78	-3.22
Nucleoprotein	TVLDHILQK		-29.99	-0.87	SFLMLCLHHAY		-8.47	-2.24
	VVFSTSDGK		-34.14	-3.19	FLLMLCLHHAYQ		-4.62	0.00
	QTNAMVTLR		-45.40	-3.31	VVFSTSDGKEYT		-12.06	-0.28

Table 3

List of the vaccines constructed for Ebola virus strain Mayinga-76. The bolded letters represent the linker sequences.

Name of the vaccines	Vaccine constructs	Number of amino acids
Ebola vaccine-1 (EV-1)	EAAAKGIINTLQKYYCRVRRGRCAVSLCLPKEEQIGKCSTRGRKCCRRKKEAAAKAKF-VAAATLKAAGGGSMVNISGPKGGSILLPKSGKGGSISFHPKLRGGGSSTHNTPVYKGGGSATQEQQHHRGGGSSGYYSTTIRGGGSTVLD-HILQKGGGSVFTSDGKGGSQTNAMVTLRGPGQKTYSFDSATAI GPGWLPQLGVADQKTYGPGPGTPQFLQLNETIGPGPGPQFLQLNE-TIYGPGPGLLQLNETIYTSGPGPGSFLMLCLHHAYGPGPGLLMLCLHHAYQGPGPGVVFSTSDGKEYTKRSNSTIARGGSNTGFLTPESVNGDTPSNPLRPIADDTID-HASHTPGSVKKNKSGKGNSADLTSPEKKKKVTKSNGQPKKNGVATDVSATKRWGFRSGVPKVNVYEAGEWAEEKKKKPDG-SATKRWGFRSGVPKVNVYEAGEWAEEKKKKPDGSECLPAAPD GIRGKKGNAKNSQ-SPARTSSDPGTNTTEDHKKAMPEEETTEANAKKAPLTNDNRR-NEPGSTS PRMLKKVVFSTSDGKEYTYPDSLEEYPPWLTEKEAMNEKKAKFVAAWTLKAAGGGS	596
Ebola vaccine-2 (EV-2)	EAAAKMAKLSTDELLDAFKEMTLLLESDFVKKEETFEVTAAAPVAVAAGAAPAGAA-VEAAEQQSEFDVILEAAGDKICVGIVKVRREIVSGLGLKEAKDLVIDGAPKPLLEKVAKEADEA-KAKLEAAGATVTKVEAAAKAKFVAAWTLKAAGGGSMVNISGPKGGS-SILLPNKSGKGGSISFHPKLRGGGSSTHNTPVYKGGGSATQEQQHHRGGGSSGYYST-TIRGGGSTVLDHILQKGGGSVVFSTSDGKGGSQTNAMVTLRGPGQKTYSFDS-TAAGPGPGKTYSFDSATAI GPGWLPQLGVADQKTYGPGPGTPQFLQLNE-TIYGPGPQFLQLNETIYTSGPGPGSFLMLCLHHAYGPGPGLLMLCLHHAYQGPGPGVVFSTSDGKEYTKRSNSTIARGGSNTGFLT-PESVNGDTPSNPLRPIADDTIDHASHTPGSVKKNKSGKGNSADLT-SPEKKKKVTKSNGQPKKNGVATDVSATKRWGFRSGVPKVNVYEAGEWAEEKKKKPDG-SECLPAAPD GIRGKKGNAKNSQSPARTSSDPGTNTTEDHKKAMPEEETTEANAK-KAPLTNDRNRNEPGSTS PRMLKKVVFSTSDGKEYTYPDSLEEYPPWLTEKEAMNEKKAKFVAAWTLKAAGGGS	681
Ebola vaccine-3 (EV-3)	EAAAKMAENPNIIDLPAPLLAALGAADLALATVNDLIANLRERAETRTRVEER-RARLTKFQEDLPEQFIELRDKFTTEELRKAAGYLEAATNRNELVERGEAALQRRLRSQTAFEDASARAEGYVDQAVELTQEALGTVASQTRAVGERAAKLVGIELEAAAKAKFVAAWTLKAAGGGSMVNISGPKGGSILLPNKSGKGGSISFHPKLRGGGSSTHNTPVYKGGGSATQVEQHHRGGGSSGYYSTTIRGGGSTVLDHILQKGGGSVVFSTSDGKGGSQTNAMVTLRGPGP-QKTYSFDSATAI GPGWLPQLGVADQKTYGPGPGTPQFLQLNE-TIYGPGPQFLQLNETIYTSGPGPGSFLMLCLHHAYGPGPGLLMLCLHHAYQGPGPGVVFSTSDGKEYTKRSNSTIARGGSNTGFLT-PESVNGDTPSNPLRPIADDTIDHASHTPGSVKKNKSGKGNSADLT-SPEKKKKVTKSNGQPKKNGVATDVSATKRWGFRSGVPKVNVYEAGEWAEEKKKKPDG-SECLPAAPD GIRGKKGNAKNSQSPARTSSDPGTNTTEDHKKAMPEEETTEANAK-KAPLTNDRNRNEPGSTS PRMLKKVVFSTSDGKEYTYPDSLEEYPPWLTEKEAMNEKKAKFVAAWTLKAAGGGS	710

ext. coefficient (EV-3 had the lowest value of $55475 \text{ M}^{-1} \text{ cm}^{-1}$) and all of them were found to be stable. Moreover, all of the vaccine constructs were also predicted to have the similar half-life of 1 h in the mammalian cells. EV-2 was reported to have the highest GRAVY value of -0.504 among the three vaccine constructs. The antigenicity, allergenicity and physicochemical property analysis of the three vaccine constructs are listed in Supplementary Table S10.

3.9. Secondary and tertiary structure prediction of the vaccine constructs

The secondary structure analysis of the vaccine constructs was carried out to determine the percentage of the amino acids in the alpha-helix, beta-sheet and coil formation within the vaccine structure. EV-3 had the highest percentage of the amino acids (35.2 %) in the coil formation and EV-2 the highest percentage of amino acids (53.2 %) in the beta-strand formation. However, EV-1 had the highest percentage of

amino acids (14.5 %) in the alpha-helix formation (Supplementary Figure S4 and Table 4).

The 3D structures of the vaccine constructs were predicted by the online server RaptorX. All the three vaccines had 5 domains and EV-1 had the lowest p-value of 4.94e-12. The p-value estimates the relative quality of a protein model and the smaller p-value corresponds to higher quality of the protein model (Källberg et al., 2012; Ma et al., 2013; Peng and Xu, 2011). As a result, EV-1 vaccine was predicted to be the best quality 3D model among the three Ebola vaccine constructs. The homology modeling of the three Ebola vaccine constructs was carried out using 5HJ3C as template from protein data bank to predict or determine the 3D structures of the query proteins by the RaptorX server (Källberg et al., 2012). The results of the 3D structure analysis are listed in Supplementary Table S11 and illustrated in Supplementary Fig. S5.

Table 4

Results of the secondary structure analysis of the vaccine constructs.

Name of the vaccine	Alpha helix (percentage and number of amino acids)	Beta sheet (percentage and number of amino acids)	Coil structure (percentage and number of amino acids)
EV-1	19.8% (118)	65.6 % (391)	14.5 % (87)
EV-2	27.8% (189)	58.0 % (395)	14.2 % (97)
EV-3	35.2 % (250)	53.2 % (378)	11.5 % (82)

3.10. 3D structure refinement and validation

The three Ebola vaccine constructs were refined and validated in the 3D structure refinement and validation step. The Ramachandran plot generated by the PROCHECK server is divided into four regions i.e., the most favored region (represented by red color), the additional allowed region (represented by yellow color), the generously allowed region (represented by light yellow color) and the disallowed region (represented by white color). According to PROCHECK server, a valid protein with the best quality should have over 90 % of its amino acids in the most favored region. The additional allowed region and generously allowed region might contain some percentage of the amino acids of protein, but the highest percentage should reside within the most favored region of a high quality protein structure. Moreover, no amino acid should reside in the disallowed region (Sateesh et al., 2010; Laskowski et al., 1993; Zobayer et al., 2011).

The 3D structure refinement (by 3Drefine server) and validation (by PROCHECK server) analysis showed that EV-1 vaccine had 65.0 % of the amino acids in the most favored region, 31.1 % of the amino acids in the additional allowed regions, 3.3 % of the amino acids in the generously allowed regions and 0.7 % of the amino acids in the disallowed regions. The EV-2 vaccine had 66.0 % of the amino acids in the most favored region, 29.9 % of the amino acids in the additional allowed regions, 3.4 % of the amino acids in the generously allowed regions and 0.7 % of the amino acids in the disallowed regions. The EV-3 vaccine had 72.1 % of the amino acids in the most favored regions, 24.0 % of the amino acids in the additional allowed regions, 2.7 % of the amino acids in the generously allowed regions and 1.2 % of the amino acids in the disallowed regions (Supplementary Figure S6). And thus, EV-3 showed the best result in the Ramachandran plot analysis among the three vaccine constructs.

3.11. Protein disulfide engineering

In this step, possible disulfide bonds were predicted for the 3D structures of the three vaccine constructs. The DbD2 server identifies the pairs of amino acids that have the capability to form disulfide bonds based on the adjusted selection criteria. In this experiment, only those amino acid pairs that had bond energy valueless than 2.00 kcal/mol were selected to get predictions which are close to the native state because according to a study with proteins, about 90 % of the native disulfide bonds have energy value of less than 2.2 kcal/mol (Craig and Dombkowski, 2013). The EV-1 generated 28 amino acid pairs that had the capability to form disulfide bonds. However, only 6 pairs i.e., 20 Gly and 127 Gln, 59 Val and 79 Gly, 101 Ser and 114 His, 195 Ser and 238 Pro, 292 Gly and 293 Ser, 318 His and 364 Gly were selected since they had bond energy less than 2.00 kcal/mol. EV-2 generated 29 pairs of amino acids that had the capability to form disulfide bonds and only 4 pairs were selected i.e., 60 Glu and 67 Glu, 68 Phe and 118 Ala, 240 His and 255 Asp, 498 Val and 560 Arg. EV-3 generated 28 pairs of amino acids capable of forming disulfide bonds and only 5 pairs of the amino acids were selected i.e., 214 Ile and 261 Gly, 229 Asn and 293 Asn, 332 Ala and 344 Val, 438 Gly and 439 Pro, 459 Thr and 476 Val. The selected amino acid pairs formed the mutant version of the original vaccines in this experiment (Supplementary Figure S7).

3.12. Protein-Protein docking study

The protein-protein docking of the vaccine constructs and the MHC alleles were carried out by several online tools for enhancing the accuracy of the prediction: i.e., ClusPro 2.0, PatchDock and HawkDock server. The docked complexes that were generated by ClusPro 2.0 and PatchDock tools were further analyzed by PRODIGY tool of HADDOCK webserver and FireDock server respectively. The PRODIGY tool predicts the binding affinity score (in kcal/mol) whereas FireDock predicts the global energy of the docked complexes. However, HawkDock generates

ranking scores along with the binding free energy (in Kcal/mol). The binding free energy was calculated after the MM-GBSA scoring in the HawkDock server.

EV-1 showed the best results among the three vaccines in the docking experiment carried out by the ClusPro 2.0 and PRODIGY servers. It had the lowest binding affinity when docked with DRB3*0202 (-17.3 kcal/mol), DRB1*0101 (-19.2 kcal/mol), DRB3*0101 (-19.2 kcal/mol), DRB1*0401 (-21.5 kcal/mol), DRB1*0301 (-17.9 kcal/mol) and TLR-8 (-21.9 kcal/mol). When the docking was performed by the PatchDock server, EV-3 showed the best performances with the lowest score while docking with DRB3*0202 (-19.59), DRB5*0101 (-23.35), DRB3*0101 (-9.89) and TLR-8 (-22.21). However, EV-1 also showed the best performances and lowest scores when docked with all the selected MHC alleles by the HawkDock server and also when analyzed in the MM-GBSA study. As a consequence, EV-1 was selected as the best vaccine among the three constructs since it showed the best values among the three vaccines in most of the steps of docking studies (Fig. 3). Table 5 lists the results of the docking study of the three Ebola vaccine constructs.

3.13. Molecular dynamics simulation

Fig. 4 illustrates the results of molecular dynamics simulation and normal mode analysis (NMA) of EV-1-TLR-8 docked complex. The molecular dynamics simulation study is an important experiment to be carried out in vaccine designing by reverse vaccinology since it is performed to determine the stability and physical movements of atoms and molecules of the best selected vaccine construct (Chauhan et al., 2019). For this reason, the simulation study was conducted in this study to determine the movements of molecules and atoms in the EV-1 vaccine construct. The deformability graph of the complex illustrate the peaks in the graphs which represent the regions of the protein with deformability (Fig. 4 b). The B-factor graph of the complexes provide an easy understanding and visualization of the comparison between the NMA and the PDB field of the docked complex (Fig. 4 c). The graph of eigenvalue of the complex is illustrated in Fig. 4 d. EV-1 and TLR-8 docked complex generated an eigenvalue of 2.129193e-05. The variance graph indicates the individual variance by red colored bars and cumulative variance by green colored bars (Fig. 4 e). Fig. 4 f illustrates the co-variance map of the complex where the correlated motion between a pair of residues is indicated by red color, uncorrelated motion is indicated by white color and anti-correlated motion is indicated by blue color. EV-1 had a good number of correlated amino acids (marked in red color) and also showed a large number of stiffer regions (marked in darker gray color). The elastic map of the complex represents the connection between the atoms and darker gray regions indicate stiffer regions (Figure 04g) (López-Blanco et al., 2014; López-Blanco et al., 2011; Kovacs et al., 2004).

3.14. Codon adaptation and in silico cloning

For *in silico* cloning and plasmid construction, the protein sequence of the best selected vaccine from docking study; EV-1 was adapted by the JCcat server. Since EV-1 vaccine had 596 amino acids, the reverse translation of the protein sequence generated a DNA sequence of 1788 nucleotides. The codon adaptation index (CAI) value of 0.943 of EV-1 indicated that the DNA sequence contained higher proportion of the codons that are most likely to be present and used in the cellular machinery of the target organism *E. coli* strain K12 (codon bias). And as a consequence, the production of the EV-1 vaccine is expected to be efficient inside the host organism (Carbone et al., 2003; Sharp and Li, 1987). The GC content of the improved sequence was 50.73 %. The predicted DNA sequence of EV-1 was inserted into the pET-19b vector plasmid between the SgrAI and SphI restriction sites. Since the DNA sequence did not have restriction sites for SgrAI and SphI restriction enzymes, both of the restriction sites were conjugated at the N-terminal

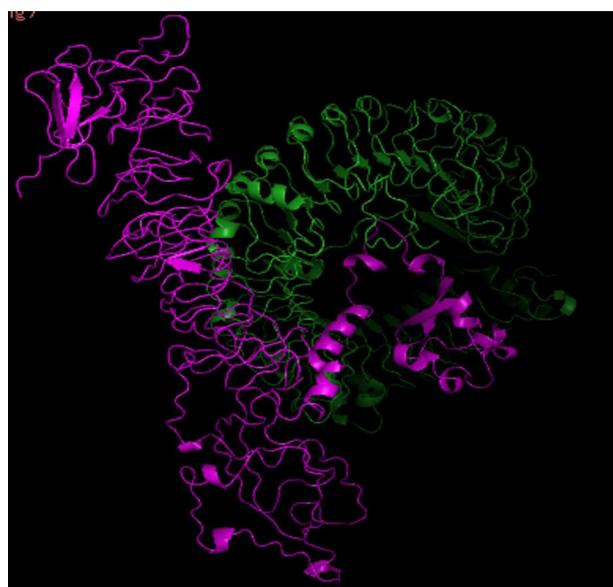


Fig. 3. Figure showing the interaction between the ligand protein, EV-1 and receptor protein, TLR-8. The ligand protein is indicated by pink color and the receptor protein is indicated by green color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 5
Results of the docking study of all the vaccine constructs.

Name of the vaccines	Name of the Targets	PDB IDs of the targets	Binding affinity, ΔG (kcal mol ⁻¹)	Global energy	HawkDock score (the lowest score)	MM-GBSA free energy, in kcal mol ⁻¹)
EV-1	DRB3*0202	1A6A	-17.3	-3.60	-6220.65	-58.62
	DRB5*0101	1H15	-19.8	1.34	-6003.66	-98.32
	DRB1*0101	2FSE	-19.3	-22.18	-7455.95	-126.19
	DRB3*0101	2Q6W	-19.2	8.86	-6347.92	-74.11
	DRB1*0401	2SEB	-21.5	-33.85	-5793.01	-60.95
	DRB1*0301	3C5J	-17.9	11.21	-6835.73	-69.93
	TLR8	3W3M	-21.9	-2.29	-6148.16	-66.70
EV-2	DRB3*0202	1A6A	-16.9	-18.38	-4926.81	-1.94
	DRB5*0101	1H15	-19.8	-13.53	-4935.45	-71.05
	DRB1*0101	2FSE	-19.2	-18.66	-5684.59	-52.92
	DRB3*0101	2Q6W	-18.9	-2.82	-6046.53	-42.83
	DRB1*0401	2SEB	-21.5	-40.90	-5243.99	-16.20
	DRB1*0301	3C5J	-17.5	-0.08	-4720.13	-36.24
	TLR8	3W3M	-20.7	-5.71	-5295.64	-27.63
EV-3	DRB3*0202	1A6A	-17.1	-19.59	-5205.55	-54.82
	DRB5*0101	1H15	-19.8	-23.35	-4481.78	-52.61
	DRB1*0101	2FSE	-19.3	-21.01	-5585.34	-85.44
	DRB3*0101	2Q6W	-19.0	-9.89	-4905.59	-23.38
	DRB1*0401	2SEB	-21.3	-10.08	-5048.95	-46.65
	DRB1*0301	3C5J	-17.0	9.69	-5153.60	-47.36
	TLR8	3W3M	-21.2	-22.21	-5773.78	-28.08

and C-terminal sites respectively before inserting the sequence into the plasmid pET-19b vector. The newly constructed cloned plasmid was 7360 base pair long including the constructed DNA sequence of the EV-1 vaccine along with the restriction sites (Supplementary Figure S8 and Fig. 5).

4. Discussions

Although, vaccine is one of the most important and widely produced pharmaceutical products in the world, the development and production

of vaccine is a costly process. Beyond this, sometimes it takes many years to develop a proper vaccine candidate against a particular pathogen. In modern times, different methods and tools of bioinformatics, immunoinformatics and reverse vaccinomics are used for vaccine development which save both time and cost of the vaccine development process (María et al., 2017). Conventional vaccines are developed based on inactivated or killed antigens or pathogens and such processes sometimes take up to 15 years or more to develop a successful vaccine. Moreover, adverse immune responses that result from the improper attenuation of the pathogens is also a potential drawback of conventional vaccines. Such difficulties in the development of conventional vaccines have led the way to develop many computer programs and software in the recent years which can be effectively used in designing and developing potential vaccines in a very short time. These software and database-based methods also have acquired great acceptance among the scientific community in the recent times (Poland et al., 2009; Purcell et al., 2007; Rappuoli et al., 2016; Rappuoli, 2000). In this study the rapidly accepted *in silico*-based methods were exploited to design possible potential vaccines against the Mayinga-76 strain of Zaire Ebola virus

The specific viral strain of the Ebola virus was identified and selected from the NCBI database. Three proteins of the Mayinga-76 strain i.e., matrix protein VP40, envelope glycoprotein and nucleoprotein were selected as the potential targets for the vaccines. The matrix protein VP40 of Ebola virus plays very important roles in the life cycle of a virus in regulating the viral transcription, virion assembly and budding of viruses from the infected cells (Madara et al., 2015). Moreover, the envelope glycoprotein of the Ebola virus mediates the proper viral attachment and entry into the target cell (Martinez et al., 2013). Studies have also revealed the important roles of nucleoprotein in the replication of Ebola virus (Watanabe et al., 2006). Since these three proteins play essential roles in the infection, proliferation and life cycle of the Ebola virus therefore, these proteins are potential targets for vaccine development aiming to block the viral replication and proliferation.

Primarily, the three protein sequences i.e., envelope glycoprotein, matrix protein VP40 and nucleoprotein, were retrieved from the UniProt (<https://www.uniprot.org/>) database. Different, physicochemical properties i.e., number of amino acids, molecular weight, theoretical pI, extinction co-efficient, instability index, aliphatic index, GRAVY were then determined using ProtParam server. The instability index of a compound refers to the probability of that particular compound to be stable and if a compound has instability index of over 40, then that particular compound is considered to be unstable (Guruprasad et al., 1990). The extinction coefficient refers to the amount of light that is absorbed by a compound at a certain wavelength (Pace et al., 1995; Gill and Von Hippel, 1989). The aliphatic index of a protein can be defined as the relative volume occupied by the aliphatic amino acids i.e., alanine, valine etc. in the side chains (Ikai, 1980). Only the envelope glycoprotein was found to be the stable protein with an instability index score of 38.36. Envelope glycoprotein was also predicted to absorb the highest amount of light at a certain wavelength as it was found to have the highest extinction co-efficient of 74464.46 M⁻¹ cm⁻¹. However, matrix protein VP40 was reported to have the highest amount of aliphatic amino acids in its side chains which was reflected by the highest score of 96.32, among the three selected proteins. Theoretical pI is also an important physicochemical property of the proteins which depicts the pH at which the proteins should have no net charge (Shi et al., 2005). The GRAVY value for a peptide or protein is represented as the sum of hydropathy values of all the amino acids which is then divided by the number of residues in the sequence (Kyte and Doolittle, 1982). Matrix protein VP40 had the highest GRAVY value. Envelope glycoprotein showed best results for most of the parameters of the physicochemical property analysis

All the proteins were found to be quite antigenic in the bacteria, parasitic and tumor model. The three proteins showed excellent

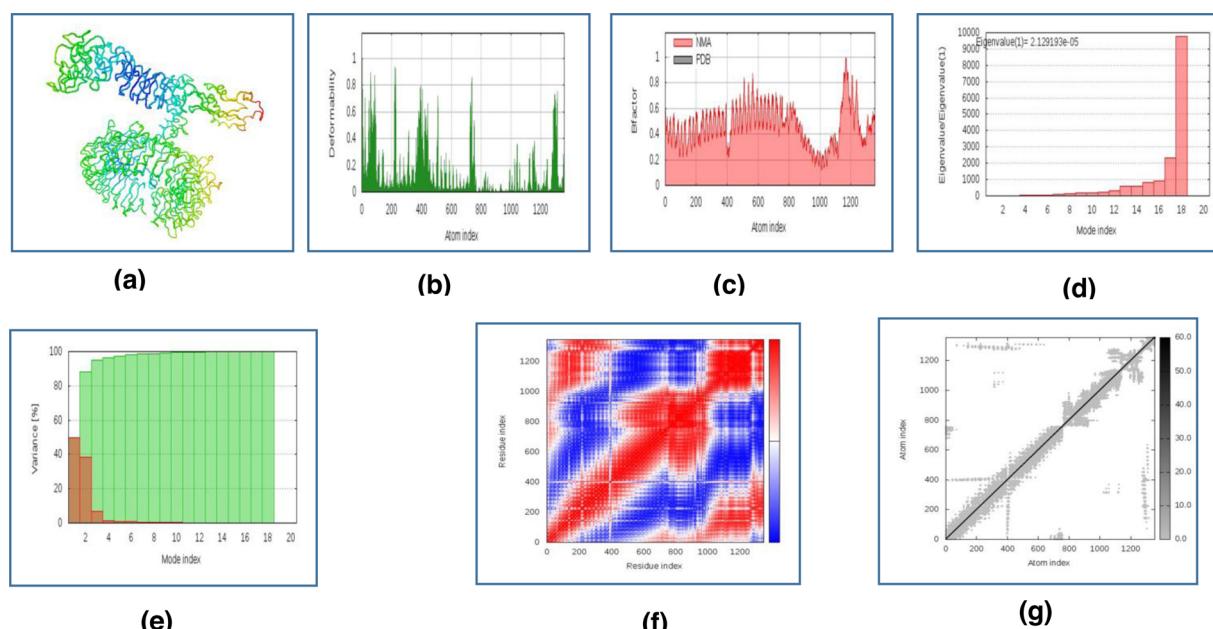


Fig. 4. Figure displaying the results of molecular dynamics simulation study of EV-1 and TLR-8 docked complex. Here, (a) NMA mobility, (b) deformability, (c) B-factor, (d) eigenvalues, (e) variance (red color indicates individual variances and green color indicates cumulative variances), (f) co-variance map (correlated (red), uncorrelated (white) or anti-correlated (blue) motions) and (g) elastic network (darker gray regions indicate more stiffer regions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

antigenic scores in all the three models. Matrix protein VP40 generated antigenicity scores of 0.5692 (bacteria), 0.4598 (tumor), 0.5496 (parasite) in the antigenicity test. Envelope glycoprotein showed antigenicity scores of 0.5852 (bacteria), 0.4933 (tumor), 0.5364 (parasite) and nucleoprotein generated antigenicity scores of 0.5315 (bacteria), 0.5442 (tumor), 0.4692 (parasite) in the antigenicity test. The scores were well over the threshold of 0.4. Especially, all the three proteins had antigenic scores of over 0.5 in the bacterial model. The antigenic scores of the proteins over threshold value pointed to the fact that all the three proteins were predicted to be potential antigens. The three models differ from each other in their prediction accuracy, sensitivity and specificity. The tumor model has better accuracy, sensitivity and specificity than the two other models (Doytchinova and Flower, 2007a,b; Doytchinova and Flower, 2008; Shey et al., 2019).

T lymphocytic cells and B lymphocytic cells are the main immune cells that play significant roles to confer immunity in human body. Firstly, an antigen is recognized by an antigen presenting cell i.e., macrophage, dendritic cell etc. Then the antigen is presented to the helper T cell through the MHC class-II molecule present on the surface of the antigen presenting cell. Since, the helper T cell contains CD4 + molecule on its surface, it is also known as CD4 + T cell. After being activated by the antigen presenting cell, the T-helper cell then activates the B cells and cause the production of memory B cells and antibody producing plasma B cells. The plasma B cell produces a large number of antibodies and the memory B cell functions as the immunological memory. However, T-helper cell also activates macrophage and CD8 + cytotoxic T cells. These activated cells then destroy the target infected cells (Goerdt and Orfanos, 1999; Tanchot and Rocha, 2003; Pavli et al., 1993; Arpin et al., 1995; Cano and Lopera, 2013).

The possible T cell and B cell epitopes of the selected viral proteins were predicted by the IEDB server and these epitopes are expected to induce strong immune response against the target virus. The transmembrane topology of the epitopes were determined to identify whether the epitopes would be present inside or outside of the cell membrane. Antigenicity is defined as the ability of a foreign substance to act as antigen and activate and stimulate the T cell and B cell responses through their antigenic determinant portion or epitope (Fishman et al., 2015). The allergenicity of a substance refers to the ability of that

substance to act as allergen and induce potential allergic reactions within the body (Andreae and Nowak-Wegrzyn, 2017). Epitopes of a protein that remain conserved across various strains are given much priority during vaccine designing than genomic regions that are highly variable because the conserved epitopes of protein(s) provide broader protection across various strains and species (Bui et al., 2007). The epitopes were selected based on few criteria i.e., the epitopes should be highly antigenic so that, they could induce high antigenic response, the epitopes must be non-allergenic in nature so as to, they would not be able to induce any allergenic reaction in an individual and the epitopes should be non-toxic as well to confer safety. The epitopes with 100 % conservancy and over 50 % minimum identity were selected for vaccine construction so that, the conserved epitopes would be able to provide protection against multiple strains. The conservancy analysis was carried out using Zaire Ebola virus strains i.e., Gabon 94 and Kikwit 95, for comparison. Since the selected epitopes were 100 % conserved across these strains, the newly constructed vaccine for Ebola virus Mayinga 76 strain, should also provide considerable immunity against the Gabon 94 and Kikwit 95 strains too. Finally, the cluster analysis of the MHC class-I alleles and MHC class-II alleles was carried out to determine their relationship with each other based on their predicted binding specificity (Thomsen et al., 2013).

In the next step, the protein-peptide docking was carried out between the epitopes and the MHC alleles in order to determine the ability of the epitopes to bind with their respective MHC allele. The MHC class-I epitopes were docked against the MHC class-I allele (PDB ID: 5WJL) and the MHC class-II epitopes were docked against the MHC class-II allele (PDB ID: 5JLZ). At first, the 3D structures of the selected epitopes were generated and after that, the generated 3D structures were docked with HLA-A*11 – 01 allele and HLA DRB1*04 – 01. GISFHPKLR, SGY-YSTTIR, QTNAMVTLR, KTYSF DSTTAII, TPQFLQLNETI, VVFSTSDGKEYT epitopes generated the best results in the protein-protein docking study among the selected epitopes. However, successful interaction of all the epitopes with the MHC alleles was observed, since all of them generated comparable results in the docking study.

Then the complete vaccines were constructed from the selected epitopes. Three different vaccines were constructed and the three vaccines differed from each other only in their adjuvant sequence and the

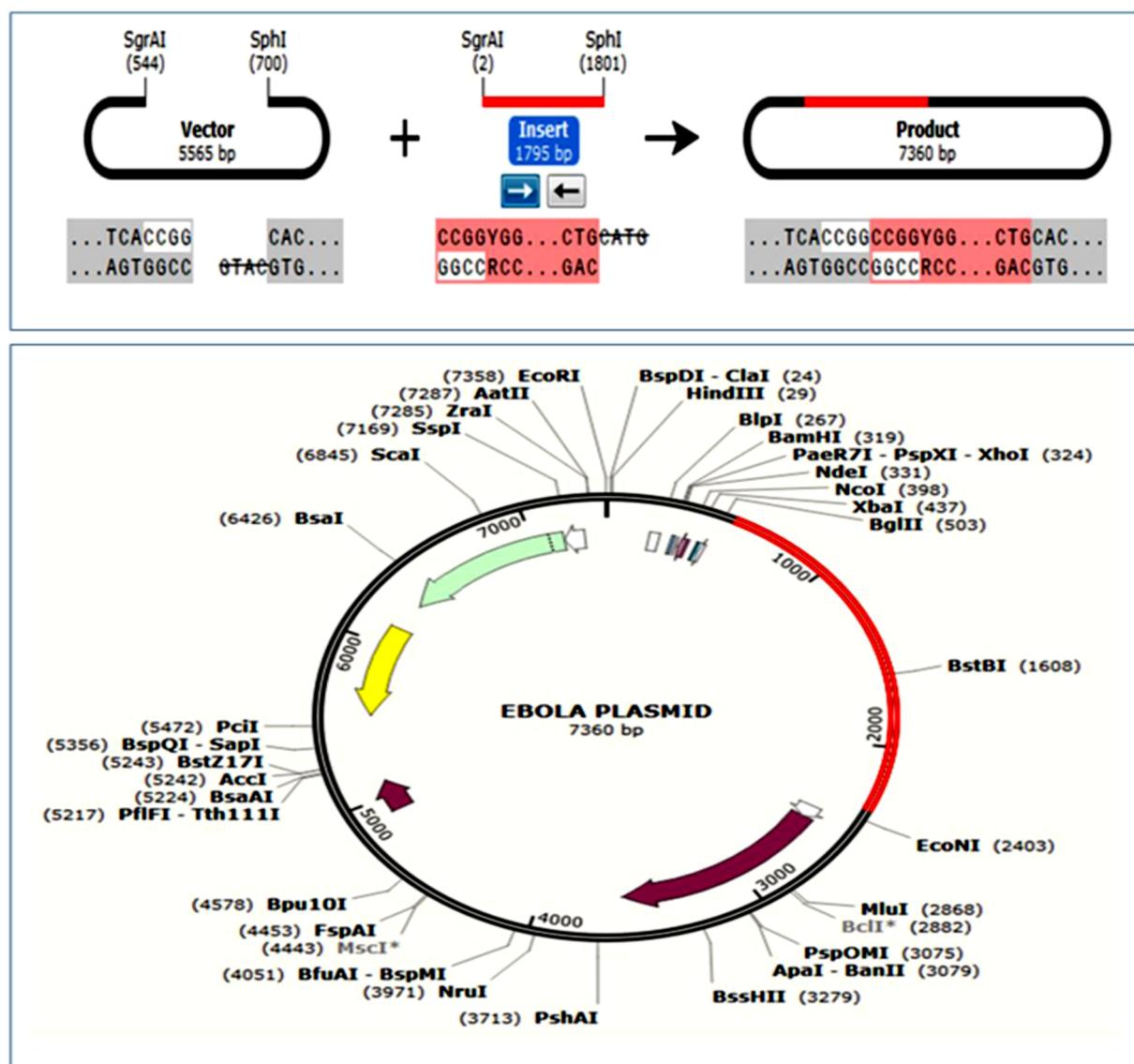


Fig. 5. *In silico* restriction cloning of the EV-1 vaccine sequence in the pET-19b plasmid between the SgrAI and SphI restriction enzyme sites. The red colored marked sites contain the DNA inserts of the vaccine. The cloning was carried out using the SnapGene tool. The newly constructed plasmid can be inserted into *E. coli* strain K12 for efficient vaccine production. In the plasmid, the larger purple colored arrow represents the lacI gene (from 2518 bp to 3600 bp), the smaller purple colored arrow represents the rop gene (from 4912 bp to 5103 bp), yellow colored arrow represents the origin of replication (from 5533 bp to 6121 bp), the light green colored arrow represents the AmpR (ampicillin resistance) gene (from 6292 bp to 7152 bp), the white rectangle represents the T7 terminator (from 213 bp to 260 bp), the light blue colored arrow represents the multiple cloning site (from 319 bp to 335 bp) and the desired gene has been inserted (marked by red color) between the 503 bp and 2146 bp nucleotide. Various restriction enzyme sites are mentioned in the plasmid structure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

vaccines were designated as EV-1, EV-2 and EV-3. EV-1 had beta defensing adjuvant, EV-2 had L7/L12 ribosomal protein adjuvant and EV-3 had HABA protein adjuvant respectively. EAAAK, GGGS, GPGPG and KK linkers were used to conjugate the selected epitopes.

The antigenicity, allergenicity and other physicochemical properties were determined for three different vaccine constructs using different online tools after the successful vaccine construction. All the vaccine constructs were found to be potential antigen. Moreover, they were also predicted to be non-allergenic and expected to cause no allergenic reaction inside the body. EV-3 had the highest extinction co-efficient of $59,945 \text{ M}^{-1} \text{ cm}^{-1}$, which reflected that EV-3 should absorb the highest amount of light of a particular wavelength. EV-2 had the highest aliphatic index and as a result it had the highest amount of aliphatic amino acids in its side chains among all of the vaccine constructs. All the vaccine constructs had instability index of less than 40 and so, all of them were considered to be stable (Guruprasad et al., 1990). Since all of

the vaccines had predicted half-life of 1 h, so their half degradation is expected to take 1 h inside the human body. Three vaccine constructs were predicted to have quite similar physicochemical properties.

EV-3 had the highest percentage of amino acids in the alpha-helix conformation (35.2 %). EV-1 had the highest percentage of amino acids in the beta-sheet conformation (65.6 %). EV-3 had the lowest percentage of amino acids (11.5 %) in the coil structure of the protein. All the vaccine constructs also showed quite good results in the secondary structure prediction experiment. All the vaccine constructs had 5 domains and EV-2 had the highest p-value of 1.37e-07. Therefore, EV-2 had the predicted tertiary model with the lowest quality.

EV-3 showed the best performances in the protein structure refinement and validation experiment with 72.1 % of the amino acids in the most favored regions and 24.0 % of the amino acids in the additional allowed regions as well as 2.7 % of the amino acids in the generously allowed regions and 1.2 % of the amino acids in the disallowed

regions. After validation of the 3D protein structures, the disulfide engineering of the vaccine constructs was performed using Disulfide by Design 2 v12.2 server and the amino acid pairs with binding energy value less than 2.0 kcal/mol were picked up for disulfide bond formation.

Molecular docking has become an important tool in recent days in the field of structural molecular biology and drug-design which acts on the lock and key mechanism. In molecular docking, interaction or predominant binding mode between a ligand and a receptor is predicted with the aid of different types of computer software. The docking software predicts the binding affinity of a ligand-receptor complex and the greater binding affinity corresponds to better predictions. In molecular docking, the receptor is usually a protein and the ligand could be a small chemical molecule, nucleic acid, peptide or a small protein. Two basic steps are involved in the docking study i.e., first the ligand conformation, position and orientation within the active site of the receptor is assessed and then the binding affinity of the docked complex is analyzed. Thus, molecular docking helps to predict the best ligand-receptor interaction from a group of ligand-receptor complexes. Protein-protein docking is used extensively in reverse vaccinology to find out the best possible vaccine construct (Meng et al., 2011; Morris and Lim-Wilby, 2008; Solanki and Tiwari, 2018; Sarkar et al., 2019).

Protein-protein docking was carried out to find the best vaccine construct against several MHC alleles i.e., DRB1*0101, DRB3*0202, DRB5*0101, DRB3*0101, DRB1*0401 and DRB1*0301. The constructed vaccines were also docked against TLR-8. The aim of the docking experiment is to find out the best vaccine among all of the newly constructed vaccines. For enhancing the accuracy of our prediction, three different online tools, EV-1 showed the best and excellent performances in the overall docking experiments. For this reason, EV-1 was considered the best vaccine construct which was selected for dynamics simulation study and *in silico* codon adaptation.

The molecular dynamics simulation study was carried for the docked TLR-8 and EV-1 complex using the online tool iMODS (<http://imods.chaconlab.org/>). The study showed that the complex had quite lesser chance of deformability with high eigenvalue of 2.129193e-05. The deformability graph (Fig. 4 b) also confirmed that the location of the hinges in the structures were not so significant and the complex is quite stable with lower degree of deforming for each individual amino acid residue. EV-1 had a good number of correlated amino acids and also showed a large number of stiffer regions (Fig. 4 f, g). Nevertheless, the complex showed good results in the molecular dynamics simulation study.

Finally, the designed EV-1 vaccine was adapted for *in silico* cloning in *E. coli* strain K12. The obtained CAI value of 0.943 supported that, the codon adaptation was carried out perfectly and the DNA sequence contained the codons that are most likely to be used by the *E. coli* cells. The vaccine DNA sequence was inserted into the SgrAI and SphI restriction sites of the pET-19b plasmid vector with an overall size of 7360 base pair for efficient expression and production of the desired vaccine, EV-1 in *E. coli* cells.

The genome-based technologies for vaccine development will continue to dominate the field of vaccine development since it provides scientists the opportunity to develop vaccines by optimizing the target antigens. Conventional vaccines i.e., the attenuated vaccines or the inactivated vaccines may fail to provide potential immunity towards a target antigen and the conventional approach of vaccine development have raised many safety concerns in many clinical trials. The subunit vaccines like the vaccines predicted in the study could overcome such difficulties (Tameris et al., 2013; Merten, 2002; Hasson et al., 2015; Kaufmann et al., 2014; Stratton et al., 2002).

Overall, this study recommends EV-1 as the best vaccine based on the strategies employed in the study to be an effective worldwide treatment to be triggered against Ebola virus (strain Mayinga-76) infection. However, further *in vivo* and *in vitro* experiments might be required to strengthen the findings of this study.

5. Conclusion

Ebola virus disease is a type of haemorrhagic fever that is caused by the infection of Ebola virus. Ebola is considered as one of the most fatal viruses in the world with terrifying mortality rate as recorded in the previous years. No epitope-based subunit vaccine has yet been discovered with satisfactory results to trigger immunity against Ebola virus. In this study, a possible subunit vaccine against Ebola virus Mayinga-76 was designed using different tools of bioinformatics, immunoinformatics and vaccinomics. At first, the potential proteins of the viral structure were identified. Then, the potential epitopes were identified through different analytical processes and these epitopes were used for vaccine construction. Upon continual experimentation, three possible vaccines were constructed initially and then one best vaccine construct was determined based on molecular docking. Thereafter, the molecular dynamics simulation study and codon adaptation as well as *in silico* cloning were performed on the best selected vaccine. However, *in vivo* and *in vitro* researches should be carried out on the findings of this experiment to finally confirm the safety, efficacy and potentiality of the vaccine constructs. Hopefully, this study will raise research interest among the scientists of the respective field and uphold their efforts to find an effective vaccine against Ebola virus.

Author statement

Md. Asad Ullah conceptualized the research idea and helped others to conduct the research.

Bishajit Sarkar (corresponding author) designed the vaccine with the help of Md. Asad Ullah and Syed Sajidul Islam.

Syed Sajidul Islam analysed the results with the help of Bishajit Sarkar and Md. Asad Ullah.

Declaration of Competing Interest

Md. Asad Ullah declares that he has no conflict of interest. Bishajit Sarkar declares that he has no conflict of interest. Syed Sajidul Islam declares that he has no conflict of interest

Acknowledgements

Authors are thankful to Swift Integrity Computational Lab, Dhaka, Bangladesh, a virtual platform of young researchers, for providing the tools.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.imbio.2020.151949>.

References

- Andreae, D., Nowak-Wegrzyn, A., 2017. The effect of infant allergen/immunogen exposure on long-term health. Early Nutrition and Long-Term Health. Woodhead Publishing, pp. 131–173. <https://doi.org/10.1016/B978-0-08-100168-4.00006-9>.
- Angov, E., 2011. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol. J. 6, 650–659. <https://doi.org/10.1002/biot.201000332>.
- Arpin, C., Dechancet, J., Van Kooten, C., Merville, P., Grouard, G., Briere, F., Bancereau, J., Liu, Y.J., 1995. Generation of memory B cells and plasma cells in vitro. Science 268, 720–722. <https://doi.org/10.1126/science.7537388>.
- Atapour, A., Mokarram, P., MostafaviPour, Z., Hosseini, S.Y., Ghasemi, Y., Mohammadi, S., Nezafat, N., 2019. Designing a fusion protein vaccine against HCV: an *in silico* approach. Int. J. Pept. Res. Thera. 25, 861–872. <https://doi.org/10.1007/s10989-018-9735-4>.
- 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. <https://doi.org/10.3390/ijms18010139>.
- Bhattacharya, D., Cheng, J., 2013. 3Drefine: consistent protein structure refinement by optimizing hydrogen bonding network and atomic-level energy minimization. Proteins. 81, 119–131. <https://doi.org/10.1002/prot.24167>.

- Bhattacharya, D., Nowotny, J., Cao, R., Cheng, J., 2016. 3Drefine: an interactive web server for efficient protein structure refinement. *Nucleic Acids Res.* 44, W406–W409. <https://doi.org/10.1093/nar/gkw336>.
- Biovia, D.S., Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Richmond, T.J., 2016. Dassault systèmes BIOVIA, discovery studio visualizer, v. 17.2, San Diego: dassault systèmes. *J. Chem. Phys.* 10, 0021–9991.
- Buchan, D.W., Jones, D.T., 2019. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* 47, W402–W407. <https://doi.org/10.1093/nar/gkz297>.
- Bui, H.H., Sidney, J., Li, W., Fusseder, N., Sette, A., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinform.* 8, 361. <https://doi.org/10.1186/1471-2105-8-361>.
- Cano, R.L.E., Lopera, H.D.E., 2013. Introduction to T and B lymphocytes. *Autoimmunity: From Bench to Bedside [Internet]*. El Rosario University Press.
- Carbone, A., Zinov'yev, A., Képes, F., 2003. Codon adaptation index as a measure of dominating codon bias. *Bioinformatics* 19, 2005–2015. <https://doi.org/10.1093/bioinformatics/btg272>.
- Chauhan, V., Rungta, T., Goyal, K., Singh, M.P., 2019. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Sci. Rep.* 9, 1–15. <https://doi.org/10.1038/s41598-019-39299-8>.
- Choi, W.Y., Hong, K.J., Hong, J.E., Lee, W.J., 2015. Progress of vaccine and drug development for Ebola preparedness. *Clin. Exp. Vaccine Res.* 4, 11–16. <https://doi.org/10.7774/cerv.2015.4.1.11>.
- Chong, L.C., Khan, A.M., 2019. Vaccine Target Discovery. <https://doi.org/10.1016/b978-0-12-809633-8.20100-3>.
- Craig, D.B., Dombrowski, A.A., 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinform.* 14, 346. <https://doi.org/10.1186/1471-2105-14-346>.
- DeLano, W.L., 2002. PyMol: an open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography* 40, 82–92.
- Dimitrov, I., Flower, D.R., Doytchinova, I., 2013. AllerTOP-a server for in silico prediction of allergens. *BMC Bioinform.* 14, S4. <https://doi.org/10.1007/s00894-014-2278-5>.
- Dimitrov, I., Naneva, L., Doytchinova, I., Bangov, I., 2014. AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* 30, 846–851. <https://doi.org/10.1093/bioinformatics/btt619>.
- Doytchinova, I.A., Flower, D.R., 2007a. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine* 25, 856–866. <https://doi.org/10.1016/j.vaccine.2006.09.032>.
- Doytchinova, I.A., Flower, D.R., 2007b. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* 8, 4. <https://doi.org/10.1186/1471-2105-8-4>.
- Doytchinova, I.A., Flower, D.R., 2008. Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. *Open Vaccine J.* 1, 4.
- Duhovny, D., Nussinov, R., Wolfson, H.J., 2002. Efficient unbound docking of rigid molecules. International Workshop on Algorithms in Bioinformatics. Springer, Berlin, Heidelberg, pp. 185–200. https://doi.org/10.1007/3-540-45784-4_14.
- Dyer, O., 2019. Two Ebola treatments halve deaths in trial in DRC outbreak. *BMJ* 366. <https://doi.org/10.1136/bmj.l5140>.
- Feldmann, H., Geisbert, T.W., 2011. Ebola haemorrhagic fever. *Lancet* 377, 849–862. [https://doi.org/10.1016/S0140-6736\(10\)60667-8](https://doi.org/10.1016/S0140-6736(10)60667-8).
- Feldmann, H., Klenk, H.D., Sanchez, A., 1993. Molecular biology and evolution of filoviruses. Unconventional Agents and Unclassified Viruses. Springer, Vienna, pp. 81–100. https://doi.org/10.1007/978-3-7091-9300-6_8.
- Feng, T., Chen, F., Kang, Y., Sun, H., Liu, H., Li, D., Zhu, F., Hou, T., 2017. HawkRank: a new scoring function for protein–protein docking based on weighted energy terms. *J. Cheminformatics* 9, 66. <https://doi.org/10.1186/s13321-017-0254-7>.
- Fishman, J.M., Wiles, K., Wood, K.J., 2015. The acquired immune system response to biomaterials, including both naturally occurring and synthetic biomaterials. *Host Response to Biomaterials*. pp. 151–187. <https://doi.org/10.1016/B978-0-12-800196-7.00008-6>.
- Funderburg, N., Lederman, M.M., Feng, Z., Drage, M.G., Jadlowsky, J., Harding, C.V., Weinberg, A., Sieg, S.F., 2007. Human β-defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2. *Proc. Natl. Acad. Sci.* 104, 18631–18635. <https://doi.org/10.1073/pnas.0702130104>.
- Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server. In the Proteomics Protocols Handbook 571–607. Humana Press <https://doi.org/10.1385/1-59259-890-0:571>.
- Gill, S.C., Von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319–326. <https://doi.org/10.1093/oxfordjournals.jbchem.a13168>.
- Goertt, S., Orfanos, C.E., 1999. Other functions, other genes: alternative activation of antigen-presenting cells. *Immunity* 10, 137–142.
- Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, 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. <https://doi.org/10.1093/nar/gki376>.
- Gujula, K.R., 2008. Prediction and Comparison of hiv-1 Protease Inhibitor Binding Energies by Various Molecular Docking Methods (Doctoral dissertation). .
- Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G.P., Open Source Drug Discovery Consortium, 2013. In silico approach for predicting toxicity of peptides and proteins. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0073957>.
- Guruprasad, K., Reddy, B.B., Pandit, M.W., 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng. Des. Sel.* 4, 155–161. <https://doi.org/10.1093/protein/4.2.155>.
- 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. <https://doi.org/10.1016/j.meegid.2016.12.010>.
- Hancock, R.E., Nijnik, A., Philpott, D.J., 2012. Modulating immunity as a therapy for bacterial infections. *Nat. Rev. Microbiol.* 10, 243–254. <https://doi.org/10.1038/nrmicro2745>.
- Hasson, S.S.A.A., Al-Busaidi, J.K.Z., Sallam, T.A., 2015. The past, current and future trends in DNA vaccine immunisations. *Asian Pac. J. Trop. Biomed.* 5, 344–353. [https://doi.org/10.1016/s2221-1691\(15\)30366-x](https://doi.org/10.1016/s2221-1691(15)30366-x).
- Holtsberg, F.W., Shulenin, S., Vu, H., Howell, K.A., Patel, S.J., Gunn, B., Karim, M., Lai, J.R., Frei, J.C., Nyakatura, E.K., Zeitlin, L., 2016. Pan-ebolavirus and pan-filovirus mouse monoclonal antibodies: protection against Ebola and Sudan viruses. *J. Virol.* 90, 266–278. <https://doi.org/10.1128/jvi.02171-15>.
- Hou, T., Wang, J., Li, Y., Wang, W., 2010. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* 51, 69–82.
- Ikai, A., 1980. Thermostability and aliphatic index of globular proteins. *J. Biochem.* 88, 1895–1898.
- Jones, D.T., 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195–202. <https://doi.org/10.1006/jmbi.1999.3091>.
- Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., Xu, J., 2012. Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7, 1511. <https://doi.org/10.1038/nprot.2012.085>.
- Kanapathipillai, R., Henao Restrepo, A.M., Fast, P., Wood, D., Dye, C., Kieny, M.P., Moorthy, V., 2014. Ebola vaccine—an urgent international priority. *N. Engl. J. Med.* 371, 2249–2251. <https://doi.org/10.1056/NEJMp1412166>.
- Kaufmann, S.H., McElrath, M.J., Lewis, D.J., Del Giudice, G., 2014. Challenges and responses in human vaccine development. *Curr. Opin. Immunol.* 28, 18–26. <https://doi.org/10.1016/j.co.2014.01.009>.
- Khatoon, N., Pandey, R.K., Prajapati, V.K., 2017. Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci. Rep.* 7, 1–12. <https://doi.org/10.1038/s41598-017-08842-w>.
- Kiley, M.P., Bowen, E.T.W., Eddy, G.A., Isaacsom, M., Johnson, K.M., McCormick, J.B., Murphy, F.A., Pattyn, S.R., Peters, D., Prozesky, O.W., Regnery, R.L., 1982. Filoviridae: a taxonomic home for Marburg and Ebola viruses? *Intervirology* 18, 24–32. <https://doi.org/10.1159/000149300>.
- Klenk, H.D., Feldmann, H., eds. 2004. Ebola and Marburg viruses: molecular and cellular biology. Garland Science.
- Kovacs, J.A., Chacón, P., Abagyan, R., 2004. Predictions of protein flexibility: first-order measures. *Proteins: Struct., Funct., Bioinf.* 56, 661–668. <https://doi.org/10.1002/prot.20151>.
- Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., Vajda, S., 2013. How good is automated protein docking? *Proteins: Struct., Funct., Bioinf.* 81, 2159–2166. <https://doi.org/10.1002/prot.24403>.
- Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorney, D., Yueh, C., Beglov, D., Vajda, S., 2017. The ClusPro web server for protein–protein docking. *Int. J. Internet Protoc. Technol.* 12, 255. <https://doi.org/10.1038/nprot.2016.169>.
- Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157, 105–132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0).
- Lamiable, A., Thévenet, P., Rey, J., Vavrusa, M., Derreumaux, P., Tufféry, P., 2016. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res.* 44, W449–W454. <https://doi.org/10.1093/nar/gkw329>.
- Larsen, J.E.P., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. *Immunoeng. Res.* 2, 2. <https://doi.org/10.1186/1745-7580-2-2>.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. <https://doi.org/10.1107/s002188992009944>.
- Laskowski, R.A., MacArthur, M.W., Thornton, J.M., 2006. PROCHECK: validation of protein-structure coordinates. *Int. Tables Crystallogr.* 25 (2), 722–725. <https://doi.org/10.1107/97809553602060000882>.
- Lee, S.J., Shin, S.J., Lee, M.H., Lee, M.G., Kang, T.H., Park, W.S., Soh, B.Y., Park, J.H., Shin, Y.K., Kim, H.W., Yun, C.H., 2014. A potential protein adjuvant derived from *Mycobacterium tuberculosis* Rv0652 enhances dendritic cells-based tumor immunotherapy. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0104351>.
- Leroy, E.M., Baize, S., Volchkov, V.E., Fisher-Hoch, S.P., Georges-Courbot, M.C., Lansoud-Soukate, J., Capron, M., Debre, P., Georges, A.J., McCormick, J.B., 2000. Human asymptomatic Ebola infection and strong inflammatory response. *Lancet* 355, 2210–2215. [https://doi.org/10.1016/S0140-6736\(00\)2405-3](https://doi.org/10.1016/S0140-6736(00)2405-3).
- Leroy, E.M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassouni, A., Yaba, P., Délicat, A., Paweska, J.T., Gonzalez, J.P., Swanepoel, R., 2005. Fruit bats as reservoirs of Ebola virus. *Nature* 438, 575–576. <https://doi.org/10.1038/438575a>.
- Lester, S.N., Li, K., 2014. Toll-like receptors in antiviral innate immunity. *J. Mol. Biol.* 426, 1246–1264. <https://doi.org/10.1016/j.jmb.2013.11.024>.
- López-Blanco, J.R., Garzón, J.I., Quintana-Ortí, E.S., Chacón, P., 2011. iMod: multipurpose normal mode analysis in internal coordinates. *Bioinformatics* 27, 2843–2850. <https://doi.org/10.1093/bioinformatics/btr497>.
- López-Blanco, J.R., Aliaga, J.I., Quintana-Ortí, E.S., Chacón, P., 2014. iMODS: internal coordinates normal mode analysis server. *Nucleic Acids Res.* 42, W271–W276. <https://doi.org/10.1093/nar/gku339>.
- Ma, J., Wang, S., Zhao, F., Xu, J., 2013. Protein threading using context-specific alignment potential. *Bioinformatics* 29, i257–i265. <https://doi.org/10.1093/bioinformatics/btt210>.
- Madara, J.J., Han, Z., Ruthel, G., Freedman, B.D., Harty, R.N., 2015. The multifunctional Ebola virus VP40 matrix protein is a promising therapeutic target. *Future Virol.* 10, 537–546. <https://doi.org/10.2217/fvl.15>.
- María, R.R., Arturo, C.J., Alicia, J.A., Paulina, M.G., Gerardo, A.O., 2017. The Impact of Bioinformatics on Vaccine Design and Development. InTech, Rijeka, Croatia. <https://doi.org/10.5772/intechopen.69273>.
- Martinez, O., Tantral, L., Mulherkar, N., Chandran, K., Basler, C.F., 2011. Impact of Ebola mucin-like domain on antiglycoprotein antibody responses induced by Ebola virus-like particles. *J. Infect. Dis.* 204, S825–S832. <https://doi.org/10.1093/infdis/jir295>.
- Martinez, O., Ndungo, E., Tantral, L., Miller, E.H., Leung, L.W., Chandran, K., Basler, C.F., 2013. A mutation in the Ebola virus envelope glycoprotein restricts viral entry in a

- host species-and cell-type-specific manner. *J. Virol.* 87, 3324–3334. <https://doi.org/10.1128/jvi.01598-12>.
- Martínez-Romero, C., García-Sastre, A., 2015. Against the clock towards new Ebola virus therapies. *Virus Res.* 209, 4–10. <https://doi.org/10.1016/j.virusres.2015.05.025>.
- Meng, X.Y., Zhang, H.X., Mezei, M., Cui, M., 2011. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comp. Aided Drug. Des.* 7, 146–157. <https://doi.org/10.2174/157340911795677602>.
- Merten, O.W., 2002. Virus contaminations of cell cultures—a biotechnological view. *Cytotechnology* 39, 91–116. <https://doi.org/10.1023/a:1022969101804>.
- Möller, S., Croning, M.D., Apweiler, R., 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646–653. <https://doi.org/10.1093/bioinformatics/17.7.646>.
- Morris, G.M., Lim-Wilby, M., 2008. Molecular docking. *Molecular Modeling of Proteins*. Humana Press, pp. 365–382. https://doi.org/10.1007/978-1-59745-177-2_19.
- Morris, A.L., MacArthur, M.W., Hutchinson, E.G., Thornton, J.M., 1992. Stereochemical quality of protein structure coordinates. *Proteins: Struct., Funct., Bioinf.* 12, 345–364. <https://doi.org/10.1002/prot.340120407>.
- Nabel, G.J., Yang, Z.Y., Sullivan, N., Sanchez, A., US Department of Health and Human Services (HHS), 2012. Development of a preventive vaccine for filovirus infection in primates. U.S. Patent 8,124,592.
- Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., Gray, T., 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 4, 2411–2423. <https://doi.org/10.1002/pro.5560041120>.
- Pandey, R.K., Sundar, S., Prajapati, V.K., 2016. Differential expression of miRNA regulates T cell differentiation and plasticity during visceral leishmaniasis infection. *Front. Microbiol.* 7, 206. <https://doi.org/10.3389/fmicb.2016.00206>.
- Patel, J.S., Quates, C.J., Johnson, E.L., Ytreberg, F.M., 2019. Expanding the watch list for potential Ebola virus antibody escape mutations. *PLoS One* 14. <https://doi.org/10.1371/journal.pone.0211093>.
- Pavli, P., Hume, D.A., Van De Pol, E., Doe, W.F., 1993. Dendritic cells, the major antigen-presenting cells of the human colonic lamina propria. *Immunology* 78, 132.
- Peng, J., Xu, J., 2011. RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins: Struct., Funct., Bioinf.* 79, 161–171. <https://doi.org/10.1002/prot.23175>.
- Petersen, M.T.N., Jonson, P.H., Petersen, S.B., 1999. Amino acid neighbours and detailed conformational analysis of cysteines in proteins. *Protein Eng.* 12, 535–548. <https://doi.org/10.1093/protein/12.7.535>.
- Petersen, B., Lundegaard, 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. <https://doi.org/10.1371/journal.pone.0015079>.
- Poland, G.A., Ovsyannikova, I.G., Jacobson, R.M., 2009. Application of pharmacogenomics to vaccines. *Pharmacogenomics*, 10, 837–852. <https://doi.org/10.2217/pgs.09.25>.
- Pourrut, X., Kumulungui, B., Wittmann, T., Moussavou, G., Délicat, A., Yaba, P., Nkoghe, D., Gonzalez, J.P., Leroy, E.M., 2005. The natural history of Ebola virus in Africa. *Microbes Infect.* 7, 1005–1014. <https://doi.org/10.1016/j.micinf.2005.04.006>.
- 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, 778–791. <https://doi.org/10.1080/07391102.2015.1054430>.
- 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. <https://doi.org/10.1038/nrd2224>.
- 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. <https://doi.org/10.1016/j.imbio.2015.12.004>.
- Rappuoli, R., 2000. Reverse vaccinology. *Curr. Opin. Microbiol.* 3, 445–450. [https://doi.org/10.1016/s1369-5274\(00\)00119-3](https://doi.org/10.1016/s1369-5274(00)00119-3).
- Rappuoli, R., Bottomley, M.J., D'Orso, U., Finco, O., De Gregorio, E., 2016. Reverse vaccinology 2.0: human immunology instructs vaccine antigen design. *J. Exp. Med.* 213, 469–481. <https://doi.org/10.1084/jem.20151960>.
- Regnery, R.L., Johnson, K.M., Kiley, M.P., 1980. Virion nucleic acid of Ebola virus. *J. Virol.* 36, 465–469.
- Regules, J.A., Beigel, J.H., Paolino, K.M., Voell, J., Castellano, A.R., Hu, Z., Muñoz, P., Moon, J.E., Ruck, R.C., Bennett, J.W., Twomey, P.S., 2017. A recombinant vesicular stomatitis virus Ebola vaccine. *N. Engl. J. Med.* 376, 330–341. <https://doi.org/10.1056/NEJMoa1414216>.
- Saha, S., Raghava, G.P.S., 2006. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res.* 34, W202–W209. <https://doi.org/10.1093/nar/gkl343>.
- Sanchez, A., Trappier, S.G., Mahy, B.W., Peters, C.J., Nichol, S.T., 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci.* 93, 3602–3607. <https://doi.org/10.1073/pnas.93.8.3602>.
- Saphire, E.O., Schendel, S.L., Gunn, B.M., Milligan, J.C., Alter, G., 2018. Antibody-mediated protection against Ebola virus. *Nat. Immunol.* 19, 1169–1178. <https://doi.org/10.1038/s41590-018-0233-9>.
- Sarkar, B., Islam, S.S., Zohora, U.S., Ullah, M.A., 2019. Virus like particles-A recent advancement in vaccine development. *Korean J. Microbiol.* 55, 327–343. <https://doi.org/10.7845/kjm.2019.9089>.
- Sateesh, P., Rao, A., Sangeeta, S.K., Babu, M.N., Grandhi, R.S., 2010. Homology modeling and sequence analysis of anxC3. *1. Int. J. Eng. Sci. Technol.* 2, 1125–1130.
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* 33, W363–W367. <https://doi.org/10.1093/nar/gki481>.
- Sharp, P.M., Li, W.H., 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15, 1281–1295. <https://doi.org/10.1093/nar/15.3.1281>.
- Shen, Y., Maupetit, J., Derreumaux, P., Tuffery, P., 2014. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. *J. Chem. Theory Comput.* 10, 4745–4758. <https://doi.org/10.1021/ct500592m>.
- Shey, R.A., Ghogomu, S.M., Esho, K.K., Nebangwa, N.D., Shintouo, C.M., Nongley, N.F., Asa, B.F., Ngale, M., Vanhamme, L., Soupgui, J., 2019. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci. Rep.* 9, 1–18. <https://doi.org/10.1038/s41598-019-40833-x>.
- Shi, Q., Zhou, Y., Sun, Y., 2005. Influence of pH and ionic strength on the steric mass-action model parameters around the isoelectric point of protein. *Biotechnol. Prog.* 21, 516–523. <https://doi.org/10.1021/bp0497350>.
- 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, 1–19. <https://doi.org/10.1038/s41598-018-26689-7>.
- Stern, L.J., Calvo-Calle, J.M., 2009. HLA-DR: molecular insights and vaccine design. *Curr. Pharm. Des.* 15, 3249–3261.
- Stratton, K., Almario, D.A., McCormick, M.C., Institute of Medicine (US) Immunization Safety Review Committee, 2002. *Immunization Safety Review: SV40 Contamination of Polio Vaccine and Cancer. Inimmunization Safety Review: SV40 Contamination of Polio Vaccine and Cancer. National Academies Press (US)*.
- Sturniolo, T., Bono, E., Ding, J., Radziszani, L., Tuereci, O., Sahin, U., Braxenthaler, M., Gallazzi, F., Prott, M.P., Sinigaglia, F., Hammer, J., 1999. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat. Biotechnol.* 17, 555–561. <https://doi.org/10.1038/9858>.
- Sun, H., Li, Y., Tian, S., Xu, L., Hou, T., 2014. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys. Chem. Chem. Phys.* 16, 16719–16729. <https://doi.org/10.1039/C4CP01388C>.
- Tameris, M.D., Hatherill, M., Landry, B.S., Scriba, T.J., Snowden, M.A., Lockhart, S., Shea, J.E., McClain, J.B., Hussey, G.D., Hanekom, W.A., Mahomed, H., 2013. Safety and efficacy of MVA85A, a new tuberculous vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *Lancet* 381, 1021–1028. [https://doi.org/10.1016/s0140-6736\(13\)60177-4](https://doi.org/10.1016/s0140-6736(13)60177-4).
- Tanchot, C., Rocha, B., 2003. CD8 and B cell memory: same strategy, same signals. *Nat. Immunol.* 4, 431–432. <https://doi.org/10.1038/ni0503-431>.
- Thévenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P., Tuffery, P., 2012. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res.* 40, W288–W293. <https://doi.org/10.1093/nar/gks419>.
- Thompson, J.M., Iwasaki, A., 2008. Toll-like receptors regulation of viral infection and disease. *Adv. Drug Deliv. Rev.* 60, 786–794. <https://doi.org/10.1016/j.addr.2007.11.003>.
- Thomsen, M., Lundsgaard, C., Buus, S., Lund, O., Nielsen, M., 2013. MHCcluster, a method for functional clustering of MHC molecules. *Immunogenetics* 65, 655–665. <https://doi.org/10.1007/s00251-013-0714-9>.
- Toussi, D.N., Massari, P., 2014. Immune adjuvant effect of molecularly-defined toll-like receptor ligands. *Vaccines* 2, 323–335. <https://doi.org/10.3390/vaccines2020323>.
- Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., Kozaev, D., 2017. New additions to the C lus P ro server motivated by CAPRI. *Proteins: Struct., Funct., Bioinf.* 85, 435–444. <https://doi.org/10.1002/prot.25219>.
- Vangone, A., Bonvin, A.M., 2015. Contacts-based prediction of binding affinity in protein–protein complexes. *elife* 4, e07454.
- Vita, R., Mahajan, S., Overton, J.A., Dhandha, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., Peters, B., 2019. The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res.* 47, D339–D343. <https://doi.org/10.1093/nar/gky1006>.
- Volchkov, V., Volchkova, V., Eckel, C., Klenk, H.D., Bouloy, M., LeGuenno, B., Feldmann, H., 1997. Emergence of subtype Zaire Ebola virus in Gabon. *Virology* 232, 139–144. <https://doi.org/10.1006/viro.1997.8529>.
- Wamala, J.F., Lukwago, L., Malimbo, M., Nguku, P., Yoti, Z., Musenero, M., Amone, J., Mbabazi, W., Nanyunja, M., Zaramba, S., Opio, A., 2010. Ebola hemorrhagic fever associated with novel virus strain, Uganda, 2007–2008. *Emerg. Infect. Dis.* 16, 1087. <https://doi.org/10.3201/eid1607.091525>.
- Watanabe, S., Noda, T., Kawaoka, Y., 2006. Functional mapping of the nucleoprotein of Ebola virus. *J. Virol.* 80, 3743–3751. <https://doi.org/10.1128/jvi.80.8.3743-3751.2006>.
- Weng, G., Wang, E., Wang, Z., Liu, H., Zhu, F., Li, D., Hou, T., 2019. HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA. *Nucleic Acids Res.* 47, W322–W330. <https://doi.org/10.1093/nar/gkz397>.
- World Health Organization, 2014. *Experimental Ebola Vaccines: WHO Consultation on Ebola Vaccines*.
- 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. <https://doi.org/10.1186/1423-0127-17-88>.
- Xue, L.C., Rodrigues, J.P., Kastritis, P.L., Bonvin, A.M., Vangone, A., 2016. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics* 32, 3676–3678. <https://doi.org/10.1093/bioinformatics/btw514>.
- Yuriev, E., Ramsland, P.A., 2013. Latest developments in molecular docking: 2010–2011 in review. *J. Mol. Recognit.* 26, 215–239. <https://doi.org/10.1002/jmr.2266>.
- Zobayer, N., Hossain, A.A., Canonica, G.W., Blaiss, M., Shimamura, T., Shiroishi, M., Weyand, S., Tsujimoto, H., Winter, G., Hill, S.J., Ganellin, C.R., 2011. The importance of short structural motifs in protein structure analysis. *J. Biol. Sci.* 18, 47–53.