

Blueprint of epitope-based multivalent and multipathogenic vaccines: targeted against the dengue and zika viruses

Bishajit Sarkar, Md. Asad Ullah, Yusha Araf, Sowmen Das & Mohammad Jakir Hosen

To cite this article: Bishajit Sarkar, Md. Asad Ullah, Yusha Araf, Sowmen Das & Mohammad Jakir Hosen (2021) Blueprint of epitope-based multivalent and multipathogenic vaccines: targeted against the dengue and zika viruses, *Journal of Biomolecular Structure and Dynamics*, 39:18, 6882-6902, DOI: [10.1080/07391102.2020.1804456](https://doi.org/10.1080/07391102.2020.1804456)

To link to this article: <https://doi.org/10.1080/07391102.2020.1804456>



[View supplementary material](#)



Published online: 08 Aug 2020.



[Submit your article to this journal](#)



Article views: 193



[View related articles](#)



[View Crossmark data](#)



[Citing articles: 1](#) [View citing articles](#)

Blueprint of epitope-based multivalent and multipathogenic vaccines: targeted against the dengue and zika viruses

Bishajit Sarkar^a , Md. Asad Ullah^a , Yusha Araf^b , Sowmen Das^c and Mohammad Jakir Hosen^b 

^aDepartment of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Jahangirnagar University, Dhaka, Bangladesh;

^bDepartment of Genetic Engineering and Biotechnology, School of Life Sciences, Shahjalal University of Science and Technology, Sylhet, Bangladesh; ^cDepartment of Computer Science and Engineering, School of Physical Sciences, Shahjalal University of Science and Technology, Sylhet, Bangladesh

Communicated by Ramaswamy H. Sarma

ABSTRACT

Both dengue virus (DENV) and zika virus (ZIKV) belong to the highly infectious *Flaviviridae* family that has already caused several outbreaks and epidemics in many countries. DENV and ZIKV cause two of the most wide spread mosquito-borne viral diseases in the world, dengue fever (DENF) and zika fever (ZIKF), respectively. In many regions around the world, both of these diseases can outbreak together and can be lethal as well as life-threatening. Unfortunately, there is no functional and satisfactory vaccine available to combat these viruses. Therefore, in this study, we have attempted to design a blueprint of potential multivalent and multipathogenic vaccines using immunoinformatics approach, which can combat both the DENV and ZIKV infections, simultaneously. Initially, three vaccines were designed; containing highly antigenic, non-allergenic, and non-toxic epitopes of T-cell (100% conserved) and B-cell from all the four DENV serotypes and ZIKV. In total, nine cytotoxic T-lymphocytic (CTL), nine helper T-lymphocytic (HTL), and seven B-cell lymphocytic (BCL) epitopes were used to construct three vaccines using three different adjuvants, designated as 'V1', 'V2', and 'V3'. Later, V3 was found to be the best vaccine construct, determined by molecular docking analysis. Thereafter, several *in silico* validation studies including molecular dynamics simulation and immune simulation were performed which indicated that V3 might be quite stable and should generate substantial immune response in the biological environment. However, further *in vivo* and *in vitro* validation might be required to finally confirm the safety and efficacy of our suggested vaccine constructs.

ARTICLE HISTORY

Received 20 May 2020

Accepted 24 July 2020

KEYWORDS

Dengue virus; zika virus; vaccine; immunoinformatics; multivalent vaccines

1. Introduction

Dengue and zika fevers are two of the most prevalent mosquito-borne diseases around the world, caused by the closely related dengue virus (DENV) and zika virus (ZIKV), respectively. Both these viruses belong to the *Flaviviridae* family and have almost similar genetic makeup with a single-stranded positive-sense RNA (Agumadu & Ramphul, 2018; Simmons et al., 2012). DENV has four main serotypes (namely DENV-1, DENV-2, DENV-3, and DENV-4) that share ~65% similarity among their genomes and these serotypes have made the diagnostic as well as treatments challenging (Mustafa et al., 2015). However, ZIKV does not have any serotype but it has two lineages (i.e. the Asian lineage and the African lineage) and these two lineages share more than 95% amino acid sequence identity (Dowd et al., 2016).

At present, dengue has become one of the major global burdens since the incidence of dengue cases has increased dramatically in recent decades. About 96 million people per year get infected with DENV (Bhatt et al., 2013). And about 3.9 billion people from 129 countries are at risk of dengue

infections (Brady et al., 2012; Guzman et al., 2010). On the other hand, the ZIKV infection was regarded as a self-limited, mild illness in the past, however, two back to back outbreaks in 2007 and 2013 has led the scientific community to mark it as a global threat. Unlike DENV, ZIKV can also be sexually transmitted. Furthermore, the suspected link of ZIKV with congenital abnormalities like Guillain-Barré syndrome had led it to be public health emergency of international concern (Abbink et al., 2018; Kindhauser et al., 2016). Every year, about 12,500 deaths occur worldwide due to the severe DENV infection including 500,000 hospitalizations (Dupont-Rouzeyrol et al., 2015). Although the death rate due to ZIKV infection is not so high like the other mosquito-borne diseases, however, still the infection rate of ZIKV is alarming in 87 countries and territories identified by the WHO, as of July, 2019. Along with the documented infections, suspected birth defects due to the ZIKV infections are also being reported in these regions (WHO, 2019). The co-existence and co-circulation of both DENV and ZIKV are quite common in all the tropical and sub-tropical regions around the world (Carrillo-Hernández et al., 2018). Since these two viruses are



circulated by the same vector, therefore, the ZIKV outbreak has been reported in those regions where DENV outbreak is prominent (Tang et al., 2018). Moreover, co-infection of both DENV and ZIKV has also been reported in several studies, particularly predominant in the endemic regions of these two viruses (Chaves et al., 2018).

Scientists are working hard to develop potential vaccines against DENV and ZIKV. Several vaccine candidates for both DENV and ZIKV are in various stages of clinical and pre-clinical trials around the world but none of them looks promising due to low efficacy and safety (Abbink et al., 2018; da Silveira et al., 2019; Guzman et al., 2010). Although, 'Dengvaxia' of the Sanofi Pasteur has been marketed as the first DENV vaccine and it can confer immunity towards the all four serotypes of DENV, but its live attenuated form has made it controversial to use. Moreover, the use of 'Dengvaxia' has become limited as it can only work in the individuals who have been infected by at least one serotype of DENV in the past (Abbink et al., 2018; Cohen, 2019; Fatima & Syed, 2018). Unfortunately, no ZIKV vaccine is also currently available in the market. Therefore, the development of potential vaccines against the DENV and ZIKV has become a burning issue among the scientific community around the world (Fernandez & Diamond, 2017).

In this study, multivalent and multipathogenic epitope-based subunit vaccines are designed using several *in silico* tools including reverse vaccinology and immunoinformatics to fight against both the DENV and ZIKV, targeting the envelope protein E of both DENV and ZIKV. These *in silico* methods are quick, easy, and cost-effective approaches of vaccine designing where the novel antigens of a pathogen are identified by analyzing its genetic makeup and the vaccines are designed targeting the novel antigens (Chong & Khan, 2019; María et al., 2017). Our results may open a new avenue to good management of patients where both the DENV and ZIKV co-exist.

2. Materials and methods

2.1. Identification and collection of viral protein sequences

The viral protein sequences, envelope protein E of DENV serotypes-1, 2, 3, and 4 as well as ZIKV, were identified and downloaded in FASTA format from the National Center for Biotechnology Information (NCBI) database (<https://www.ncbi.nlm.nih.gov/>). **Figure 1** depicts a flowchart that describes the step-by-step methods used to conduct the whole study.

2.2. Physicochemical property analysis of the proteins

ExPASy's online tool ProtParam (<https://web.expasy.org/protparam/>) was used for the identification and prediction of different physicochemical properties of the selected proteins including number of amino acids, number of positively charged amino acids, number of negatively charged amino

acids, theoretical pl, extinction co-efficient, and estimated half-life (Gasteiger et al., 2005).

2.3. T-cell and B-cell epitope determination

The T-cell and B-cell epitopes of the viral proteins were predicted for vaccine construction. Both of these cells are important for proper immune response. The cytotoxic T-cells specifically recognize antigens and aid to eliminate them from the body. The helper T-cells function in the activation of the B-cells, macrophages and sometimes even the cytotoxic T-cells. Again, the B-cells produce antibodies which are also essential for immunity against antigens or pathogens (Chaudhri et al., 2009; Cooper, 2015; Zhu & Paul, 2008). The epitopes that were found to be highly antigenic, non-allergenic, 100% conserved across other strains and species, and non-toxic, were finally selected for vaccine construction. Since multivalent and multipathogenic vaccines were designed in this study, the envelope protein E from DENV-1 was used as the model for determining the T-cell epitopes. As the epitopes from DENV-1, that were found to be 100% conserved among DENV serotypes-2, 3, and 4 and the epitopes from ZIKV that were predicted to be 100% conserved among DENV serotypes-1 and 2, were considered for designing the multivalent and multipathogenic vaccines, these vaccines are expected to be effective against all the four serotypes of DENV and also the ZIKV. Moreover, the B-cell epitopes were determined separately from each of the DENV serotypes and ZIKV envelope protein E and as before, the highly antigenic, non-allergenic, 100% conserved, and non-toxic epitopes were taken for vaccine construction. The prediction of the epitopes was conducted using Immune Epitope Database or IEDB (<https://www.iedb.org/>) server (Vita et al., 2019). The MHC class-I restricted CD8+ cytotoxic T-lymphocytic (CTL) epitopes were predicted using the stabilized matrix method (SMM) method for HLA-A*11-01 allele (Peters & Sette, 2005). Top 20 MHC class-I epitopes were selected for further analysis. Moreover, MHC class-II restricted CD4+ helper T-lymphocytic (HTL) epitopes were predicted using the Sturniolo method for HLA DRB1*04-01 allele which predicts the results in percentile scores. The lower the percentile score, the better the binding affinity of the epitopes with their targets and vice versa (Sturniolo et al., 1999). Top 20 best predicted MHC class-II epitopes were taken for further analysis. The B-cell lymphocytic (BCL) epitopes of the proteins were predicted by BepiPred linear epitope prediction method and the BCL epitopes with more than 10 amino acids were selected. The BepiPred method uses the combination of hidden Markov model and propensity scale method to predict the locations of linear BCL epitopes (Larsen et al., 2006).

2.4. Antigenicity, allergenicity, conservancy, and toxicity prediction

After determining the epitopes in Sub-section 2.3, they were analyzed through several online tools for predicting their antigenicity, allergenicity, conservancy, and toxicity. VaxiJen

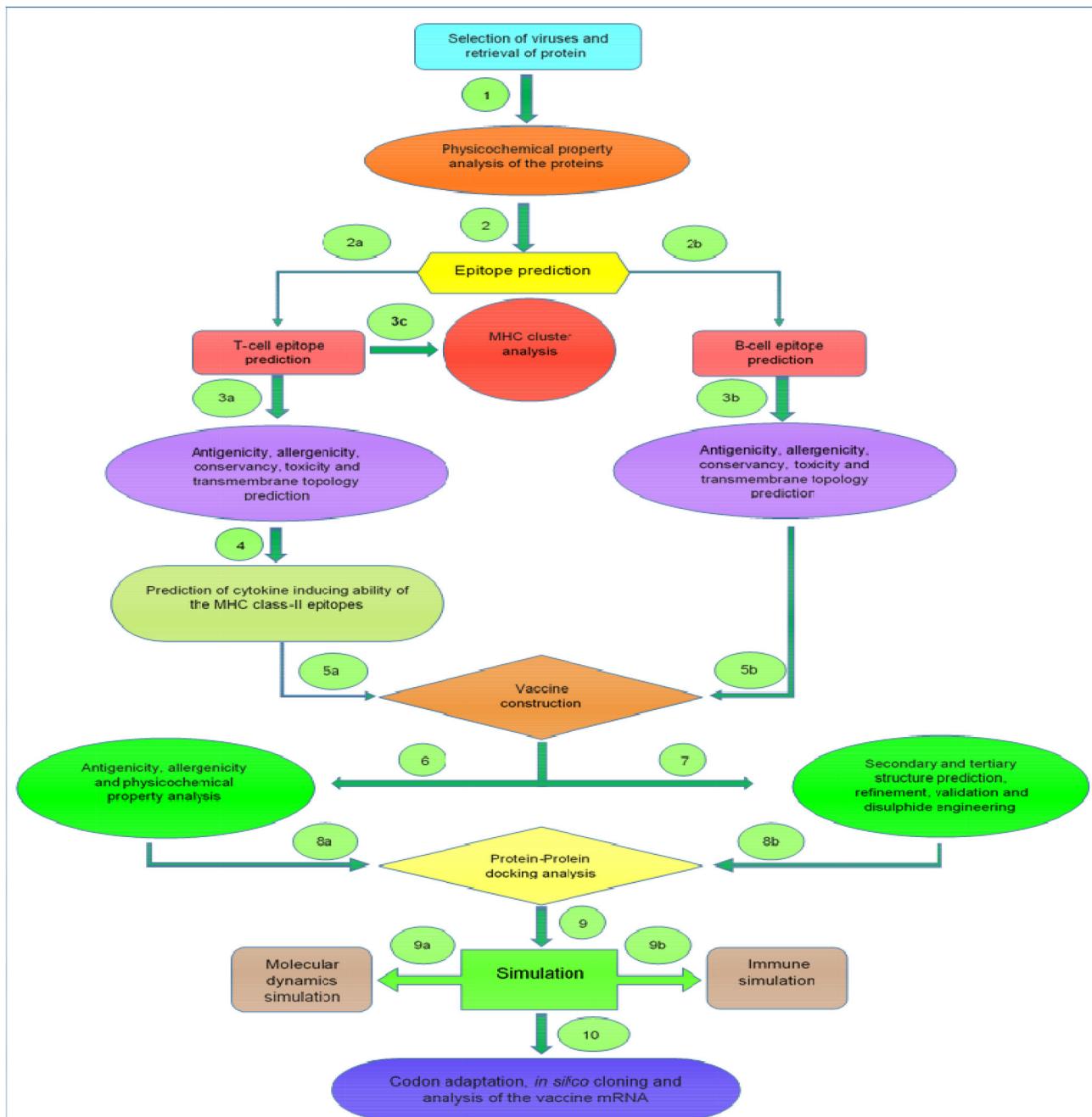


Figure 1. The step-by-step procedures adapted in the study for designing multivalent and multi-pathogenic vaccines.

version 2.0 (<http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.htm>) was used for determining the antigenicity of the epitopes (Doytchinova & Flower, 2007a, 2008). Two different tools i.e. AllerTOP version 2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>) as well as AllergenFP version 1.0 (<http://ddg-pharmfac.net/AllergenFP/>) were used for allergenicity prediction of the epitopes (Dimitrov, Bangov, et al., 2014; Dimitrov, Naneva, et al., 2014). ; The conservancy module of the IEDB server (<http://tools.iedb.org/conservancy/>) was used for the conservancy prediction and ToxinPred server (<http://crdd.osdd.net/raghava/toxinpred/>) was used for toxicity prediction of the epitopes. The support vector machine (SVM) method was used for the toxicity prediction keeping all the parameters default (Bui et al., 2006; Gupta et al., 2013). The epitopes that followed the previously mentioned criteria were

considered as the best-selected epitopes or the most promising epitopes and analyzed further for vaccine construction.

2.5. IFN-gamma, IL-4, and IL-10 inducing capacity and transmembrane topology prediction

Different types of cytokines like the interferon-gamma (IFN-gamma), IL-4, IL-10, etc., are activated and produced by the helper T-cells. Upon activation or production, these cytokines later activates other immune cells like the macrophages, cytotoxic T-cells, etc. (Luckheeram et al., 2012). Therefore, the cytokine inducing capability can be considered as an important characteristic of the MHC class-II or HTL epitopes. The IFNepitope (<http://crdd.osdd.net/raghava/ifnepitope/>) server was employed for the IFN-gamma induction capability prediction of all the predicted epitopes. For

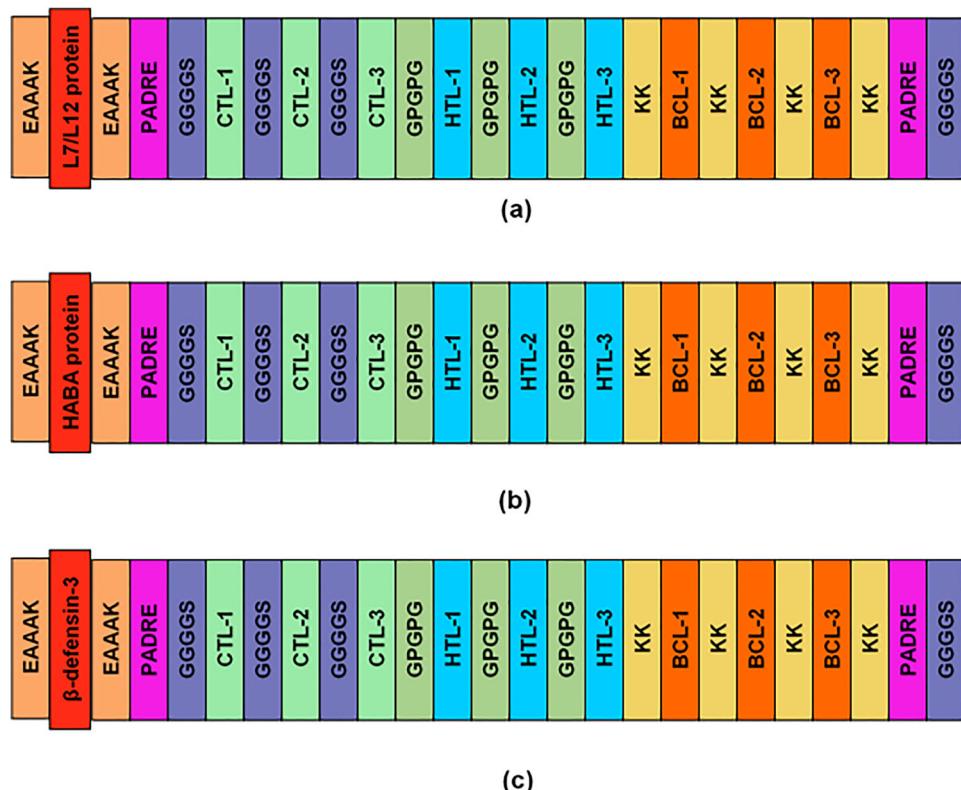


Figure 2. A schematic representation of the three vaccine constructs with appropriate linkers (EAAAK, GGGGS, GPGPG and KK), PADRE sequence, adjuvants (L7/L12 protein, HABA protein and human beta-defensin-3) and epitopes (CTL, HTL and BCL) in their sequential manner. (a) is the first vaccine constructed with the L7/L12 protein adjuvant, (b) is the second vaccine constructed with the HABA adjuvant protein, and (c) is the third vaccine constructed with the human beta-defensin-3 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.

predicting the IFN-gamma inducing ability, the Design module was used and Hybrid (Motif + SVM) prediction approach was selected (Dhanda et al., 2013). Thereafter, the IL-4 and IL-10 inducing properties of the HTL epitopes were predicted using IL4pred and IL10pred servers, setting the threshold values at 0.2 and -0.3, respectively (default) (Dhanda et al., 2013; Nagpal et al., 2017). After that, the transmembrane topology experiment of all the epitopes was performed using the TMHMM version 2.0 server (<http://www.cbs.dtu.dk/services/TMHMM/>) (Möller et al., 2001).

2.6. Cluster analysis of the MHC alleles

Cluster analysis of various MHC alleles predicts the relationship among them by exploiting their binding specificities. In our study, the online tool MHCcluster version 2.0 (<http://www.cbs.dtu.dk/services/MHCcluster/>) was used for the cluster analysis. The server generates the results in phylogenetic manner and gives output in the form of MHC specificity tree and MHC specificity heat-map (Thomsen et al., 2013). During the analysis, the number of peptides and bootstrap calculations were set to the default values and all the HLA supertype representatives (MHC class-I) and HLA-DR representatives (MHC class-II) were selected.

2.7. Vaccine construction

The most promising epitopes from Sub-section 2.3 were conjugated together for constructing the possible vaccines.

Different linkers i.e. EAAAK, GGGGS, GPGPG, and KK were used for joining the CTL, HTL, and BCL epitopes. The EAAAK linkers effectively separate different domains of bifunctional fusion proteins (Arai et al., 2001). Moreover, the GPGPG linkers have the capability to facilitate the immune processing and presentation of their conjugated epitopes. GPGPG linkers also prevent the generation of junctional epitopes (Saadi et al., 2017). On the other hand, the bi-lysine (KK) linkers permit the independent immunological activities of the epitopes of a vaccine construct and the GGGGS linkers aid in conferring resistance to proteases, thus making the vaccines more stable (Gu et al., 2017; Wen et al., 2013).

Three different vaccines were constructed, containing three different adjuvants: L7/L12 ribosomal protein, HABA protein (*Mycobacterium tuberculosis*, accession number: AGV15514.1), and human beta-defensin-3. The adjuvants were added to increase the potency of the vaccines since the adjuvants are known to stimulate the activities of toll-like receptors (TLRs)-1, 2, and 4 (Funderburg et al., 2007; Hancock et al., 2012; Hajighahramani et al., 2017; Lee et al., 2014; Pandey et al., 2016; Rana & Akhter, 2016; Toussi & Massari, 2014). Pan-HLA-DR epitope (PADRE) sequence was also used in the vaccine construction since it enhances the CTL response of the vaccine (Ullah et al., 2020; Wu et al., 2010). **Figure 2** represents a schematic diagram of the three vaccines with their epitopes, adjuvants and linkers in their appropriate orientations. The three vaccines differ from each other only in their adjuvant sequences.

2.8. Antigenicity, allergenicity, and physicochemical property analyses of the vaccines

The constructed vaccines should be highly antigenic to provide better immune response. So, the antigenicity of the three vaccine constructs was predicted using two different servers i.e. VaxiJen version 2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htm>) and the ANTIGENpro module of the SCRATCH protein predictor (<http://scratch.proteomics.ics.uci.edu/>), where the parameters were kept default (Doytchinova & Flower, 2007a, 2008; Magnan et al., 2010). Thereafter, the allergenicity of the three vaccine constructs was determined by three different servers to enhance the prediction accuracy i.e. AlgPred (<http://crdd.osdd.net/raghava/algpred/>), AllerTop version 2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>), and AllergenFP v1.0 (<http://ddg-pharmfac.net/AllergenFP/>). In the AlgPred server, the MEME/MAST Motif prediction approach was used (Dimitrov, Bangov, et al., 2014; Dimitrov, Naneva, et al., 2014; Saha & Raghava, 2006). The solubility of the vaccine proteins upon over-expression in *Escherichia coli* was predicted by the SOLpro module of the SCRATCH protein predictor (<http://scratch.proteomics.ics.uci.edu/>) and cross-checked by the Protein-sol server (<https://protein-sol.manchester.ac.uk/>) (Hebditch et al., 2017). Furthermore, ProtParam (<https://web.expasy.org/protparam/>) tool was again used for predicting the physicochemical properties of the constructed vaccines (Gasteiger et al., 2005).

2.9. Secondary and tertiary structure prediction of the vaccines

Several tools i.e. PRISPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) (using PRISPRED version 4.0 prediction method) (Buchan & Jones, 2019; Jones, 1999), GOR IV (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html) (Garnier et al., 1996), SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Geourjon & Deleage, 1995), and SIMPA96 (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_simpa96.html) (Levin et al., 1986) were used for the secondary structure prediction of the vaccines so that the accuracy of the prediction would be improved. These tools are easy and efficient online servers to predict the amount of amino acids in alpha-helix, beta-sheet and coil structure formations and in all the tools, the default values were used for conducting the prediction. After that, the RaptorX (<http://raptorg.uchicago.edu/>) server was used for predicting the tertiary structures of the vaccine constructs. The server generates *p* value for each of the predicted structure which evaluates the quality of the model. The lower *p* value represents better model quality (Källberg et al., 2012; Ma et al., 2013; Peng & Xu, 2011).

2.10. Tertiary structure refinement and validation

When the tertiary structures are predicted by computational methods, they may lack their true, native structures. For this reason, the protein structures are refined so that these structures may closely resemble the native proteins. The 3D

structures of the vaccine constructs were refined by GalaxyRefine module of the GalaxyWEB server (<http://galaxy.seoklab.org/>). This server uses CASP10 tested refinement method and dynamics simulation to provide better-refined structures (Ko et al., 2012; Nugent et al., 2014). Later, two online tools i.e. PROCHECK (<https://servicesn.mbi.ucla.edu/PROCHECK/>) and ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) were used for validating the refined structures. The PROCHECK server evaluates the quality of the protein by analyzing the Ramachandran plot and the ProSA-web expresses the protein quality by generating z-score. A z-score within the range of the z-scores of all experimentally determined protein chains in current PDB database corresponds to better quality of a query protein (Laskowski et al., 2006; Morris et al., 1992; Wiederstein & Sippl, 2007).

2.11. Vaccine protein disulfide engineering

Disulfide engineering of the vaccine constructs was performed for predicting the potential amino acid pairs capable of undergoing disulfide bond formation. Disulfide bonds are important characteristic of a protein since they aid in improving the conformational stability of that protein. The disulfide engineering was conducted by Disulfide by Design 2 (<http://cptweb.cpt.wayne.edu/DbD2/>) online server (Craig & Dombkowski, 2013). During the analysis, the χ_3 angle was kept -87° or $+97^\circ \pm 5$ (to discard the undesired disulfides which were predicted by the default angle). Moreover, $\text{C}\alpha\text{-C}\beta\text{-S}\gamma$ angle was set at its default value of $114.6^\circ \pm 10$, since the $\text{C}\alpha\text{-C}\beta\text{-S}\gamma$ angle was estimated to reach a peak near 115° and covers a range from angles 105° – 125° in known disulfides. Finally, the pairs showing energy value less than 2.2 kcal/mol were mutated to cysteine residues to form the disulfide bonds (Petersen et al., 1999). The energy value was kept at 2.2 kcal/mol because 90% of native disulfide bonds are generally found to have energy value less than 2.2 Kcal/mol (Craig & Dombkowski, 2013).

2.12. Protein–protein docking

In protein–protein docking, the vaccine constructs were docked with multiple TLRs to find out one best vaccine construct, selected based on its superior performances in the docking experiment. The protein–protein docking was carried out using numerous online docking tools for improving the accuracy of the docking. In this experiment, the vaccine constructs were docked with TLR-1 (PDB ID: 6NIH), TLR-2 (PDB ID: 3ATC), TLR-3 (PDB ID: 2AOZ), TLR-4 (PDB ID: 4G8A), TLR-8 (PDB ID: 3W3M), and TLR9 (PDB ID: 4QDH). Best performed vaccine in the docking study was considered as the ‘best vaccine construct’. At first, the docking was carried out by ClusPro version 2.0 (<https://cluspro.bu.edu/login.php>) (Kozakov et al., 2017; Vajda et al., 2017). The lower the energy score generated by the server, the higher the binding affinity of the targets with their receptor(s) (Yuriev & Ramsland, 2013). Thereafter, the docking was again performed by PatchDock (<https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php>) and later refined and re-scored by

FireDock server (<http://bioinfo3d.cs.tau.ac.il/FireDock/php.php>). The lower the global energy, the better the binding affinity among the vaccine proteins and their receptors and vice versa (Atapour et al., 2019; Duhovny et al., 2002; ; Schneidman-Duhovny et al., 2005). Finally, the docking was carried out using HawkDock server (<http://cadd.zju.edu.cn/hawkdock/>) for the third time. The Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) study was also conducted using the same server. According to the server, the lower scores and lower energy correspond to better binding affinities (Feng et al., 2017; Hou et al., 2011; Sun et al., 2014; Weng et al., 2019). The best-docked vaccine complex with TLR-8 was visualized by Discovery Studio Visualizer (Biovia et al., 2000). The TLR-8 docked with the best vaccine (TLR8-best vaccine complex) was selected randomly from the group of the TLRs for visualization and molecular dynamics (MD) simulation study. The vaccine showing best performances in the docking was considered as the best vaccine construct.

2.13. Screening of conformational B-cell epitopes

Antibody-mediated humoral immunity is initiated within the body when the B-cells interact with their epitopic counterparts. So, the vaccines having conformational B-cell epitopes should provide better immune response. The conformational B-cell epitopes of the best-predicted vaccine construct from the 2.12 sub-section were determined by IEDB ElliPro tool (<http://tools.iedb.org/ellipro/>), keeping all the parameters at default of the minimum score of 0.5 and maximum distance of 6 angstrom (Ponomarenko et al., 2008).

2.14. Molecular dynamics (MD) simulation

The MD simulation was applied on TLR8-vaccine complex to understand any state changes in a given biological environment. For this, the Linux-based program GROMACS (GROningen MAchine for Chemical Simulations) was used (Abraham et al., 2015). The MD simulation goes through many stages of topology creation and energy minimization. The protein 'pdb' file was first cleaned to remove any environmental substrates and crystal water molecules and the protein topology was generated using the force-field, optimized potential for liquid simulation-all atom (OPLS-AA) force field (Kaminski et al., 2001). The protein structure was then positioned in the center of a cube of 1 nm from the edge to generate its periodic image 2 nm apart. The volume of this simulation environment was found to be 4958.00 nm³. Water molecules, spatially placed with a force constant of 1000 kJ mol⁻¹ nm⁻², were used to simulate the biological environment. The total charge of the protein complex was calculated to be +14 with one chain having -10 and another chain having a charge of +24. This system was neutralized by adding 14 CL⁻ ions through ionization in which the Verlet scheme was selected (Spreiter & Walter, 1999). Thereafter, the energy minimization was conducted to stabilize the vaccine structure. Supplementary Figure S1 clearly shows that the system remained stable enough to conduct

further simulations; as the potential energy of the structure spontaneously reduced below the order of 10⁶. To stabilize the temperature, the number volume temperature (NVT) equilibration was performed for 500 ps and then the number pressure temperature (NPT) equilibration was carried out for 500 ps to calculate the pressure and density. The resulting structure was subjected to MD simulation for 50 ns. The root-mean-square deviation (RMSD) of backbone of the energy minimized structure was predicted and radius of gyration (Rg) was also calculated. All plots and simulation graphs were analyzed using the XmGrace and QtGrace tool (Turner & XmGrace, 2005).

2.15. Immune simulation

The immune simulation study was performed for the best vaccine construct to predict its immunogenicity and immune response profile. The C-ImmSim server (<http://150.146.2.1/C-IMMSIM/index.php>) was used for the immune simulation study which predicts the real-life-like immune responses using position-specific scoring matrix (PSSM) and machine learning techniques (Rapin et al., 2010). During the experiment, all the parameters were kept default. However, the time steps were set at 1, 84, and 170 (time step 1 is injection at time = 0 and each simulation step represents 8 h) and the number of simulation steps was set at 1050. So, three injections would be given at 4 weeks apart because 4 weeks interval is recommended between two doses of most of the commercial vaccines (Castiglione et al., 2012). The Simpson's Diversity index, D was calculated from the figures.

2.16. Codon adaptation and in silico cloning

In codon adaptation, the best vaccine construct was reverse translated to the possible DNA sequence so that the vaccine can be produced in a mass quantity in a target organism. Since the cellular mechanisms of one organism differ from another organism and the same amino acid can be encoded by different codons in different organisms (codon bias), the codon adaptation experiment is conducted to find out the best codons to encode a protein within a particular organism. The Java Codon Adaptation Tool or JCAT server (<http://www.jcat.de/>) was used for codon adaptation experiment which ensures the maximal expression of a protein in a target organism (Angov, 2011; Grote et al., 2005; Khatoon et al., 2017). During codon adaptation, prokaryotic *E. coli* strain K12 was selected and at the same time, rho-independent transcription terminators, prokaryotic ribosome binding sites, and Eael and StyI cleavage sites of restriction enzymes were avoided at the server. Then the optimized DNA sequence was retrieved and Eael and StyI restriction sites were attached to the N-terminal and C-terminal sites, respectively. Finally, the SnapGene restriction cloning software was used to insert the DNA sequence between the Eael and StyI restriction sites of pETite vector (Solanki & Tiwari, 2018). Since the plasmid has sequences for the small ubiquitin-like modifier or SUMO-tag and 6X-His tag, the expressed vaccine

Table 1. List of envelope protein E sequences retrieved from the DENV serotypes and the ZIKV.

Sl. No.	Name of the virus	Reference ID	Length (aa)
1	Dengue virus type-1	NP_722460.2	495
2	Dengue virus type-2	NP_739583.2	495
3	Dengue virus type-3	YP_001531168.2	493
4	Dengue virus type-4	NP_740317.1	495
5	Zika virus	YP_009227198.1	500

protein would also contain these tags, which might facilitate its solubilization and affinity purification (Biswal et al., 2015).

2.17. Prediction of the mRNA secondary structure of the vaccine

The mRNA secondary structure prediction was conducted by two different servers i.e. Mfold (<http://unafold.rna.albany.edu/?q=mfold>) and RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>). These servers carry out the thermodynamic analysis of the mRNA secondary structures and provide minimum free energy (ΔG Kcal/mol) as output. The lower minimum free energy corresponds to more stably folded mRNA (Gruber et al., 2008; Mathews et al., 1999, 2007; Zuker 2003). To predict the mRNA secondary structure of the best vaccine construct, at first the optimized DNA sequence from JCat server was converted to possible RNA sequence by DNA<->RNA->Protein tool (<http://biomodel.uah.es/en/lab/cybertory/analysis/trans.htm>). Then the converted RNA sequence was copied and pasted in the Mfold and RNAfold servers for prediction using the default settings for all the parameters.

3. Results

3.1. Retrieval of envelope protein sequences

The envelope protein E sequences of all the four DENV serotypes-1, 2, 3, and 4 as well as ZIKV, were retrieved in FASTA format from the NCBI database (Table 1).

3.2. Physicochemical parameters of the retrieved proteins

The physicochemical property analysis showed that the proteins of DENV-1, 2, and 4 had quite similar predicted theoretical pI. The envelope glycoprotein E of ZIKV had the highest number of positively charged amino acids (51) as well as the lowest number of negatively charged amino acids (47). Proteins from all the four DENV serotypes had similar predicted half-life of 30 h in mammalian cell culture system. Moreover, both envelope proteins from DENV-4 and ZIKV were found to have the same predicted extinction coefficient of $72,140 M^{-1}cm^{-1}$. All the proteins were found to be stable and the protein from DENV-3 was predicted to have the lowest GRAVY value of -0.111 (Supplementary Table S1).

Table 2. List of the best-selected epitopes based on the selection criteria.

MHC class-I epitopes	MHC class-II epitopes	B-cell epitopes
KALKLSWFK	QWFQLDPLPWTSGAS	SNTTDSRCPTQGEATLVEEQDT
ALKLSWFKK	KQWFQLDPLPWTSGA	TNTTTESRCPQTGEPSLNEEQD
FSGVSWTMK	HKQWFQLDPLPWTSG	IVTEKDSPVNIEAEPFGD
ISNTTTDSR	VQVKYEGTDAPCKIP	NITTDSRCPTQGEAVLPPEEQDQN
SIQPENLEY	QVKYEGTDAPCKIPF	TATRCPTQGEPYLKKEEQDQQ
CVTVMAQDK	LVQVKYEGTDAPCKI	KMTGKSIQOPEN
GGVFNSLGK	YYLTMMNNKHVLVHKE	QYAGTDGPCKI
VEFKDAHAK	LYYLTMMNNKHVLVHK	-
TCAKFTCSK	DLYYLTMMNNKHVLVH	-

3.3. Selection and characterization of T-cell and B-cell epitopes

When predicting the MHC class-I and MHC class-II epitopes of proteins from DENV serotypes, the envelope protein E from DENV-1 was selected as the model of the prediction. From the IEDB server, top 20 MHC class-I and MHC class-II epitopes were taken. And also top twenty MHC class-I and MHC class-II epitopes were selected for the envelope protein E of ZIKV (Supplementary Tables S2–S5). In addition, the B-cell epitopes were selected for each of the viral proteins (Supplementary Figure S2). Finally, the best considered or most promising epitopes with high antigenicity, non-allergenicity, non-toxicity, and 100% conservancy were selected for further analysis (Table 2). Furthermore, about 60% of all the epitopes were predicted to be located inside the cell-membrane.

3.4. Epitopes were IFN-gamma, IL-4 and IL-10 inducer

Comprehensive analysis of cytokine inducing ability revealed that most of the selected epitopes were good inducers of IFN-gamma, IL-4, and IL-10. All the epitopes selected for vaccine construction were also found to have at least one of these cytokine inducing capability (Supplementary Tables S2 and S4).

3.5. Cluster analysis of the MHC alleles

The cluster analysis of the MHC class-I and MHC class-II alleles that may interact with the predicted epitopes revealed significant interactions among the MHC class-I and MHC class-II alleles (Supplementary Figure S3).

3.6. Joining of the T-cell and B-cell epitopes and final vaccine construction

Epitopes of the T-cell and B-cell were joined together by different linkers for vaccine construction. The CTL epitopes or MHC class-I epitopes were conjugated by GGGGS linkers, whereas, GPGPG linkers were used to conjugate the HTL or MHC class-II epitopes. The BCL epitopes were joined together by KK linkers, and the adjuvants and PADRE sequence were linked to the rest of the vaccine through EAAAK linker. The vaccines were designed according to the outline provided in Figure 2. The three constructed vaccines were designated as: V1, V2 and V3 and they differ from each other only in their adjuvant sequences (Table 3).

Name of the vaccine	AN	AG	pI	Number of positively charged amino acids	Number of negatively charged amino acids	Extinction coefficients (in M ⁻¹ cm ⁻¹)	Half-life	II	AI	GRAVY	Solubility
V1	Antigen	Non-allergen	8.28	70	74	95,505	1 h (mammalian reticulocyte), >10 h in <i>E. coli</i>	34.53 (stable)	61.19	-0.520	Soluble (SolPro: 0.988, Protein-Sol: 0.55)
V2	Antigen	Non-allergen	8.14	76	79	99,975	1 h (mammalian reticulocyte), >10 h in <i>E. coli</i>	39.73 (stable)	59.71	-0.657	Soluble (SolPro: 0.990, Protein-Sol: 0.55)
V3	Antigen	Non-allergen	9.29	47	71	98,860	1 h (mammalian reticulocyte), >10 h in <i>E. coli</i>	39.05 (stable)	51.27	-0.704	Soluble (SolPro: 0.996, Protein-Sol: 0.55)

AN: antigenicity; AG: allergenicity; pI: theoretical pI; II: instability index; AI: aliphatic index; GRAVY: grand average of hydropathicity

3.10. Stability improvement of the vaccines

The vaccine protein disulfide engineering of the three vaccine proteins revealed that the V1 construct had three pairs of amino acids (200 Thr-209 Ser, 368 Ile-388 Cys, and 442 Met-447 Trp), V2 construct had five pairs of amino acids (1 Glu-5 Ala, 250 Gly-289 Gly, 326 Gln-345 His, 435 Leu-446 Tyr, and 618 Gly-620 Cys) and V3 construct had 4 pairs of amino acids (113 Ser-130 Asn, 170 His-174 Gly, 336 Met-341 Trp, and 362 Trp-382 Ala) with the energy value less than 2.2 kcal/mol; capable of forming potential disulfide bonds among themselves. Thus, these pairs of amino acids were selected and mutated to form the disulfide bonds ([Supplementary Figure S7](#)). V2 was predicted to be most stable than the other two vaccines, as it can form five pairs of disulfide bonds.

3.11. Protein–protein docking study

The molecular docking between the vaccine proteins and multiple TLRs was performed for selecting the best vaccine construct with the capability to interact with these TLRs efficiently during an immune response. Several different online tools were used for the docking analysis to improve the prediction accuracy. From the docking study, V3 vaccine was found to be the best vaccine construct among the three constructed vaccines. It generated the best results with all the receptors, when analyzed by the HawkDock server as well as the MM-GBSA study. Furthermore, it also generated the very good results when the docking was conducted by ClusPro version 2.0 and PatchDock servers. Most importantly, V3 was predicted to be the best performer by all the servers when docked with the TLRs. Therefore, V3 was considered as the best vaccine construct ([Table 5](#)). [Figure 3](#) illustrates the interaction of the best vaccine V3 and the receptor protein TLR-8 along with their interacting amino acids.

3.12. Screening for conformational B-cell epitopes

In conformational B-cell epitope prediction, total 317 residues were found with scores varying from 0.502 to 0.698. The epitopes were ranged from 5 to 156 amino acid residues and predicted to be located within five conformational B-cell epitopes ([Table 6](#), [Supplementary Figure S8](#)).

3.13. Molecular dynamics (MD) simulation study

The MD simulation was conducted on the TLR8-V3 docked complex. The total mass of the complex after applying the OPLS-AA force field was found to be 141,680.069 amu. A total of 157,763 water molecules were added into the system during solvation from which 14 were replaced by CL ions during ionization.

Energy minimization was completed in 2086 steps when the steepest descent converged and the force reached <1000 kJ/mol. The potential energy was measured to be -8.7690670e + 06 kJ/mol. From the temperature equilibration plot of [Supplementary Figure S9\(a\)](#), it was observed that the target value of 300 K remained stable over the remainder of the equilibration and fluctuates only by ±1 K. The pressure value found from NPT equilibration showed fluctuations around 0 bar with a range of ±125 bar ([Supplementary Figure](#)

Table 5. Results of the molecular docking between the three vaccine constructs and their selected receptors.

Name of the vaccines	Name of the targets	PDB IDs of the targets	ClusPro energy score	Global energy	HawkDock score (the lowest score)	MM-GBSA (binding free energy, in kcal mol ⁻¹)
V1	TLR-1	6NIH	-977.5	-17.58	-5106.94	-28.19
	TLR-2	3A7C	-978.6	-8.12	-5817.30	-68.15
	TLR-3	2A0Z	-1045.6	-4.71	-5450.66	-31.39
	TLR-4	4G8A	-993.4	-2.71	-5375.24	-63.80
	TLR-8	3W3M	-926.2	-0.39	5179.50	-35.32
	TLR-9	4QDH	-988.7	-19.23	-5850.40	-62.52
	TLR-1	6NIH	-865.0	-13.43	-3248.84	-18.30
	TLR-2	3A7C	-1008.1	-3.47	-3766.88	-52.33
	TLR-3	2A0Z	-904.0	-1.28	-3804.33	-36.22
V2	TLR-4	4G8A	-912.5	-7.28	-3795.29	-40.44
	TLR-8	3W3M	-834.30	-13.45	-4914.25	-29.64
	TLR-9	4QDH	-817.6	-17.30	-3538.55	-50.96
	TLR-1	6NIH	-1071.2	-25.01	-5621.91	-33.48
	TLR-2	3A7C	-1170.9	-15.56	-6146.50	-80.41
	TLR-3	2A0Z	-1098.4	-9.73	-6281.47	-71.20
	TLR-4	4G8A	-927.6	-1.16	-5877.17	-91.85
	TLR-8	3W3M	-1009.0	-26.50	-5933.40	-45.42
	TLR-9	4QDH	-982.1	-26.38	-6244.94	-69.75
Remarks	-	-	Best vaccine construct: V3 (considering the docking scores with the targets except TLR-4 and TLR-9)	Best vaccine construct: V3 (considering the docking scores with the targets except TLR-4)	Best vaccine construct: V3 (considering the docking scores with all the targets)	Best vaccine construct: V3 (considering the docking scores with all the targets)

S9(b)). This behavior is not unexpected. The running average of the data over a 50 ps window also showed similar fluctuations. As with pressure, the running average of density was plotted in red line (Supplementary Figure S9(c)) and the average value over the course of 500 ps was $1090 \pm 3 \text{ kgm}^{-3}$. The density values were mostly stable over time, indicating that the system was well equilibrated.

Trajectory analysis was conducted after a 50 ns simulation. After performing trajectory conversion to account for periodicity in the system RMSD calculation was done (Figure 4(a)). A plot of RMSD backbone revealed that the RMSD levels go up to $\sim 1.25 \text{ nm}$ and were maintained during the course of the simulation. The black line refers to the RMSD relative to the structure present in the minimized, equilibrated system. The red line is the RMSD relative to the crystal structure. Both these plots were almost identical with little to no difference during the experiment which indicated that the structure was very stable. Subtle differences between the plots indicated that the structure at $t = 0 \text{ ns}$ was slightly different from this crystal structure. This is to be expected, since it was energy-minimized, and because the position restraints were not 100% perfect. The RMS Fluctuation (RMSF) and the radius of gyration also showed quite sound and satisfactory results (Figure 4(b,c)). The radius of gyration of the vaccine reached a peak of 4.75 nm but remained within a boundary of $4.5\text{--}4.7 \text{ nm}$ over the course of the simulation. This signifies that the protein loses some of its compactness as it unfolds. RMSF, on the other hand, explains regions with high flexibility and the RMSF graph also showed quite sound performance of the TLR8-V3 docked complex in terms of flexibility during the MD simulation.

3.14. Immune simulation

The immune simulation of the best-selected vaccine V3 was conducted by the C-ImmSimm server which predicts the

generation of adaptive immunity as well as the immune interactions of epitopes with their specific targets (Rapin et al., 2010). The immune simulation study showed that after each of the three injections of the V3 vaccine, the primary immune response against the antigenic segments of the vaccine was predicted to rise significantly as indicated by the gradual elevation in concentrations of different immunoglobulins (Figure 5(a)). Furthermore, the secondary immune response was also found to be increased in response of the primary immune response. The gradual increase in the concentrations of active B-cell (Figure 5(b,c)), plasma B-cell (Figure 5(d)), helper T-cell (Figure 5(e,f)), regulatory T-cell and cytotoxic T-cell (Figure 5(g-i)) were also found, which pointed towards a very strong secondary immune response, very good immune memory generation, and the increased clearance of the antigen. Furthermore, the augmentation in the concentration of dendritic cells and macrophages indicated a very good antigen presentation by these antigen-presenting cells or APCs (Figure 5(j,k)). The vaccine was also found to be able to produce good amount of different types of cytokines like the IFN-gamma, IL-23, IL-10, and IFN-beta which are some of the most significant cytokines for generating immune response against viruses (Figure 5(l)). Overall, the immune simulation study revealed that with the predicted ability of generating high amount of immunoglobulins, cytokines as well as APCs, active B-cells, and T-cells, the multivalent and multipathogenic vaccine V3 might be able to provide good immunogenic protection against the targeted viruses.

3.15. Codon adaptation, *in silico* cloning, and analysis of the mRNA structure

The codon adaptation of the best-predicted vaccine V3 for *E. coli* strain K12, yielded the excellent codon adaptation index

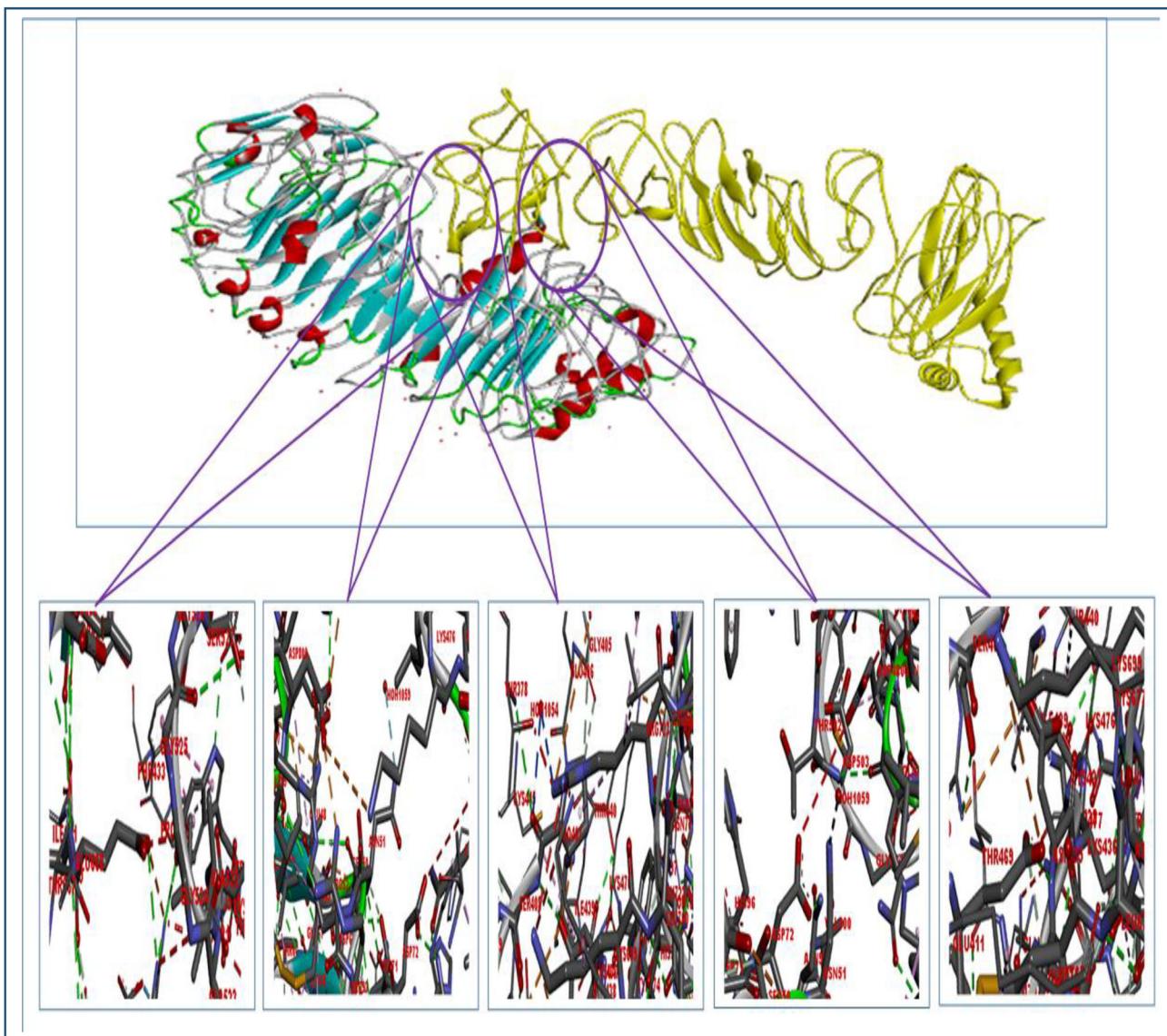


Figure 3. Interaction of the V3 vaccine construct (ligand in yellow color) with TLR-8 (receptor). Glu 411 (ligand)- Lys 699 (receptor, bond length: 1.49 Å), Thr 378 (ligand)- Arg 723 (receptor, bond length: 2.22 Å), Pro 407 (ligand)- Arg 723 (receptor, bond length: 1.17 Å), Glu 406 (ligand)- Arg 723 (receptor, bond length: 2.76 Å), Asp 480 (ligand)- Asp 72 (receptor, bond length: 1.67 Å), Lys 476 (ligand)- Asp 800 (receptor, bond length: 2.02 Å), Gly 524 (ligand)- Glu 668 (receptor, bond length: 1.65 Å), and Gly 525 (ligand)- Glu 668 (receptor, bond length: 1.98 Å).

(CAI) value of 1.0 and GC content of 52.8% (Figure 6). After codon adaptation, the newly adapted DNA sequence of the V3 vaccine construct was inserted between the Eael and StyI restriction sites of the pETite vector plasmid. Upon translation in the *E. coli* host with the pETite vector, the vaccine protein was expected to be expressed in fusion with SUMO protein and 6× His tag, which might help effective solubilization and purification of the vaccine protein (Biswal et al., 2015) (Figure 7). The newly constructed plasmid with the V3 sequence was designated as 'DZV_3' plasmid.

When the lowest minimum free energy of the best vaccine V3 was predicted by the Mfold server, the best-determined structure (among 39 generated structures by the server) of the optimized construct showed minimum free energy, ΔG value of -578.70 kcal/mol. The prediction of the Mfold server was in agreement with the result generated by RNAfold server, where the ΔG value of the mRNA structure was found to be -552.10 kcal/mol. Supplementary Figure

S10 illustrates the predicted mRNA secondary structure of the V3 vaccine.

4. Discussions

Vaccines are the widely administered and commonly produced pharmaceutical products all over the world to control and prevent the infectious diseases. Although, the classical or conventional approaches are mainly used for vaccine production, but these processes are costly and time consuming (María et al., 2017; Poland et al., 2009). However, technological revolution exploring many cutting-edge technologies as well as open genomic information hub of different pathogens have made it possible to design and develop novel 'subunit vaccines' (De Groot et al., 2002; Purcell et al., 2007; Zhang, 2018;). Such subunit vaccines are safe, effective, efficient, and inexpensive to be produced and used, compared to the conventional vaccines (Rappuoli, 2000). In this study, a

Table 6. List of the predicted B-cell epitopes of V3 vaccine construct.

No.	Residues	No of residues	Score	Figure
1	E1, A2, A3, A4, K5, G6, I7, I8, N9, T10, L11, Q12, K13, Y14, Y15, C16, R17, V18, R19, G20, G21, R22, C23, A24, V25, L26, S27, C28, L29, P30, K31, E32, E33, Q34, I35, G36, K37, C38, S39, T40, R41, G42, R43, K44, C45, C46, R47, R48, K49, K50, E51, A52, A53, A54, K55, A56, K57, F58, V59, A60, A61, W62, T63, L64, K65, A66, A67, A68, G69, G70, G71, S72, K73, A74, L75, K76, W79, F80, K81, G82, G83, G84, S85, A86, L87, W91, F92, K93, K94, G95, G96, G97, S98, F99, S100, G101, V102, M106, K107, G108, G109, G110, S111, T117, D118, S119, R120, G121, G122, E123, Y132, G134, G135, G136, S137, A143, Q144, D145, K146, G147, G148, G149, S150, G151, G152, V153, S156, G158, K159, G160, G161, G162, S163, V164, E165, K167, D168, A169, H170, A171, K172, G173, G174, G175, S176, T177, C178, A179, K180, S184, K185, G186, P187, G188, P189	156	698	Supplementary Figure S8 (a)
2	H361, D373, V385, E386, E387, Q388, D389, T390, K391, K392, T393, N394, C401, T403, Q404, G405, E411, Q413, D414, K415, K416, I417, V418, T419, E420, K421, D422, S423, P424, V425, N426, I427, E428, A429, E430, P431, P432, F433, G434, D435, K436, K437, N438, I439, T440, T441, D442, S443, R444, C445, G448, G449, E450, A451, V452, L453, P454, E455, Q457, D458, Q459, N460, K461, K462, T463, A464, T465, R466, C467, P468, T469, Q470, G471, E472, P473, Y474, L475, K476, E477, E478, Q479, D480, Q481, Q482, K483, K484, K485, M486, T487, G488, K489, S490, I491, Q492, P493, E494, N495, K496, K497, Q498, Y499, A500, G501, T502, D503, G504, P505, C506, K507, I508, K509, K510, A511, K512, F513, V514, A515, A516, W517, T518, L519, K520, A521, A522, A523, G524, G525, G526, S527	129	0.684	Supplementary Figure S8 (b)
3	T201, S202, G203, A204, S205, G206, T222, S223, G224, A225, G226, P227, G228	13	0.514	Supplementary Figure S8 (c)
4	K366, K367, S368, N369, T370	5	0.508	Supplementary Figure S8 (d)
5	G276, Y296, E297, G298, T299, D300, A301, T314, M315, N316, N317, K318, H319, K339	14	0.502	Supplementary Figure S8 (e)

package of immunoinformatics tool was exploited to design multivalent and multipathogenic vaccines which could work against multiple serotypes of the DENV (DENV-1, 2, 3, and 4) and ZIKV simultaneously.

The envelope protein E sequences of the DENV and ZIKV were targeted for vaccine design, as it is an essential protein for the virion assembly and viral replication especially for *Flaviviruses* (Klein et al., 2013). The physicochemical properties of the retrieved sequences were further analyzed. The theoretical pl refers to the pH at which a protein does not possess any net charge. The extinction co-efficient of a compound describes the amount of light that is absorbed by it at a specific wavelength (Gill & Von Hippel, 1989; Pace et al., 1995). The instability index of a compound refers to the probability of that particular compound to be stable and a compound with instability index over 40 is considered to be unstable (Guruprasad et al., 1990). The aliphatic index of a protein refers to the relative amount of amino acids in its side chains occupied by the aliphatic amino acids like alanine, valine, etc. (Ikai, 1980). Furthermore, the positive GRAVY value represents the hydrophobic nature of a compound, whereas the negative GRAVY value indicates the hydrophilic nature of a compound (Chang & Yang, 2013; Kyte & Doolittle, 1982). Envelope protein E from DENV-1, DENV-2, and DENV-4 had almost similar theoretical pl. All the proteins had very good instability index and aliphatic index profiles and therefore, all of them were predicted to be stable and have good amount of aliphatic amino acids in their side chains. Again, since all the proteins were found to have negative GRAVY values, it can be considered that all of the proteins were hydrophilic in nature.

A multi-epitope vaccine should contain CTL, HTL, and BCL epitopes in its structure to be an effective vaccine because these epitopes will stimulate the cytotoxic T-cells, helper T-cells, and antibody-producing B-cells. These cells are the

most important types of cells that function in immunity (Zhang, 2018). The B-cells produce antibodies and thus stimulate the humoral immune response. B-cells also keep the memory of a previous infection by a specific pathogen or antigen. However, the humoral immune response is not so robust like the cell-mediated immune response and may get weaker overtime (Bacchetta et al., 2005; Cooper & Nemerow, 1984). In such conditions, the cell-mediated immune response provides much broader and life-long immunity by secreting various antiviral cytokines and specifically identifying and destroying the infected cells (Arpin et al., 1995; Cano & Lopera, 2013; Garcia et al., 1999). For this reason, the T- and B-cell epitopes were determined to construct the multivalent, multipathogenic vaccines against both DENV and ZIKV. Since the IEDB server generated a lot of possible epitopes, some criteria were followed for selecting the best epitopes i.e. high antigenicity, non-allergenicity, non-toxicity, and 100% conservancy across strains and species. The epitopes should be highly antigenic because only the highly antigenic epitopes can generate potential immune responses. They also have to be non-allergenic and non-toxic so that they will not be able to cause any harmful allergenic and toxic reactions (Sarkar et al., 2020). Furthermore, their predicted 100% conservancy across different strains and species had ensured their broad-spectrum activity among multiple pathogens. The epitopes following these criteria were considered as the best-selected epitopes and used for vaccine construction. Again, almost all of the selected HTL epitopes were found to be at least one cytokine inducer (among the INF-gamma, IL-10, and IL-4). These cytokines are required for the activation, growth and differentiation of many immune cells like T-cells, B-cells, macrophages etc. (Luckheeram et al., 2012; Romagnani, 1997), so these epitopes might further improve the immunogenic activities of the constructed vaccines.

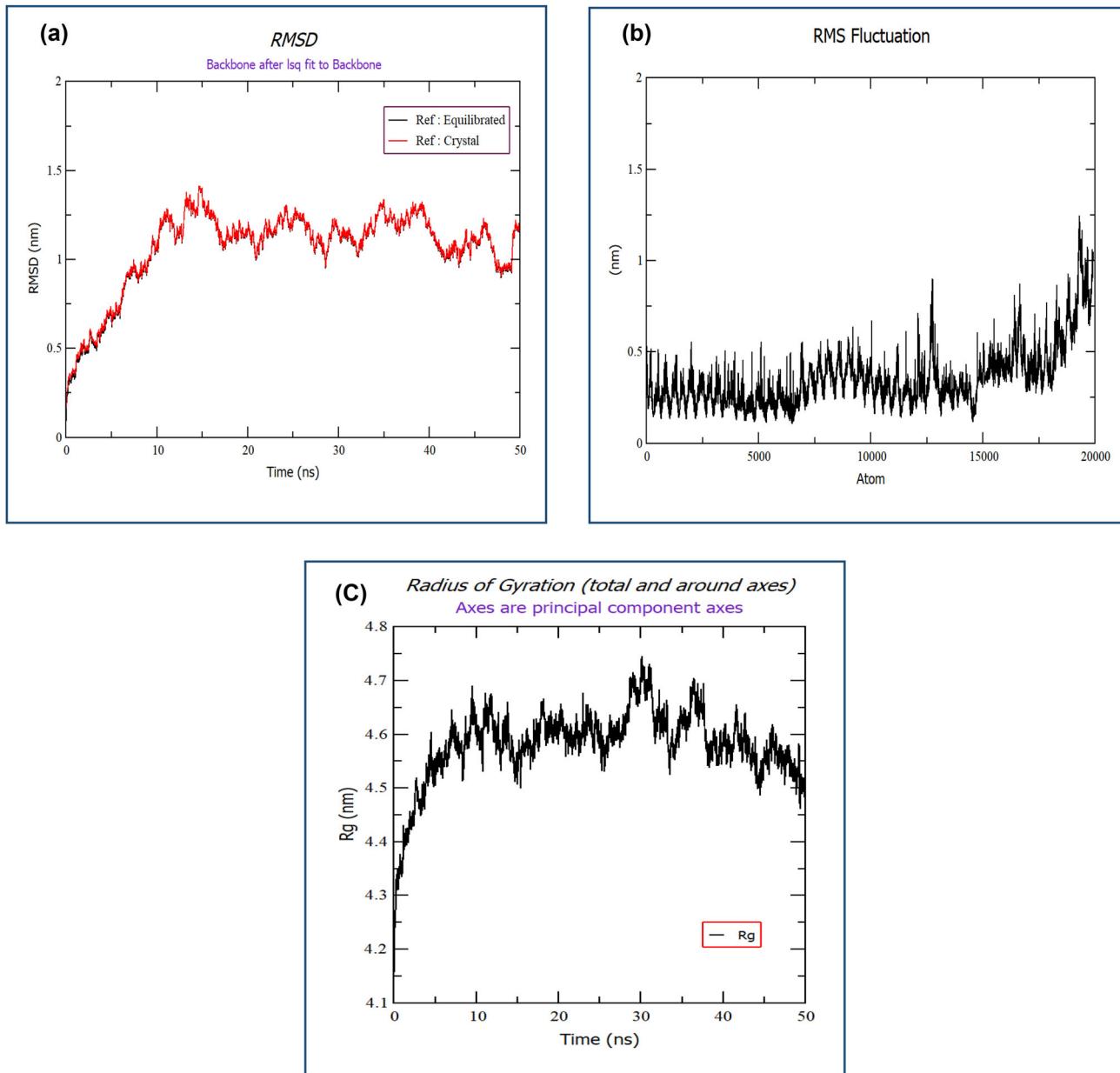


Figure 4. (a) RMSD plot of backbone. RMSD graph shows the structure maintains as Table 1.25 nm deviation with minimum fluctuations, (b) RMS Fluctuations of all the atoms about its average position. The peaks and dips in the graph denote the flexibility of the corresponding region in the molecular structure, (c) Radius of gyration around the principal component axes.

Thereafter, the cluster analysis of the MHC class-I and MHC class-II alleles was carried out which showed potential relationship among the selected MHC alleles. After the cluster analysis, all the best selected or most promising epitopes were used for vaccine construction. The EAAAK, GGGGS, GPGPG, and KK linkers were used at appropriate positions and adjuvants and PADRE sequence were used to enhance the vaccine potentiality. The three vaccines (V1, V2, and V3) differed from each other only in their adjuvant sequences.

After the successful vaccine construction, the antigenicity, allergenicity, and different physicochemical properties of the constructed vaccines were predicted. All the vaccine constructs were found to be potent antigens as well as non-

allergens. The physicochemical property analysis predicted that all the vaccines might be quite stable and all of them had predicted half-life of more than 10 h in *E. coli*. Therefore, these vaccines might not cause any problem during its mass production and purification in the *E. coli* cell culture system. Again, the aliphatic index represents the protein's thermal stability and higher aliphatic index of a protein corresponds to more thermostable structure (Panda & Chandra, 2012). Since all the vaccine constructs were predicted to have quite high aliphatic indexes, all of them were predicted to be quite thermostable. The theoretical pI of the vaccine constructs (all the vaccines had pI over 8.0) showed that they might be basic in nature. Moreover, with the negative GRAVY values, all of them were predicted to be hydrophilic in nature.

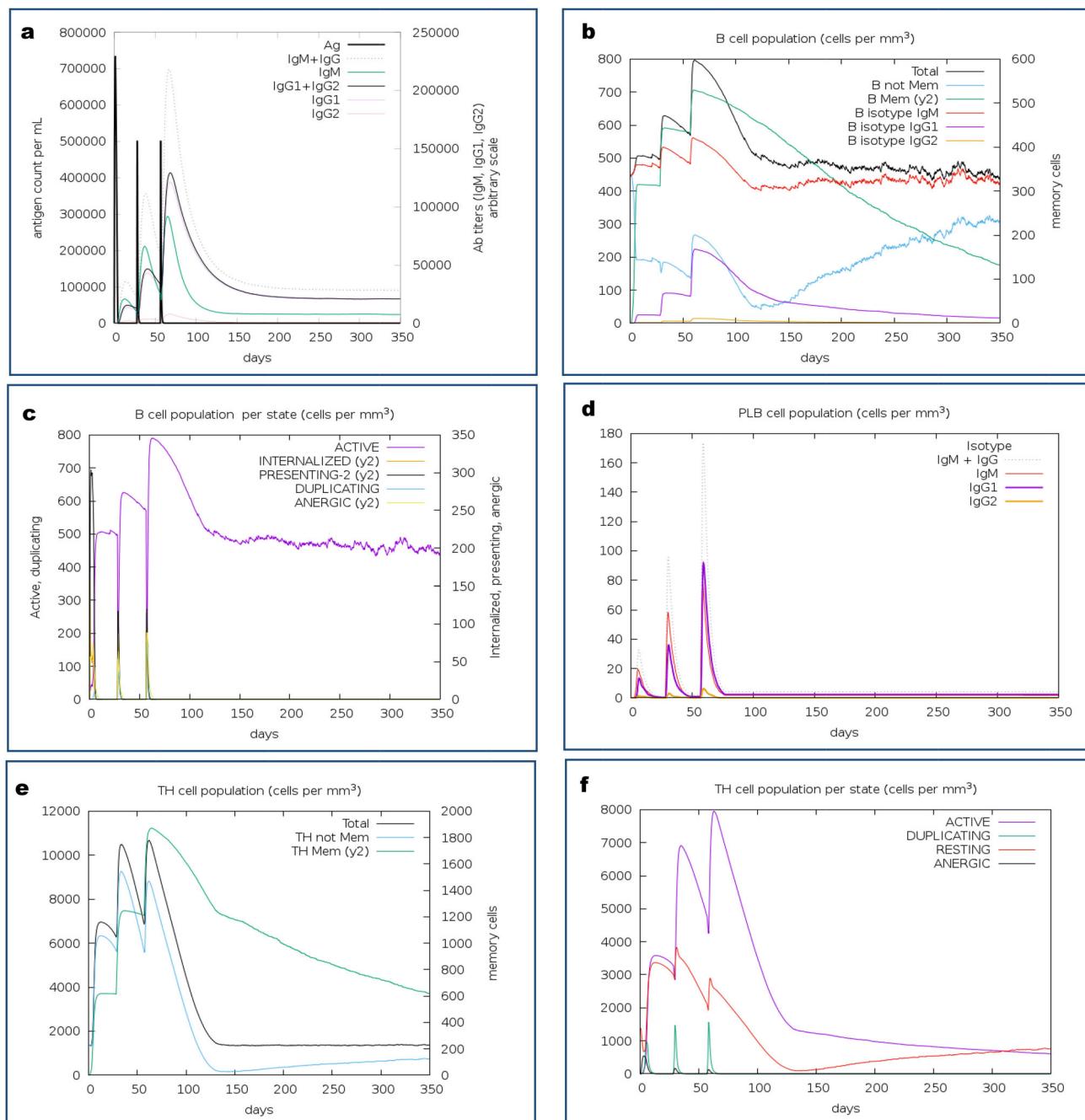


Figure 5. C-IMMSIMM representation of the immune simulation of the best-predicted vaccine, V3. a. The immunoglobulin and immunocomplex response to the V3 vaccine inoculations (black colored lines) and specific subclasses are indicated by colored lines, b. Elevation in the B-cell population over the course of the three injections, c. Increase of the B-cell population per state over the course of vaccination, d. Augmentation in the plasma B-cell population over the course of the injections, e. Increase in the helper T-cell population over the course of the three injections, f. Elevation of the helper T-cell population per state over the course of the vaccination, g. Rise in the regulatory T lymphocyte over the course of the three injections, h. Increase in the cytotoxic T lymphocyte population over the course of the injections, i. Elevation in the active cytotoxic T lymphocyte population per state over the course of the three injections, j. Rise in the active dendritic cell population per state over the course of the three injections, k. Increase in the macrophage population per state over the course of the three injections, l. Increase in the concentrations of different types of cytokines over the course of the three injections.

Solubility is a critical factor for the post-production studies of proteins. The more soluble a protein on overexpression, the easier its purification is. As all the proteins were found to have good solubility upon over-expression in *E. coli* by both servers (SolPro and Protein-Sol), so it can be declared that their purification steps should be much easier. Considering all these aspects of the physicochemical property analysis, the predicted vaccine constructs might be suitable to be chosen as the potential vaccine candidates.

The secondary and tertiary structure prediction of the three vaccine constructs revealed that the difference in their adjuvant sequences had caused some significant changes in their protein structures. The three vaccines had differences among the percentage of amino acids in their α -helix, β -strand, and coil structure formations. Again, with the lowest *p*-value of 3.48e-13, V1 vaccine was predicted to have the best quality protein structure among the three vaccine proteins by the RaptorX server. After the tertiary structure

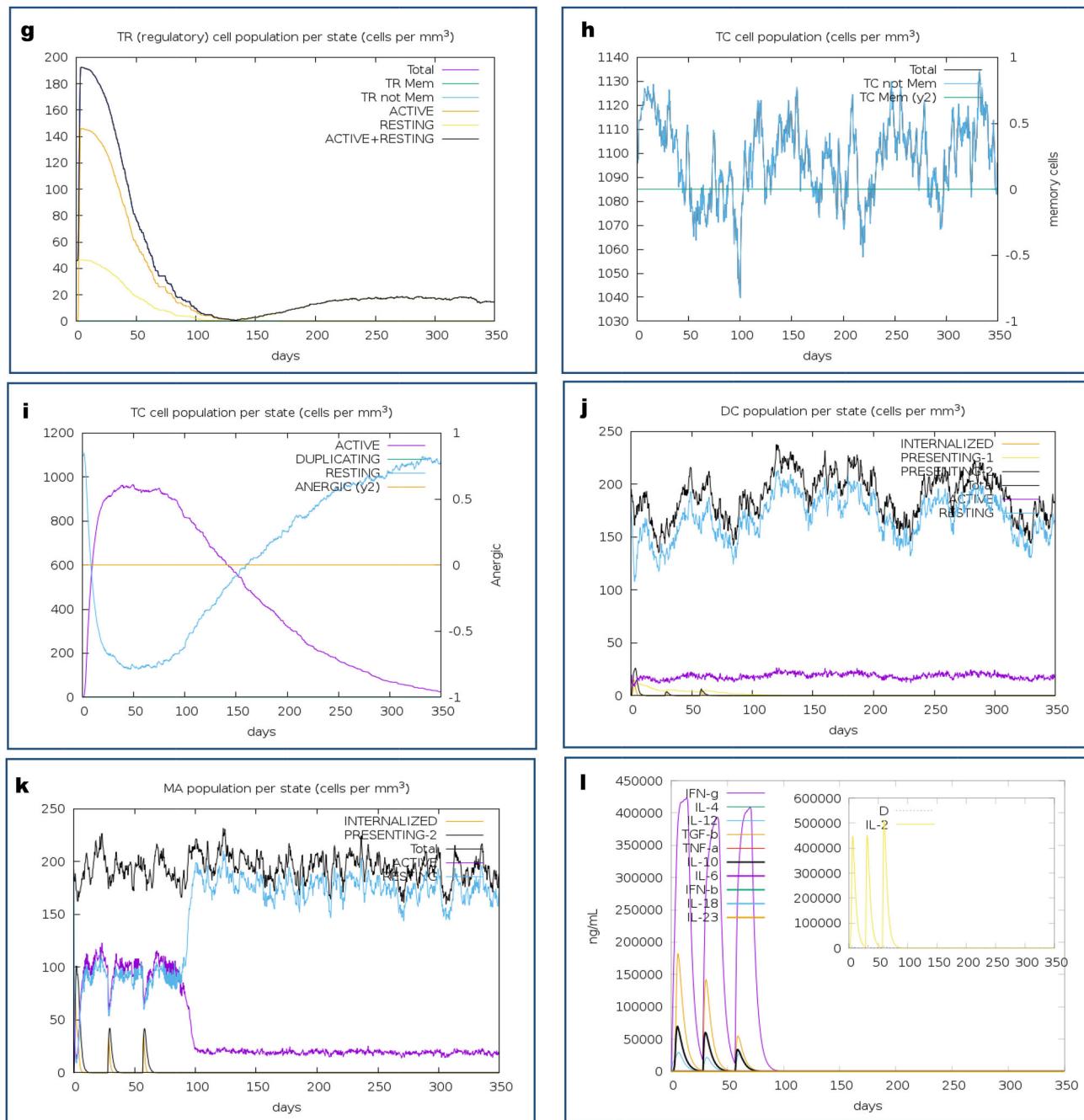


Figure 5. Continued

prediction, the structures were refined and validated, which revealed that the quality of all the three vaccine constructs was significantly improved after refinement in the context of GDT-HA, MolProbity, Rama favored amino acid percentage, and z-scores. Thereafter, when the protein disulfide engineering of the vaccine constructs was carried out to improve their stability, V2 vaccine (with five possible disulfide bonds) was found to be more stable than the other two vaccine constructs (each with four disulfide bonds).

The protein-protein docking between the constructed vaccines and various TLRs was conducted by several different online tools to improve the prediction accuracy. The docking is one of the most important steps in this experiment because from the docking study, one best vaccine construct

was identified for further analysis. From the docking study, V3 vaccine was considered as the best vaccine construct for its sound and satisfactory performances by all the servers. A vaccine construct should have conformational B-cell epitopes for providing strong humoral immunity because the conformational B-cell epitopes activate and stimulate the B-cells when these cells encounter the epitopes. Therefore, the discontinuous B-cell epitopes of the V3 vaccine construct were predicted (Ponomarenko et al., 2008). Total 317 residues were found within five conformational B-cell epitopes with scores varying from 0.502 to 0.698.

The MD simulation study was performed for the docked TLR-8 and V3 complex to simulate a biological environment for the protein complex and analyze the physical movements

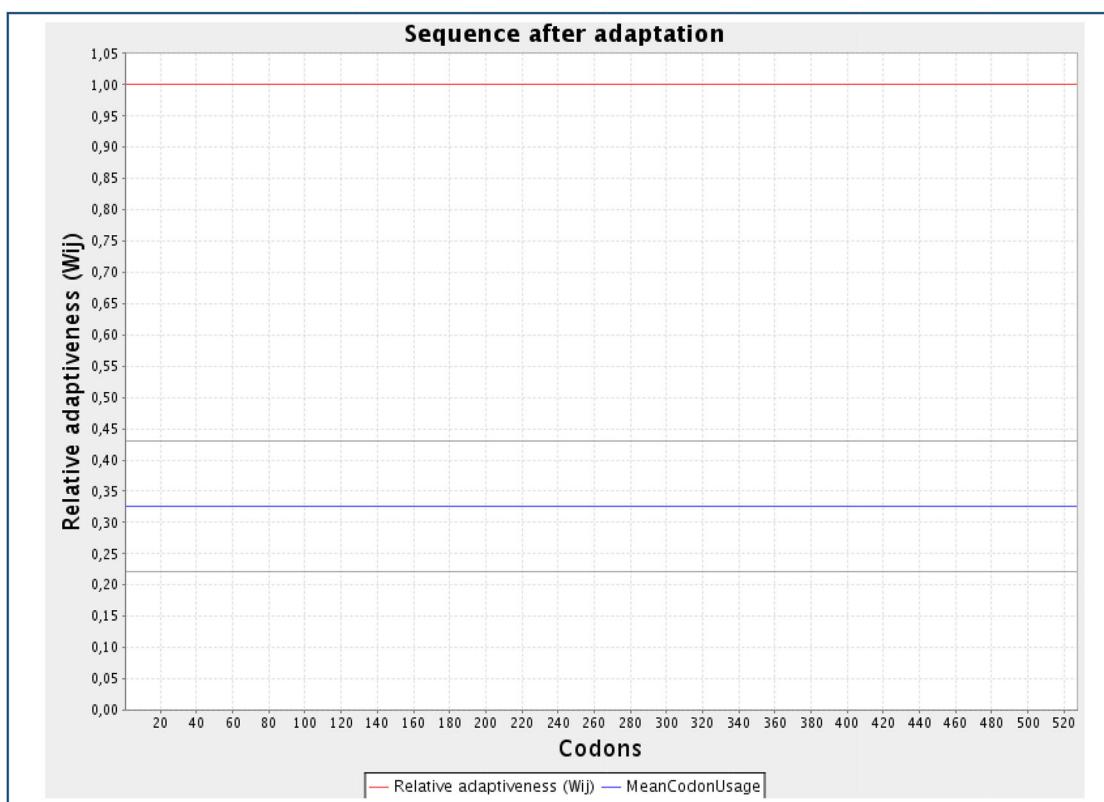


Figure 6. The results of the codon adaption experiment of V3 by the JCat server.

of the atom for a fixed length of time to observe the view of dynamic evolution of the system. The results of an MD simulation determine how stable the protein complex is in terms of changing pressure, temperature and motion. In our experiment, a low average potential energy of $-8.7690670e + 06$ kJ/mol and an average temperature fluctuation of only ± 1 K, was found from our target 300 K. These values indicated that the energy minimized structure was stable during the experiment. After simulating for 50 ns, an average RMSD of ~ 1.25 nm in both the energy minimized and crystal structure, was found. Again, the time series deviation of less than ~ 0.1 nm also indicated the good stability of the protein complex. After the MD simulation study, the immune simulation of the best vaccine construct V3 was performed to predict the real-life like immune interactions of V3 with different types of immune cells (Rapin et al., 2010). The immune simulation of the V3 construct showed that the response produced by V3 might be consistent with the typical immune response, where the primary immune response was found to be triggered after each of the vaccine injections, which produced the secondary immune response in the later stages. An increase in the concentrations of the memory B-cells, plasma B-cell, cytotoxic T-cells, helper T-cells, and different types of antibodies indicated that a very good humoral and cell-mediated immune responses might be built in the body after each of the injections. Furthermore, the elevated level of memory B-cell was also reported to last for several months after the three injections. The augmentation in the concentration of helper T-cells would also aid in the growth and proliferation of B-cells, thus improving the

adaptive immunity (Almofti et al., 2018; Carvalho et al., 2002). Furthermore, after each of the injections, the levels of macrophages and dendritic cells were also found to be increased, which pointed towards very good antigen presentation by these antigen-presenting cells (Hoque et al., 2019; Kambayashi & Laufer, 2014; Shey et al., 2019). The negligible Simpson index (D) suggested that a diverse immune response might be produced by the vaccine since it contained multiple B-cell and T-cell epitopes (Rapin et al., 2010). All these events indicated the production of very strong immune responses within the body by the V3 vaccine construct.

Finally, the codon adaptation and *in silico* cloning studies were conducted to identify the possible codons for expressing the vaccine V3 in *E. coli* strain K12. The *E. coli* is the recommended system for the mass production of recombinant proteins (Pei et al., 2005). The optimal limit of the CAI value should be close to 1.0 and the optimal range for GC content of an optimized DNA sequence was measured to be 30–70% (Khatoon et al., 2017; Morla et al., 2016; Shey et al., 2019). With the excellent predicted CAI value of 1.0, the DNA sequence might have very good amount of favorable codons that should be able to express the desired amino acids of V3 in *E. coli* strain K12. Again, the DNA sequence also had favorable amount of GC content of 52.8%. Thereafter, the recombinant pETite plasmid vector, containing the V3 vaccine insert was constructed. When the stability of the mRNA secondary structure of the vaccine protein was determined, both Mfold and RNAfold servers generated negative and much lower minimal free energies of -578.70 and

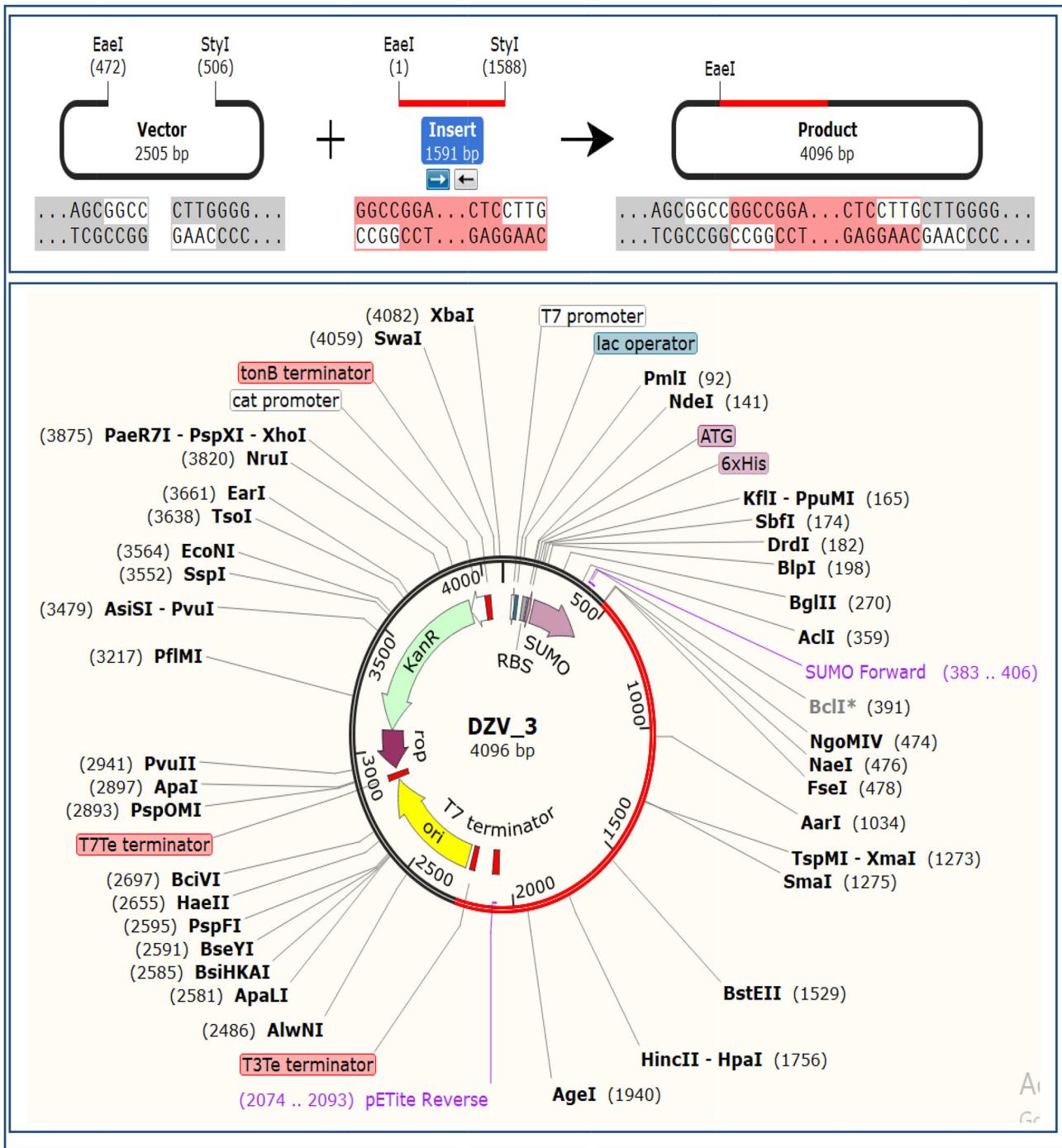


Figure 7. The results of the *in silico* cloning study of the V3 vaccine construct. The adapted DNA sequence of the V3 vaccine was inserted into the pETite plasmid.

–552.10 kcal/mol, respectively, which indicated that V3 vaccine protein might be quite stable upon transcription (Hamasaki-Katagiri et al., 2017).

The genome-based technologies for vaccine development will continue to dominate the field of vaccinology in the future because it provides the researchers a great opportunity to develop vaccines by facilitating the optimization of the target antigens. Conventional vaccines such as the attenuated vaccines or inactivated vaccines may fail to provide potential immunity towards a target antigen (Merten, 2002; Tameris et al., 2013). Therefore, the epitope-based subunit vaccines like the vaccines predicted in this study, could overcome such

difficulties. Although till now, no polyvalent subunit vaccine is available in the market, but monovalent epitope-based subunit vaccine like the RTS,S/AS01E vaccine (Mosquirix™) has gained market access recently. Exploiting the immunoinformatics approach during vaccine development, this vaccine is used as a preventative measure against malaria and it also yielded quite satisfactory results in many studies. Such vaccines will certainly pave the way of developing potential polyvalent epitope-based subunit vaccines to fight against various infectious diseases (Oyarzún & Kobe, 2016).

Overall, this study recommends V3 vaccine as the best vaccine construct based on the strategies employed in the

study to be an effective counter measure against the four DENV serotypes-1, 2, 3 and 4 as well as the ZIKV. However, results obtained from the *in silico* studies can only provide 'predictions' and need further wet lab validations before final use. Although the modern methods of bioinformatics can predict results with quite high accuracy, still *in vivo* and *in vitro* studies are required to finally confirm the outcomes of such *in silico* studies. However, our findings should definitely open new avenues to design multipathogenic and multivalent vaccines to combat against both the DENV and ZIKV.

5. Conclusion

In this study, epitope-based multivalent and multipathogenic vaccines were designed that might confer immunogenic protection against the four serotypes of DENV-1, 2, 3, and 4, and also the ZIKV at the same time. Since the vaccines contained multiple T-cell as well as B-cell epitopes from all of these five selected viruses, they might be able to provoke both humoral and cell-mediated immunogenic responses within the body. *In silico* validations also indicated that these multivalent and multipathogenic vaccines might be quite safe, effective, and responsive to use. Since all of these predictions were done based on the computational methods, wet lab validations are needed to finally confirm the outcomes of this study. If satisfied results are achieved, then these peptide-based vaccine candidates might become relatively cheap and effective options to reach the entire world to combat the infections of DENV and ZIKV.

Acknowledgments

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

Disclosure statement

The authors declare that they have no conflict of interest.

ORCID

Bishajit Sarkar  <http://orcid.org/0000-0001-8439-6994>
 Md. Asad Ullah  <http://orcid.org/0000-0002-6615-6433>
 Yusha Araf  <http://orcid.org/0000-0002-0144-5875>
 Mohammad Jakir Hosen  <http://orcid.org/0000-0001-6444-0214>

References

- Abbink, P., Stephenson, K. E., & Barouch, D. H. (2018). Zika virus vaccines. *Nature Reviews Microbiology*, 16(10), 594–600. <https://doi.org/10.1038/s41579-018-0039-7>
- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>
- Agumadu, V. C., & Ramphul, K. (2018). Zika virus: A review of literature. *Cureus*, 10(7), e3025. <https://doi.org/10.7759/cureus.3025>
- Almofti, Y. A., Abd-Elrahman, K. A., Gassmallah, S. A., & Salih, M. A. (2018). Multi epitopes vaccine prediction against severe acute respiratory syndrome (SARS) coronavirus using immunoinformatics approaches. *American Journal of Microbiological Research*, 6(3), 94–114. <https://doi.org/10.12691/ajmr-6-3-5>
- Angov, E. (2011). Codon usage: Nature's roadmap to expression and folding of proteins. *Biotechnology Journal*, 6(6), 650–659. <https://doi.org/10.1002/biot.201000332>
- Arai, R., Ueda, H., Kitayama, A., Kamiya, N., & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering*, 14(8), 529–532. <https://doi.org/10.1093/protein/14.8.522>
- Arpin, C., Dechanet, J., Van Kooten, C., Merville, P., Grouard, G., Briere, F., Banchereau, J., & Liu, Y. J. (1995). Generation of memory B cells and plasma cells in vitro. *Science*, 268(5211), 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. *International Journal of Peptide Research and Therapeutics*, 25(3), 861–872. <https://doi.org/10.1007/s10989-018-9735-4>
- Bacchetta, R., Gregori, S., & Roncarolo, M. G. (2005). CD4+ regulatory T cells: Mechanisms of induction and effector function. *Autoimmunity Reviews*, 4(8), 491–496. <https://doi.org/10.1016/j.autrev.2005.04.005>
- Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., Drake, J. M., Brownstein, J. S., Hoen, A. G., Sankoh, O., Myers, M. F., George, D. B., Jaenisch, T., Wint, G. R. W., Simmons, C. P., Scott, T. W., Farrar, J. J., & Hay, S. I. (2013). The global distribution and burden of dengue. *Nature*, 496(7446), 504–507. <https://doi.org/10.1038/nature12060>
- Biovia, D. S., Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., & Richmond, T. J. (2000). BIOVIA, discovery studio visualizer, v. 17.2, San Diego: Dassault systèmes, 2016. *The Journal of Chemical Physics*, 10, 0021–9991.
- Biswal, J. K., Bisht, P., Mohapatra, J. K., Ranjan, R., Sanyal, A., & Pattnaik, B. (2015). Application of a recombinant capsid polyprotein (P1) expressed in a prokaryotic system to detect antibodies against foot-and-mouth disease virus serotype O. *Journal of Virological Methods*, 215–216, 45–51. <https://doi.org/10.1016/j.jviromet.2015.02.008>
- Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein, J. S., Hoen, A. G., Moyes, C. L., Farlow, A. W., Scott, T. W., & Hay, S. I. (2012). Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Neglected Tropical Diseases*, 6(8), e1760. <https://doi.org/10.1371/journal.pntd.0001760>
- Buchan, D. W., & Jones, D. T. (2019). The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Research*, 47(W1), W402–7. <https://doi.org/10.1093/nar/gkz297>
- Bui, H. H., Sidney, J., Dinh, K., Southwood, S., Newman, M. J., & Sette, A. (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*, 7(1), 153. <https://doi.org/10.1186/1471-2105-7-153>
- Cano, R. L., & Lopera, H. D. (2013). Introduction to T and B lymphocytes. *Autoimmunity: from bench to bedside*. El Rosario University Press.
- Carvalho, L. H., Sano, G. I., Hafalla, J. C., Morrot, A., De Lafaille, M. A., & Zavala, F. (2002). IL-4-secreting CD4+ T cells are crucial to the development of CD8+ T-cell responses against malaria liver stages. *Nature Medicine*, 8(2), 166–170. <https://doi.org/10.1038/nm0202-166>
- Carrillo-Hernández, M. Y., Ruiz-Saenz, J., Villamizar, L. J., Gómez-Rangel, S. Y., & Martínez-Gutierrez, M. (2018). Co-circulation and simultaneous co-infection of dengue, chikungunya, and zika viruses in patients with febrile syndrome at the Colombian-Venezuelan border. *BMC Infectious Diseases*, 18(1), 61. <https://doi.org/10.1186/s12879-018-2976-1>
- Castiglione, F., Mantile, F., De Berardinis, P., & Prisco, A. (2012). How the interval between prime and boost injection affects the immune response in a computational model of the immune system. *Computational and Mathematical Methods in Medicine*, 2012, 1–9. <https://doi.org/10.1155/2012/842329>
- Chang, K. Y., & Yang, J. R. (2013). Analysis and prediction of highly effective antiviral peptides based on random forests. *PLoS One*, 8(8), e70166. <https://doi.org/10.1371/journal.pone.0070166>
- Chaudhri, G., Quah, B. J., Wang, Y., Tan, A. H., Zhou, J., Karupiah, G., & Parish, C. R. (2009). T cell receptor sharing by cytotoxic T lymphocytes

- facilitates efficient virus control. *Proceedings of the National Academy of Sciences of the United States of America*, 106(35), 14984–14989. <https://doi.org/10.1073/pnas.090655410>
- Chaves, B. A., Orfano, A. S., Nogueira, P. M., Rodrigues, N. B., Campolina, T. B., Nacif-Pimenta, R., Pires, A. C. A. M., Júnior, A. B. V., Paz, A. D. C., Vaz, E. B. D. C., Guerra, M. D. G. V. B., Silva, B. M., de Melo, F. F., Norris, D. E., de Lacerda, M. V. G., Pimenta, P. F. P., & Secundino, N. F. C. (2018). Coinfection with Zika virus (ZIKV) and dengue virus results in preferential ZIKV transmission by vector bite to vertebrate host. *The Journal of Infectious Diseases*, 218(4), 563–571. <https://doi.org/10.1093/infdis/jiy196>
- Chong, L. C., & Khan, A. M. (2019). Vaccine target discovery. *Encyclopedia of bioinformatics and computational biology* (pp. 241). Elsevier.
- Cohen, J. (2019). Controversy over dengue vaccine risk. *Science*, 365(6457), 961–962. <https://doi.org/10.1126/science.365.6457.961>
- Cooper, M. D. (2015). The early history of B cells. *Nature Reviews Immunology*, 15(3), 191–197. <https://doi.org/10.1038/nri3801>
- Cooper, N. R., & Nemerow, G. R. (1984). The role of antibody and complement in the control of viral infections. *Journal of Investigative Dermatology*, 83(1), S121–S7. <https://doi.org/10.1038/jid.1984.33>
- Craig, D. B., & Dombkowski, A. A. (2013). Disulfide by design 2.0: A web-based tool for disulfide engineering in proteins. *BMC Bioinformatics*, 14(1), 346. <https://doi.org/10.1186/1471-2105-14-346>
- da Silveira, L. T., Tura, B., & Santos, M. (2019). Systematic review of dengue vaccine efficacy. *BMC Infectious Diseases*, 19(1), 750. <https://doi.org/10.1186/s12879-019-4369-5>
- De Groot, A. S., Sbai, H., Aubin, C. S., McMurry, J., & Martin, W. (2002). Immuno-informatics: Mining genomes for vaccine components. *Immunology and Cell Biology*, 80(3), 255–269. <https://doi.org/10.1046/j.1440-1711.2002.01092.x>
- Dhandha, S. K., Vir, P., & Raghava, G. P. (2013). Designing of interferon-gamma inducing MHC class-II binders. *Biology Direct*, 8(1), 30. <https://doi.org/10.1186/1745-6150-8-30>
- Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v.2-a server for in silico prediction of allergens. *Journal of Molecular Modeling*, 20(6), 2278. <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(6), 846–851. <https://doi.org/10.1093/bioinformatics/btt619>
- Dowd, K. A., DeMaso, C. R., Pelc, R. S., Speer, S. D., Smith, A. R. Y., Goo, L., Platt, D. J., Mascola, J. R., Graham, B. S., Mulligan, M. J., Diamond, M. S., Ledgerwood, J. E., & Pierson, T. C. (2016). Broadly neutralizing activity of Zika virus-immune sera identifies a single viral serotype. *Cell Reports*, 16(6), 1485–1491. <https://doi.org/10.1016/j.celrep.2016.07.049>
- Doytchinova, I. A., & Flower, D. R. (2008). Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. *The Open Vaccine Journal*, 1, 4. <https://doi.org/10.2174/187503540080101002>
- Doytchinova, I. A., & Flower, D. R. (2007a). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8(1), 4. <https://doi.org/10.1186/1471-2105-8-4>
- Duhovny, D., Nussinov, R., & Wolfson, H. J. (2002). Efficient unbound docking of rigid molecules. In *International workshop on algorithms in bioinformatics* (pp. 185–200). Springer.
- Dupont-Rouze, M., O'Connor, O., Calvez, E., Daurès, M., John, M., Grangeon, J. P., & Gourinat, A. C. (2015). Co-infection with Zika and dengue viruses in 2 patients, New Caledonia, 2014. *Emerging Infectious Diseases*, 21(2), 381–382. <https://doi.org/10.3201/eid2102.141553>
- Fatima, K., & Syed, N. I. (2018). Dengvaxia controversy: Impact on vaccine hesitancy. *Journal of Global Health*, 8(2), 020312. <https://doi.org/10.7189/jogh.08.020312>
- 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. *Journal of Cheminformatics*, 9(1), 66. <https://doi.org/10.1186/s13321-017-0254-7>
- Fernandez, E., & Diamond, M. S. (2017). Vaccination strategies against Zika virus. *Current Opinion in Virology*, 23, 59–67. <https://doi.org/10.1016/j.coviro.2017.03.006>
- Funderburg, N., Lederman, M. M., Feng, Z., Drage, M. G., Jadlowsky, J., Harding, C. V., Weinberg, A., & Sieg, S. F. (2007). Human -defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18631–18635. <https://doi.org/10.1073/pnas.0702130104>
- Garcia, K. C., Teyton, L., & Wilson, I. A. (1999). Structural basis of T cell recognition. *Annual Review of Immunology*, 17(1), 369–397. <https://doi.org/10.1146/annurev.immunol.17.1.369>
- Garnier, J., Gibrat, J. F., & Robson, B. (1996). [32] GOR method for predicting protein secondary structure from amino acid sequence. *Methods in enzymology* (Vol. 266, pp. 540–553). Academic Press.
- Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook* (pp. 571–607). Humana press.
- Geourjon, C., & Deleage, G. (1995). SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences: Cabios*, 11(6), 681–684. <https://doi.org/10.1093/bioinformatics/11.6.681>
- Gill, S. C., & Von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry*, 182(2), 319–326. [https://doi.org/10.1016/0003-2697\(89\)90602-7](https://doi.org/10.1016/0003-2697(89)90602-7)
- 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 Research*, 33, W526–W531. <https://doi.org/https://doi.org/10.1093/nar/gki376>
- Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R., & Hofacker, I. L. (2008). The vienna RNA websuite. *Nucleic Acids Research*, 36, W70–W74. <https://doi.org/10.1093/nar/gkn188>
- Gu, Y., Sun, X., Li, B., Huang, J., Zhan, B., & Zhu, X. (2017). Vaccination with a paramyosin-based multi-epitope vaccine elicits significant protective immunity against *Trichinella spiralis* infection in mice. *Frontiers in Microbiology*, 8, 1475. <https://doi.org/10.3389/fmicb.2017.01475>
- 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(9), e73957. <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 Engineering*, 4(2), 155–161. <https://doi.org/10.1093/protein/4.2.155>
- Guzman, M. G., Halstead, S. B., Artsob, H., Buchy, P., Farrar, J., Gubler, D. J., Hunsperger, E., Kroeger, A., Margolis, H. S., Martínez, E., Nathan, M. B., Pelegrino, J. L., Simmons, C., Yoksan, S., & Peeling, R. W. (2010). Dengue: A continuing global threat. *Nature Reviews Microbiology*, 8(12), S7–S16. <https://doi.org/10.1038/nrmicro2460>
- Hamasaki-Katagiri, N., Lin, B. C., Simon, J., Hunt, R. C., Schiller, T., Russek-Cohen, E., Komar, A. A., Bar, H., & Kimchi-Sarfaty, C. (2017). The importance of mRNA structure in determining the pathogenicity of synonymous and non-synonymous mutations in haemophilia. *Haemophilia: The Official Journal of the World Federation of Hemophilia*, 23(1), e8–e17. <https://doi.org/10.1111/hae.13107>
- Hancock, R. E., Nijnik, A., & Philpott, D. J. (2012). Modulating immunity as a therapy for bacterial infections. *Nature Reviews Microbiology*, 10(4), 243–254. <https://doi.org/10.1038/nrmicro2745>
- Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R., & Warwick, J. (2017). Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics*, 33(19), 3098–3100. <https://doi.org/10.1093/bioinformatics/btx345>
- Hoque, M. N., Istiaq, A., Clement, R. A., Sultana, M., Crandall, K. A., Siddiki, A. Z., & Hossain, M. A. (2019). Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. *Scientific Reports*, 9(1), 1–4. <https://doi.org/10.1038/s41598-019-49468-4>
- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations.

- Journal of Chemical Information and Modeling*, 51(1), 69–82. <https://doi.org/10.1021/ci100275a>
- Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *The Journal of Biochemistry*, 88(6), 1895–1898. <https://doi.org/10.1093/oxfordjournals.jbchem.a133168>
- Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, 292(2), 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. *Nature Protocols*, 7(8), 1511–1522. <https://doi.org/10.1038/nprot.2012.085>
- Kambayashi, T., & Laufer, T. M. (2014). Atypical MHC class II-expressing antigen-presenting cells: Can anything replace a dendritic cell? *Nature Reviews Immunology*, 14(11), 719–730. <https://doi.org/10.1038/nri3754>
- Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., & Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*, 105(28), 6474–6487. <https://doi.org/10.1021/jp003919d>
- 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. *Scientific Reports*, 7(1), 1–2. <https://doi.org/10.1038/s41598-017-08842-w>
- Kindhauser, M. K., Allen, T., Frank, V., Santhana, R. S., & Dye, C. (2016). Zika: The origin and spread of a mosquito-borne virus. *Bulletin of the World Health Organization*, 94(9), 675–686C. <https://doi.org/10.2471/BLT.16.171082>
- Klein, D. E., Choi, J. L., & Harrison, S. C. (2013). Structure of a dengue virus envelope protein late-stage fusion intermediate. *Journal of Virology*, 87(4), 2287–2293. <https://doi.org/10.1128/JVI.02957-12>
- Ko, J., Park, H., Heo, L., & Seok, C. (2012). GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Research*, 40, W294–W297. <https://doi.org/10.1093/nar/gks493>
- Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nature Protocols*, 12(2), 255–278. <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. *Journal of Molecular Biology*, 157(1), 105–132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)
- Larsen, J. E., Lund, O., & Nielsen, M. (2006). Improved method for predicting linear B-cell epitopes. *Immunome Research*, 2(1), 2. <https://doi.org/10.1186/1745-7580-2-2>
- Laskowski, R. A., MacArthur, M. W., & Thornton, J. M. (2006). PROCHECK: Validation of protein-structure coordinates. *International tables for crystallography* (pp. 722–725). Wiley.
- 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., Jung, I. D., & Park, Y. M. (2014). A potential protein adjuvant derived from *Mycobacterium tuberculosis* Rv0652 enhances dendritic cells-based tumor immunotherapy. *PLoS One*, 9(8), e104351. <https://doi.org/10.1371/journal.pone.0104351>
- Levin, J. M., Robson, B., & Garnier, J. (1986). An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS Letters*, 205(2), 303–308. [https://doi.org/10.1016/0014-5793\(86\)80917-6](https://doi.org/10.1016/0014-5793(86)80917-6)
- Luckheeram, R. V., Zhou, R., Verma, A. D., & Xia, B. (2012). CD4+ T cells: Differentiation and functions. *Clinical and Developmental Immunology*, 2012, 1–12. <https://doi.org/10.1155/2012/925135>
- Ma, J., Wang, S., Zhao, F., & Xu, J. (2013). Protein threading using context-specific alignment potential. *Bioinformatics*, 29(13), i257–65. <https://doi.org/10.1093/bioinformatics/btt210>
- Magnan, C. N., Zeller, M., Kayala, M. A., Vigil, A., Randall, A., Felgner, P. L., & Baldi, P. (2010). High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics*, 26(23), 2936–2943. <https://doi.org/10.1093/bioinformatics/btq551>
- María, R. A., Arturo, C. V., Alicia, J. A., Paulina, M. L., & Gerardo, A. O. (2017). The impact of bioinformatics on vaccine design and development. InTech.
- Mathews, D. H., Sabina, J., Zuker, M., & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology*, 288(5), 911–940. <https://doi.org/10.1006/jmbi.1999.2700>
- Mathews, D. H., Turner, D. H., & Zuker, M. (2007). RNA secondary structure prediction. *Current Protocols in Nucleic Acid Chemistry*, 28(1), 11–12. <https://doi.org/10.1002/0471142700.nc1102s28>
- Merten, O. W. (2002). Virus contaminations of cell cultures - A biotechnological view. *Cytotechnology*, 39(2), 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(7), 646–653. <https://doi.org/10.1093/bioinformatics/17.7.646>
- Morla, S., Makrilia, A., & Kumar, S. (2016). Synonymous codon usage pattern in glycoprotein gene of rabies virus. *Gene*, 584(1), 1–6. <https://doi.org/10.1016/j.gene.2016.02.047>
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, 12(4), 345–364. <https://doi.org/10.1002/prot.340120407>
- Mustafa, M. S., Rasotgi, V., Jain, S., & Gupta, V. (2015). Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control. *Medical Journal, Armed Forces India*, 71(1), 67–70. <https://doi.org/10.1016/j.mjafi.2014.09.011>
- Nagpal, G., Usmani, S. S., Dhanda, S. K., Kaur, H., Singh, S., Sharma, M., & Raghava, G. P. (2017). Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Scientific Reports*, 7, 42851. <https://doi.org/10.1038/srep42851>
- Oyarzún, P., & Kobe, B. (2016). Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. *Human Vaccines & Immunotherapeutics*, 12(3), 763–767. <https://doi.org/10.1080/21645515.2015.1094595>
- 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 Science: A Publication of the Protein Society*, 4(11), 2411–2423. <https://doi.org/10.1002/pro.5560041120>
- Panda, S., & Chandra, G. (2012). Physicochemical characterization and functional analysis of some snake venom toxin proteins and related non-toxin proteins of other chordates. *Bioinformation*, 8(18), 891–896. <https://doi.org/10.6026/97320630008891>
- Pandey, R. K., Sundar, S., & Prajapati, V. K. (2016). Differential expression of miRNA regulates T cell differentiation and plasticity during visceral leishmaniasis infection. *Frontiers in Microbiology*, 7, 206. <https://doi.org/10.3389/fmicb.2016.00206>
- Pei, H., Liu, J., Cheng, Y., Sun, C., Wang, C., Lu, Y., Ding, J., Zhou, J., & Xiang, H. (2005). Expression of SARS-coronavirus nucleocapsid protein in *Escherichia coli* and *Lactococcus lactis* for serodiagnosis and mucosal vaccination. *Applied Microbiology and Biotechnology*, 68(2), 220–227. <https://doi.org/10.1007/s00253-004-1869-y>
- Peng, J., & Xu, J. (2011). RaptorX: Exploiting structure information for protein alignment by statistical inference. *Proteins: Structure, Function, and Bioinformatics*, 79(S10), 161–171. <https://doi.org/10.1002/prot.23175>
- Peters, B., & Sette, A. (2005). Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics*, 6(1), 132. <https://doi.org/10.1186/1471-2105-6-132>
- Petersen, M. T., Jonson, P. H., & Petersen, S. B. (1999). Amino acid neighbours and detailed conformational analysis of cysteines in proteins. *Protein Engineering*, 12(7), 535–548. <https://doi.org/10.1093/protein/12.7.535>
- Poland, G. A., Ovsyannikova, I. G., & Jacobson, R. M. (2009). Application of pharmacogenomics to vaccines. *Pharmacogenomics*, 10(5). <https://doi.org/10.2217/pgs.09.25>
- Ponomarenko, J., Bui, H. H., Li, W., Fusseder, N., Bourne, P. E., Sette, A., & Peters, B. (2008). ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*, 9(1), 514. <https://doi.org/10.1186/1471-2105-9-514>

- Purcell, A. W., McCluskey, J., & Rossjohn, J. (2007). More than one reason to rethink the use of peptides in vaccine design. *Nature Reviews Drug Discovery*, 6(5), 404–414. [10.1038/nrd2224](https://doi.org/10.1038/nrd2224) PMC[10.1038/nrd2224] [17473845]
- Rana, A., & Akhter, Y. (2016). A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. *Immunobiology*, 221(4), 544–557. <https://doi.org/10.1016/j.imbio.2015.12.004>
- Rapin, N., Lund, O., Bernaschi, M., & Castiglione, F. (2010). Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One*, 5(4), e9862. <https://doi.org/10.1371/journal.pone.0009862>
- Rappuoli, R. (2000). Reverse vaccinology. *Current Opinion in Microbiology*, 3(5), 445–450. [https://doi.org/10.1016/S1369-5274\(00\)00119-3](https://doi.org/10.1016/S1369-5274(00)00119-3)
- Romagnani, S. (1997). The th1/th2 paradigm. *Immunology Today*, 18(6), 263–266. [https://doi.org/10.1016/S0167-5699\(97\)80019-9](https://doi.org/10.1016/S0167-5699(97)80019-9)
- Saadi, M., Karkhah, A., & Nouri, H. R. (2017). Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 51, 227–234. <https://doi.org/10.1016/j.meegid.2017.04.009>
- Sarkar, B., Ullah, M. A., Johora, F. T., Taniya, M. A., & Araf, Y. (2020). Immunoinformatics-guided designing of epitope-based subunit vaccine against the SARS Coronavirus-2 (SARS-CoV-2). *Immunobiology*, 225(3), 151955. <https://doi.org/10.1016/j.imbio.2020.151955>
- Saha, S., & Raghava, G. P. (2006). AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Research*, 34, W202–W209. <https://doi.org/10.1093/nar/gkl343>
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*, 33, W363–7. <https://doi.org/10.1093/nar/gki481>
- Shey, R. A., Ghogomu, S. M., Esoh, K. K., Nebangwa, N. D., Shintouo, C. M., Nongley, N. F., Asa, B. F., Ngale, F. N., Vanhamme, L., & Souopgui, J. (2019). In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Scientific Reports*, 9(1), 1–8. <https://doi.org/10.1038/s41598-019-40833-x>
- Simmons, C. P., Farrar, J. J., van Vinh Chau, N., & Wills, B. (2012). Dengue. *The New England Journal of Medicine*, 366(15), 1423–1432. <https://doi.org/10.1056/NEJMra1110265>
- Solanki, V., & Tiwari, V. (2018). Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against *Acinetobacter baumannii*. *Scientific Reports*, 8(1), 1–9. <https://doi.org/10.1038/s41598-018-26689-7>
- Spreiter, Q., & Walter, M. (1999). Classical molecular dynamics simulation with the Velocity Verlet algorithm at strong external magnetic fields. *Journal of Computational Physics*, 152(1), 102–119. <https://doi.org/10.1006/jcph.1999.6237>
- Sturniolo, T., Bono, E., Ding, J., Raddrizzani, 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. *Nature Biotechnology*, 17(6), 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. *Physical Chemistry Chemical Physics*, 16(31), 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., & McShane, H. (2013). MVA85A 020 trial study team: Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: A randomised, placebo-controlled phase 2b trial. *The Lancet*, 381(9871), 1021–1028. [https://doi.org/10.1016/S0140-6736\(13\)60177-4](https://doi.org/10.1016/S0140-6736(13)60177-4)
- Tang, B., Huo, X., Xiao, Y., Ruan, S., & Wu, J. (2018). A conceptual model for optimizing vaccine coverage to reduce vector-borne infections in the presence of antibody-dependent enhancement. *Theoretical Biology & Medical Modelling*, 15(1), 13. <https://doi.org/10.1186/s12976-018-0085-x>
- Thomsen, M., Lundsgaard, C., Buus, S., Lund, O., & Nielsen, M. (2013). MHCcluster, a method for functional clustering of MHC molecules. *Immunogenetics*, 65(9), 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(2), 323–353. <https://doi.org/10.3390/vaccines2020323>
- Turner, P. J., & Xmgace, V. (2005). *Center for coastal and land-margin research*. Oregon Graduate Institute of Science and Technology.
- Ullah, M. A., Sarkar, B., & Islam, S. S. (2020). Exploiting the reverse vaccinology approach to design novel subunit vaccines against Ebola virus. *Immunobiology*, 225(3), 1:151949. <https://doi.org/10.1016/j.imbio.2020.151949>
- Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Kozakov, D. (2017). New additions to the ClusPro server motivated by CAPRI. *Proteins*, 85(3), 435–444. <https://doi.org/10.1002/prot.25219>
- Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., Wheeler, D. K., Sette, A., & Peters, B. (2019). The immune epitope database (IEDB): 2018 update. *Nucleic Acids Research*, 47(D1), D339–D343. <https://doi.org/10.1093/nar/gky1006>
- Wen, D., Foley, S. F., Hronowski, X. L., Gu, S., & Meier, W. (2013). Discovery and investigation of O-xylosylation in engineered proteins containing a (GGGGS) n linker. *Analytical Chemistry*, 85(9), 4805–4812. <https://doi.org/10.1021/ac400596g>
- 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 Research*, 47(W1), W322–W330. <https://doi.org/10.1093/nar/gkz397>
- World Health Organization. (2019). Zika epidemiology update, July. <https://www.who.int/emergencies/diseases/zika/epidemiology-update/en/>
- Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35, W407–W410. <https://doi.org/10.1093/nar/gkm290>
- 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. *Journal of Biomedical Science*, 17(1), 88. <https://doi.org/10.1186/1423-0127-17-88>
- Yuriev, E., & Ramsland, P. A. (2013). Latest developments in molecular docking: 2010–2011 in review. *Journal of Molecular Recognition*, 26(5), 215–239. <https://doi.org/10.1002/jmr.2266>
- Zhang, L. (2018). Multi-epitope vaccines: A promising strategy against tumors and viral infections. *Cellular & Molecular Immunology*, 15(2), 182–184. <https://doi.org/10.1038/cmi.2017.92>
- Zhu, J., & Paul, W. E. (2008). CD4 T cells: Fates, functions, and faults. *Blood*, 112(5), 1557–1569. <https://doi.org/10.1182/blood-2008-05-078154>
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31(13), 3406–3415. <https://doi.org/10.1093/nar/gkg595>