Received: 7 February 2018 | Accepted: 4 May 2018

DOI: 10.1002/jcb.27110

RESEARCH ARTICLE

WILEY Journal of Cellular Biochemistry

Designing B- and T-cell multi-epitope based subunit vaccine using immunoinformatics approach to control Zika virus infection

Rajan Kumar Pandey[1] | **Rupal Ojha**[1] | **Amit Mishra**[2] | **Vijay Kumar Prajapati**[1]

[1]Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, Ajmer, Rajasthan, India

[2]Cellular and Molecular Neurobiology Unit, Indian Institute of Technology Jodhpur, Jodhpur, India

Correspondence
Vijay Kumar Prajapati, Department of Biochemistry, Central University of Rajasthan, NH-8, Bandarsindri, Ajmer, Rajasthan 305817, India.
Email: vkprajapati@curaj.ac.in; Vijay84bhu@gmail.com

Abstract

The Zika virus is a rapidly spreading *Aedes* mosquito-borne sickness, which creates an unanticipated linkage birth deformity and neurological turmoil. This study represents the use of the combinatorial immunoinformatics approach to develop a multiepitope subunit vaccine using the structural and nonstructural proteins of the Zika virus. The designed subunit vaccine consists of cytotoxic T-lymphocyte and helper T-lymphocyte epitopes accompanied by suitable adjuvant and linkers. The presence of humoral immune response specific B-cell epitopes was also confirmed by B-cell epitope mapping among vaccine protein. Further, the vaccine protein was characterized for its allergenicity, antigenicity, and physiochemical parameters and found to be safe and immunogenic. Molecular docking and molecular dynamics studies of the vaccine protein with the toll-like receptor-3 were performed to ensure the binding affinity and stability of their complex. Finally, in silico cloning was performed for the effective expression of vaccine construct in the microbial system (*Escherichia coli* K12 strain). Aforementioned approaches result in the multiepitope subunit vaccine which may have the ability to induce cellular as well as humoral immune response. Moreover, this study needs the experimental validation to prove the immunogenic and protective behavior of the developed subunit vaccine.

KEYWORDS

adjuvant, epitope, subunit vaccine, vaccine, Zika virus (ZKV)

1 | INTRODUCTION

The Zika virus (ZKV) is an emerging mosquito-borne virus, which has shown rapid movement from obscurity to a Public Health Emergency of Global Concern due to a sudden increase in the number of reported cases of Guillain-Barré syndrome (an autoimmune disorder that has shown to causes acute or sub-acute flaccid paralysis) and congenital microcephaly.[1] The first major outbreak was reported from the Federated States of Micronesia with 185 confirmed cases. Afterward, the second major outbreak was reported from the French Polynesia in 2013 and 2014, estimated to involve 32 000 persons.[2] A recent serological survey has shown a broader geographical distribution of ZKV among the human population from India, The Philippines, Nigeria, Thailand, Malaysia, Egypt, Vietnam, and East Africa.[3]

ZKV transmission occurs via hitherto infected *Aedes* mosquitoes, but the transmission cycle differs among the different geographical areas. In African urban and

suburban locations, ZKV follows a human-to-mosquito-to-human transmission cycle, commonly comprising the *AE. aegypti* mosquitoes. Other Aedes mosquito species reported for ZKV transmission in Asia and Africa are *AE. luteocephalus*, *AE. taylori*, *AE. furcifer*, *AE. africanus*, and are likely enzootic vectors in Africa and Asia.[4,5] Among those infected with ZKV, 80% cases are asymptomatic, whereas only 20% of the clinically affected cases have been reported with mild symptoms of rash, arthralgia, fever, and conjunctivitis with no deaths reported.[6] The prevention measures used to control the ZKV infection are controlling vector, reducing sexual transmission, avoiding mosquito bites and avoiding unnecessary travel of pregnant women to ZKV transmission prone areas.

ZKV is an enveloped virus belonging to the genus *Flavivirus* and has a single-stranded, nonsegmented, positive-sense RNA genome of approximately 11 kb that can be directly translated into a long polyprotein to encode 3 structural proteins (capsid, membrane, and envelope protein) and 7 nonstructural proteins (NS1, 2K, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). All these proteins have a unique functional aspect that contributes toward the structural organization of ZKV. The structural proteins participate in virus particle formation, whereas the essential functions are performed by the nonstructural proteins.[7] Among these structural proteins, E-protein is a major component that participates in the virus to host receptor binding followed by membrane fusion and host immune recognition, with M protein positioned under the layer of E protein. The maturation process of ZKV revealed that in the immature stage, it consists of a partially ordered C protein shell that is absent in the matured particle but is rearranged during the phase of maturation.[8] On the other hand, nonstructural proteins play an essential role in the ZKV physiology. NS1 protein, in its glycosylated form, makes the association with the lipid, forms a homodimer, and gain significance in viral replication and infection. In the dimer form, NS1 resides inside the cell, but also secretes into the extracellular space by forming a hexameric lipoprotein particle, and is involved in immune invasion and pathogenesis by interacting with host immune components.[9]

Apart from the epidemiology of ZKV, there is still a lack of specific medicine and effective vaccine candidates to fight this epidemic. However, there are several precautions recommended to prevent ZKV infection including, drinking of plenty of water to prevent dehydration, proper bed rest, use of paracetamol or acetaminophen to reduce the fever and pain and so on. But, these precautions measures are not enough to prevent ZKV transmission; that's why there is an urgent need to search for effective vaccine candidates. To design an effective vaccine candidate against different ZKV strains, multiple antigenic epitopes should be selected followed by the addition of an appropriate adjuvant to enhance the host immune response. These characteristics of a vaccine can be fulfilled by using a subunit vaccine, a subclass of the vaccine family. Because of the recent headways in the technological and bioinformatics fields, multiepitope-based subunit vaccines have become a new approach to predict the possible epitope and elicit humoral (B-cell) and cell-mediated immunity (cytotoxic T-lymphocyte [CTL] and helper T-lymphocyte [HTL]).[10-13] This study also signifies the use of immunoinformatics approaches to design a multiepitope subunit vaccine with an ability to elicit specific B- and T-cell immune response against ZKV infection as reported elsewhere.[14] Structural proteins, along with NS1, were selected to predict the CTL and HTL epitopes followed by the subunit vaccine construction using a suitable adjuvant and linkers. The designed vaccine was further mapped for the presence of B-cell epitopes followed by the evaluation of its ability to elicit interferon gamma (IFN-γ) production. The nonallergenic and immunogenic properties of the vaccine candidate were also investigated using online tools. Further, secondary and tertiary structural analysis was performed to determine the tertiary structure of the designed subunit vaccine followed by refinement and model structure validation. Further, the binding affinity between the vaccine protein and toll-like receptor-3 (TLR-3) receptor was determined by molecular docking followed by molecular dynamics simulation. Moreover, codon optimization and in silico cloning were performed to investigate the futuristic polyprotein production.

2 | MATERIAL AND METHODOLOGY

2.1 | Peptide curation for multiepitope-based subunit vaccine production

A series of immunoinformatics approaches were applied to design an immunogenic multiepitope subunit vaccine against the ZKV infection. The amino acid sequences of 3 structural proteins (Capsid [GenBank: APH11492.1: 1-122], envelope [GenBank: APH11492.1:291-794], and membrane [GenBank: APH11492.1:216-290]) were obtained from the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/protein). These 3 proteins were included because of their significant role in ZKV infectivity. Capsid protein has shown a significant function in RNA binding and nucleocapsid formation.[15] Membrane protein promotes stabilization and also assists the folding and secretion of envelope protein, whereas envelope protein helps in the host receptor binding and membrane fusion for effective infectivity.[15] Along with the structural proteins, 1 nonstructural protein (NS1 [PDB ID-5K6K])

was also included because of its key role in viral RNA replication. Its amino acid sequence was obtained from the Protein Data Bank (PDB). All these curated protein sequences were further subjected to the various epitope predictions.

2.2 | MHC-II specific HTL epitope prediction

Major histocompatibility complex (MHC) molecules have vital roles in the adaptive immune response and defining the outcome of several host immune responses. MHC class II peptides have derived from the foreign protein taken up from the extracellular space. These peptides have shown to simulate both humoral and adaptive wings of the immune response against infectious diseases by utilizing the capability of CD4 + T-lymphocytes. Therefore, prediction of HTL epitopes is of quite an interest because of their involvement in massive immune responses. All 4 viral polyprotein sequences were subjected to the MHC-II epitope prediction module of Immune Epitope Database (http://tools.iedb.org/mhcii/)[12] to predict HTL epitopes. Immune Epitope Database MHC-II prediction module accepts polyprotein sequence in the FASTA format as an input to determine the ability of each subsequence to bind with specific MHC-II peptides. The selected prediction method was Consensus (SMM/NN).[16-18] HTL epitopes of 15-mer length were predicted for mouse alleles namely H2-IAd, H2-IAb, and H2-IEd.

2.3 | MHC-I specific CTL epitope prediction

The effector function of the adaptive immune system is performed by the CTLs that deal with malfunctioning or infected cells. CD8 + T cells and their surface receptors recognize the stable peptides presented by the MHC-I complex. Self-peptide does not elicit an immune response, whereas foreign antigens induce CTLs to produce a cytotoxic immune response. Hence, identification of CTLs peptide is of common interest for vaccine development and immunotherapeutic approaches against infectious diseases like Zika fever. All the shorted proteins of ZKV were subjected to the NetCTL-1.2 server (http://www.cbs.dtu.dk/services/NetCTL/)[10] to obtain the CTL epitopes. CTL epitope prediction by NetCTL 1.2 is not only restricted to 12 MHC class I supertype, but it also predicts the transporter associated with antigen processing (TAP)-transport efficiency and proteasomal C-terminal cleavage. It uses artificial neural networks for CTL epitope prediction and the same method is used to perform proteasomal cleavage. The TAP-transport efficiency was predicted using the weight matrix. During submission, the supertype was of an A1 type, whereas the weight on C-terminal cleavage and TAP-transport efficiency was 0.15 and 0.05, respectively. The threshold for the epitope prediction was kept at 0.75 to get significant sensitivity and specificity of 0.65 and 0.97, respectively.

2.4 | Multiepitope vaccine designing

The final subunit vaccine was designed by joining the resulting peptide sequences in a sequential manner using suitable linkers. Before joining these immunogenic epitopes, the presence of overlapping residues among the HTL, CTL and B-cell (BCL) epitopes was determined and only the epitopes with an overlapping region were used for designing the vaccine. Recently, mammalian β-defensin was reported to have a potential role as a mucosal adjuvant against HIV infection; therefore, because of its adjuvant properties against viral infection, it was selected and added to the N-terminal of CTL epitopes to construct the final vaccine.[19,20] Adjuvant and first CTL epitope were joined together by using the EAAAK linker, whereas intra-CTL epitopes were joined together by using the AAY linker.[12] Top 2 HTL-epitopes outputs for each input protein sequence were added next to the CTL epitope using the GPGPG linker.[12]

2.5 | B-cell epitope mapping

B-cell epitopes play important roles in humoral immunity by antibody production and memory cell formation for the future encounter of the same pathogen. Linear and conformational B-cell epitopes were mapped to the subunit vaccine using B-cell epitope prediction server (BCpred) (http://ailab.ist.psu.edu/bcpred/predict.html)[12] and ElliPro (http://tools.iedb.org/ellipro/)[13] server, respectively. Linear epitopes of 20-mer length were predicted at the default specificity of 75%. Whereas, conformational epitopes were predicted using the default parameters and outcome residues were visualized using the Chimera tool[21] to localize them in the final vaccine construct. ElliPro defines the discontinuous epitopes based on their protrusion index values and clusters them on the basis of distance R (Å). The larger R shows larger discontinuous epitopes being predicted.

2.6 | IFN-γ epitope prediction

The prediction module of the IFNepitope server (http://crdd.osdd.net/raghava/ifnepitope/)[22] was used to predict the interferon-gamma inducing HTL epitopes. By using this server, all HTL peptides were used as an input to evaluate the presence of the IFN-γ epitope. Both motif and SVM hybrid algorithms were used as prediction algorithms. The output result was obtained for each input epitope with its respective Support Vector Machine (SVM) score.

3 | VACCINE PROPERTIES EVALUATION

3.1 | Allergen prediction

To become a safe and effective vaccine candidate, the vaccine construct must be nonallergic in nature. The allergenic and nonallergenic behaviors of the vaccine construct were evaluated using 2 servers namely AllerTOP V2.0 (http://www.ddg-pharmfac.net/AllerTOP/)[12] and AlgPred (http://www.imtech.res.in/raghava/algpred/).[23] Of these, the former classifies the input protein sequence by using a k-nearest neighbor algorithm (kNN; k = 3) based on the training set consisting of 2210 known allergens from dissimilar species and 2210 nonallergens from the same species.[24] The latter integrates the SVM module to predict the allergenic nature of protein with high accuracy. The MEME/MAST allergen motif was searched by the server using MAST and the allergenic nature of protein was assigned if a matching motif was found.

3.2 | Antigenicity evaluation

The antigenicity evaluation of the final vaccine construct was performed using 2 freely accessible servers namely VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html)[25] and ANTIGENpro (http://scratch.proteomics.ics.uci.edu/).[12] The former classifies the antigen based solely on the physiochemical properties of the input protein sequence instead of the sequence alignment algorithm. The accuracy of the server is quite high and varies between 70% and 89% depending on the target organism. Whereas the latter predicts the whole protein antigenicity based on the results obtained by the protein microarray data analysis. It predicts the antigenicity independent of the pathogen but is a sequence-based approach.

3.3 | Physiochemical parameters evaluation

The final vaccine construct was subjected to the ProtParam server (http://web.expasy.org/protparam/)[26] to analyze its physiochemical properties. The parameters on which the polyprotein sequences were analyzed are theoretical pI, instability index, half-life, stability profiling, aliphatic index, and Grand Average of Hydropathy.

3.4 | Homology modeling (2D and 3D) and model validation

After the fusion of the CTL and HTL epitopes along with an immunogenic adjuvant and a suitable linker, the final vaccine construct was subjected to homology modeling using RaptorX (http://raptorx.uchicago.edu/StructurePrediction/predict/) structure prediction server.[27] RaptorX is an excellent server for predicting protein 3D structures without close homology (<30%) of the input sequence in the PDB. It uses a unique a nonlinear context-specific alignment potential and probabilistic consistency algorithm. The modeled 3D vaccine construct was further validated using online servers like RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php),[28] and ProSA (https://prosa.services.came.sbg.ac.at/prosa.php).[29] The secondary structure of the final vaccine construct was also predicted using PSIPRED.[30]

3.5 | Tertiary structure refinement

GalaxyRefine (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE)[31] is a freely accessible web server used for the improvement of template-based modeled protein. GalaxyRefine uses a side-chain rebuilding algorithm followed by packaging and overall structural relaxation using molecular dynamics simulation. It improves the global and local structural quality of template-based model proteins. The best 3D modeled structure of the final vaccine construct was selected and subjected to GalaxyRefine to refine the whole modeled protein using a mild and aggressive relaxation algorithm.

3.6 | Molecular docking to assess the interaction between vaccine construct and TLR3

Immune receptors present on the immune cells play a significant role in the induction of the correct immune response. To assess the tendency of the vaccine to induce an immunological response, blind docking was performed between the receptor [TLR3 (PDB ID: 2A0Z)] and the ligand (vaccine) using PatchDock (https://bioinfo3d.cs.tau.ac.il/PatchDock/).[32] PatchDock is a geometry-based molecular docking algorithm to obtain docking transformations and yields good molecular shape complementarity.

3.7 | Molecular dynamics simulation

Molecular dynamics simulation is an in silico approach to determine the microscopic interaction in the protein-ligand complex as well as complex stability. The docked complex of vaccine protein and TLR3 were further subjected to a molecular dynamics simulation study using Gromacs 5.0 to determine the stability of the complex.[33] First, crystal water was removed followed by topology generation using the GROMOS96 43a1 force field for the protein-ligand complex. Then the protein was centered in

a cubic box and the distance between the protein and box edge was at least 1.0 nm. The simple point charge water model was selected to add water molecules to the boundary box followed by addition of ions using a genion tool to neutralize the charged protein complex.[34] Further, the solvated, electroneutral system was relaxed through energy minimization to ensure that the system had a lack of steric clashes and inappropriate geometry. Then, canonical ensemble (NVT–amount of substance [N], volume [V] and temperature [T] are conserved) equilibration was performed for 100-ps followed by isothermal-isobaric ensemble (NPT–amount of substance [N], pressure [P] and temperature [T] are conserved) for the same time duration. Finally, molecular dynamics simulation was performed for a duration of 10 ns.

3.8 | In silico cloning of final vaccine construct

With the end goal of cloning and expression of the planned multiepitope-based subunit vaccine in an expression vector, the protein sequence of final vaccine construct was reverse transcribed to get an optimized coding sequence using Codon Usage Wrangler and Gene Synthesis Wizard (http://www.mrc-lmb.cam.ac.uk/ms/methods/codon.html).[12] *Escherichia coli* was selected as a host organism for the expression of final vaccine construct, whereas because of lack of virus as an option for the origin of the organism, another option was selected. The respective codon adaptive index (CAI), codon usage distribution and the percentage guanine-cytosine (GC) content were also obtained for the same input sequence. CAI gives a thought regarding the impact of characteristic choice on codon use inclination. To express a gene of interest in an organism, perfect CAI esteem ought to be 1.00; however, an esteem more prominent than 0.8 is likewise considered as great. But the percentage of GC content ought to lie in the range of 30%-70%. The occurrence of values outside of this range show antagonistic impacts on the transcriptional and translational efficiencies. Next, the optimized codon sequence of vaccine protein was reversed to match the direction of gene expression of the vector. Finally, restriction cloning was performed using SnapGene v3.3.4. XhoI and BamHI restriction sites were added to the N and C terminals of the optimized complementary DNA sequences followed by their insertion within the pET-28a(+) vector to ensure the polyprotein synthesis within the *E coli* expression system.

4 | RESULTS AND DISCUSSION

4.1 | Polyprotein sequence procurement

The sum of 4 polyprotein sequences was retrieved from the National Center for Biotechnology Information protein server and PDB. Among these 4 polyproteins, 3 proteins [capsid protein (accession no APH11492), envelope protein (accession No. APH11492), and membrane protein (accession No. APH11492)] were structural in nature, whereas the fourth protein [NS1 (PDB ID: 5K6K) was nonstructural. The sequence length of the capsid, membrane polyprotein, and envelope was found to be 122 amino acid, 74 amino acid, and 503 amino acid residues, respectively, whereas NS1 was found to be 376 amino acids residues long. These 4 proteins were selected because of their important role in the viral attachment to the host followed by replication and survival.[11]

4.2 | Helper T-lymphocytes epitope prediction

The HTL epitope prediction output for 4 polyproteins resulted in the 15-mer unique peptide epitopes. Top 10 peptide epitopes for each polyprotein were selected (Table 1). Among the selected epitopes, the epitopes having overlapping sequences were merged together and

TABLE 1 Predicted HTL epitopes against capsid, membrane, envelope, and NS1 protein of ZIKA virus

S. No.	Zika virus protein	Position	Allele	Percentile rank	SMM IC_{50}	HTL epitopes
1	Capsid protein	39-58	H2-IAd	0.29	75	LGHGPIRMVLAILAFLRFTA
		54-68	H2-IAd	1.55	1020	LAFLRFTAIKPSLGLINR
2	Membrane protein	32-56	H2-IAb	2.17	257	VENWIFRNPGFALAAAAIAWLLGSS
3	Envelope protein	48-64	H2-IAd	2.46	3073	TTVSNMAEVRSYCYEAS
		337-355	H2-IAd	2.31	2743	GPCKVPAQMAVDMQTLTPV
		415-430	H2-IAd	2.42	2876	VRGAKRMAVLGDTAWD
4	Nonstructural-1 (NS1)	54-78	H2-IAd	5.84	471	DRYKYHPDSPRRLAAAVKQAWEEGI
		335-350	H2-IAd	6.22	3793	WCCRECTMPPLSFRAK

HTL, helper T-lymphocyte; SMM, stabilization matrix alignment method.

got a total of 2, 1, 3, 2 peptides of variable length for capsid, membrane, envelope, and NS1, respectively. The epitopes were selected based on their lowest percentile rank because lower scored percentile rank demonstrates high affinity. The percentile ranks of predicted epitopes ranging from 0.12 to 5.8 shows high binding ability to the HTL receptor.[11,35] A peptide with an half maximal inhibitory concentration (IC_{50}) value of lower than 50 nM is considered to have high affinity, while less than 500 nm and 5000 nM for intermediate and lower affinity epitopes, respectively. The IC_{50} values obtained for the predicted epitopes ranged from 39.62 to 872 nM. These data show that the predicted HTL epitopes have higher to moderate to lower affinity for the MHC-II peptide.[10]

4.3 | CTL epitope prediction

The CTL epitopes for 4 ZIKV proteins were obtained by using the MHC-I prediction module of the NetCTL server resulting in CTL peptide epitopes of 9mer length as an output along with MHC-I binding affinity and prediction score. For capsid, membrane polyprotein, envelope, and NS1; a total of 1, 4, 10, 4 epitopes were obtained, respectively (Table 2). Overlapping sequences among these epitopes were also obtained, which were merged together to get a single peptide. The epitope prediction score was greater than the threshold (0.75), where higher the value of prediction score, stronger will be the binding of CTL epitopes to the CTL receptor and that leads to good cellular immune response. Therefore, predicted CTL epitopes have a good binding affinity to the CTL receptor.[10,11]

4.4 | Design and construction of final multiepitope vaccineThe

CTL and HTL epitopes obtained from the 4 ZIKV polyprotein were joined together to make the final vaccine construct. The final vaccine constructs consist of 456 amino acid residues and 3 domains, which include an adjuvant, CTL epitopes, and HTL epitopes.[11] β-defensin with a sequence of GIINTLQKYYCRVR GGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK was selected as an adjuvant and added to the N-terminal sequence of the vaccine construct.[19] Seventeen CTL epitopes were obtained for the capsid, membrane protein, envelope, and NS1 protein sequences. These 17 CTL epitopes were added next to the adjuvant. Furthermore, 8 MHC class II epitopes were added next to the CTL epitopes. These 3 domains were joined together by using proper linkers (EAAAK, AAY, and GPGPG) as shown in Figure 1. Here, the helical linker (EAAAK) was utilized to join the adjuvant and capsid toward the N-terminal site of the vaccine construct to reduce the interaction with another protein region along with efficient separation.[36,37] The AAY motif was

TABLE 2 Predicted CTL epitopes against capsid, membrane, envelope, and NS1 protein of ZIKA virus

S. No.	Zika virus protein	Position	CTL epitopes	Prediction score	MHC binding affinity
1	Capsid protein	105-113	GADTSVGIV	0.7708	0.1654
2	Membrane protein	17-25	QTWLESREY	1.1356	0.1974
		55-63	SSTSQKVIYL	1.3	0.268
		66-74	MILLIAPAY	0.7576	0.1069
3	Envelope protein	53-61	MAEVRSYCY	2.1654	0.4403
		85-93	QSDTQYVCK	0.9922	0.2076
		129-137	SIQPENLEY	2.1621	0.4389
		159-167	ETDENRAKV	0.872	0.1722
		195-103	GLDFSDLYY	3.0511	0.6604
		198-206	FSDLYYLTM	2.6297	0.5872
		205-213	TMNNKHWLV	1.0169	0.2026
		308-316	CTAAFTFTK	0.9226	0.1876
		324-332	GTVTVEVQY	1.1278	0.1984
		368-374	STENSKMML	1.1357	0.2251
4	Nonstructural-1 (NS1)	148-156	DVEAWRDRY	1.4467	0.2745
		216-224	EAAHSDLGY	1.4651	0.2906
		219-227	HSDLGYWIE	1.1677	0.2926
		293-301	HSEELEIRF	1.1379	0.2072

CTL, cytotoxic T-lymphocyte; MHC, major histocompatibility complex.

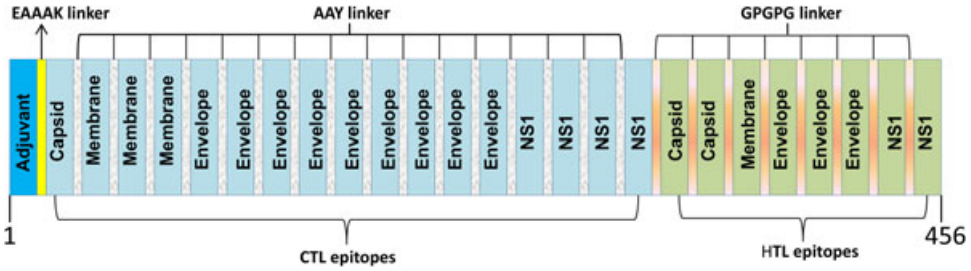

FIGURE 1 Schematic representation of multiepitope vaccine construct. A multiepitope vaccine sequence consisting of 456 amino acids residues has been constructed where adjuvant was linked at N-terminal and fused with the multiepitope sequences with the help of EAAAK linker (yellow) followed by CTL (ice blue) and HTL (green) epitopes have been joined together by AAY linkers and GPGPG linkers, respectively. CTL, cytotoxic T-lymphocyte; HTL, helper T-lymphocyte

used to join the CTL epitope as a linker to enhance the epitope separation. After the separation, the C-terminal of the CTL epitope was suitable for TAP transported binding or other chaperon molecules, which led to efficient epitope presentation,[38] whereas GPGPG linkers simulated the HTL immune response along with the conserved conformational dependent immunogenicity of helper T cells as well as antibody epitopes.[39]

4.5 | B-cell epitope mapping

B-cell epitopes have shown their identity as an antigenic determinant that is recognized and bound by a B-lymphocyte surface receptor. Because of an important role in vaccination and immunodiagnostic purposes, the final vaccine construct was subjected to BCPRED and ElliPro for linear and conformational B-cell epitopes prediction, respectively.[11,40] Linear B-cell epitope was predicted using the amino acid sequences of the vaccine construct, whereas conformational epitopes were predicted using the refined 3D model of the vaccine construct. Among the predicted B-cell epitopes, top epitopes with a score of >0.8 were selected showing good affinity for the B-cell receptor (Supporting Information Table 1 and Figure 2A). While 43 conformational B-cell epitopes were identified with a score of 0.858 (Supporting Information Table 2 and Figure 2B). Here the ElliPro score represents the protrusion index value average for the epitope residues.

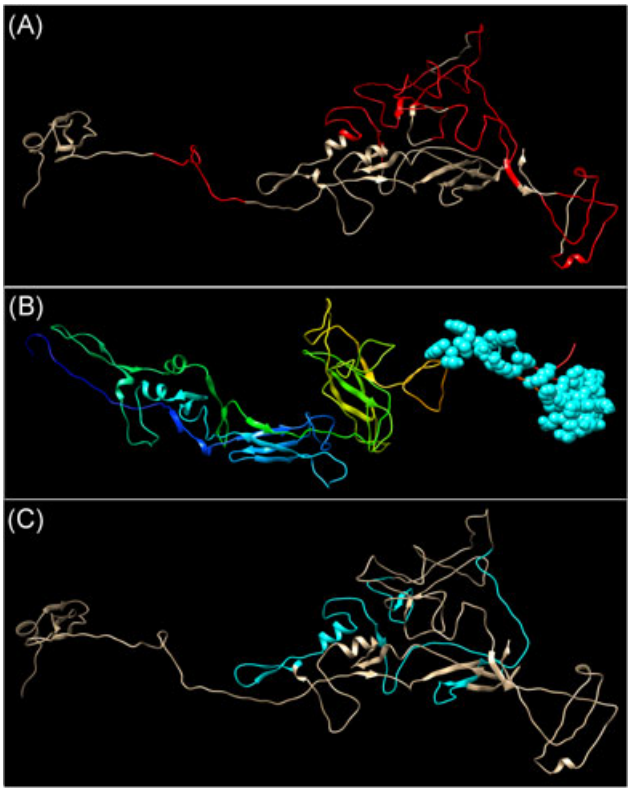

FIGURE 2 Diagrammatic representation of predicted B-cell and IFN-γ epitopes where (A) represents the modeled 3D structure of the final vaccine construct (Golden) showing linear B cell epitopes highlighted in the red color. (B) Conformational epitopes predicted for the final vaccine construct have been shown in cyan color with a ball-like appearance. (C) IFN-γ inducing epitopes have been shown in cyan color in the final multiepitope vaccine construct. IFN-γ, interferon gamma

4.6 | IFN-γ inducing epitopes prediction

With a specific end goal to outline the subunit vaccine, the MHC class II binder epitopes with the ability to induce cell-mediated immunity were identified using the IFNepitope server (http://crdd.osdd.net/raghava/ifnepitope/).[22,41] The antigenic regions that were identified to bind to MHC class II and activate the T-helper cells may lead to the release of IFN-γ and activate the downstream signaling.[42] All epitopes were found to have IFN-γ inducing ability (Figure 2C).

5 | SUBUNIT VACCINE PROPERTIES EVALUATION

5.1 | Allergenicity and immunogenicity assessment

The designed subunit vaccine was evaluated on the allergenic parameter using AllerTOP 2.0 server and AlgPred server. Both these servers have shown the nonallergenic nature of the final vaccine construct. The score obtained for the AlgPred output was −0.615 at the threshold value of −0.4.

The antigenicity associated with the subunit vaccine was predicted using VaxiJen v2.0 and ANTIGENpro servers. As per the result of vaxiJen, the vaccine antigenicity was 0.6468% at the threshold value of 0.4%, showing it as a probable antigen; whereas the result output of ANTIGENpro is similar to VaxiJen output with a score of 0.664 showing antigenic nature of vaccine construct. Therefore, the obtained result from both these servers showed high probability for the antigenic nature of the subunit vaccine.

5.2 | Physiochemical parametric evaluation

By evaluating the final vaccine constructs among the nine physiochemical parameters of ProtParam server,[43] we found that the molecular weight of our final vaccine construct was 49.6 kDa. For an ideal vaccine candidate, the molecular weight should be greater than approximately 40 to 50 kDa, which will lead to an increased uptake by the lymphatic system.[44] Therefore, the obtained molecular weight for the final vaccine construct lies in an ideal range, whereas the theoretical protrusion index was found to be 8.53, showing the basic nature of the vaccine construct. It's estimated in vitro half-life was found to be 30 hours in mammalian reticulocytes, indicating the presence of glycine at the amino terminus and on the experimental model.[45] This estimated half-life basically shows the time taken by the protein to remain half of the amount as initially synthesizing within the cell. Another parameter, instability index, was also estimated to check the stability of the protein in a test tube.[46] The instability index value of less than forty represents the stable nature of the protein and vice versa. We found that the final vaccine construct had the instability index value of 35.81, which shows their stable nature. Aliphatic index[47] was also analyzed, showing the relative volume occupied by the aliphatic side chain. It might be considered as a positive variable for the expansion of thermostability of the globular proteins. The obtained values of the aliphatic index were 72.52; showing their thermostable nature at variable temperature. The value of Grand Average of Hydropathy[48] represents the sum of the hydropathy value, showing the hydrophilic and hydrophobic nature of protein along its amino acid sequence. The obtained value of Grand Average of Hydropathy for the vaccine protein was −0.177, showing the hydrophilic nature.

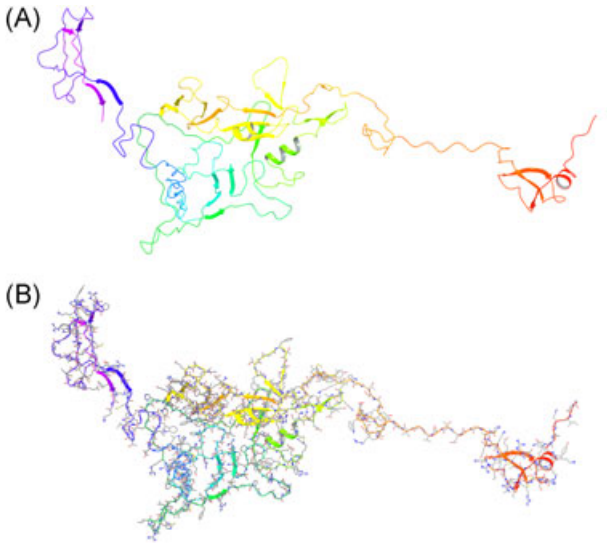

FIGURE 3 Final multiepitope vaccine model and validation. (A,B) Represents the 3D model obtained for the multi-subunit vaccine construct obtained using RaptorX

5.3 | Molecular modeling (2D and 3D) and structure validation

The secondary structure of the final vaccine construct was obtained using PSIPRED showing 16.23% of helix, 12.28% β-strand and 71.49% of coiled structure (Supporting Information Figure 1). Later on, the final vaccine construct was modeled using the RaptorX structure prediction server, resulting in the 3-dimensional structure of the input sequence.[49] The template used to model the vaccine protein was dengue virus serotype 1 soluble fragment-E (SE) (PDB ID: 4GT0). On the basis of this template, 100% residues were modeled as 3 domains. The overall unnormalized global distance test (uGDT) was 158 (34), showing the number of identical residues in the alignment. In RaptorX, the *P* value represents the relative quality of the predicted model; a smaller *P* value shows higher quality of the model. The *P* value obtained for the predicted vaccine model was 2.12e−05, showing the good quality of the model (Figure 3A,B). The Ramachandran plot assessment for vaccine models showed 89% residues in the favored region, while 7.5% in the allowed region. But after refinement using GalaxyRefine, an increase in the number of residues in the favored region was observed. After refinement, 94.1% residues were present in the favored

FIGURE 4 Docked complex of TLR-3 (PDB ID: 2A0Z) with the multiepitope-based subunit vaccine construct. Receptor (TLR-3) has been depicted in blue color, whereas magenta represents the multiepitope subunit vaccine construct as a ligand in the docked complex obtained by RaptorX molecular docking. PDB, protein data bank; TLR-3, toll-like receptor-3

region, 4.8% residues in the allowed region, and only 1.1% residues were found to be in the disallowed region (Supporting Information Figure 2A). The quality assessment of modeled vaccine constructs using PROSA shows a Z-score of − 1.27 (Supporting Information Figure 2B).

5.4 | Molecular docking between the vaccine and TLR-3 receptor

To predict the binding affinity of the vaccine construct and the TLR3 receptor, molecular docking was performed using PatchDock server. Top 10 docked structures were obtained in the PDB format with the clustering root-mean-square deviation (RMSD) of 4.0. The output model was visualized using University of California, San Francisco (UCSF-Chimera)[21] and the most plausible model was selected (Figure 4). The geometric shape complementarity score for the best plausible model was 21832, whereas the complex interface area was 3295. The atomic contact energy complex was 26.3. The best-docked model is shown in Figure 4.

5.5 | Molecular dynamics simulation study

To obtain a stable protein model and improved 3D configuration, the final vaccine protein was docked with TLR3 without any irregularities, followed by energy minimization.[50] Energy minimization relieved unphysical changes but it corrects the molecular geometry of structure without altering the overall structure. The protein vaccine was energy minimized with steps < 50 000 Fmax having energy with constant temperature 300 K and pressure 1 bar run for 100 ps. To confirm that the energy minimization was successfully performed, the potential energy of the complex was evaluated and it was found that the Epot was − 15.344e + 06. Energy minimization results in a structure with reasonable geometry and solvent orientation.[51,52] System equilibration under the canonical ensemble was performed and the temperature of the system was plotted (Supporting Information Figure 3A), and it was found that the temperature of the system reaches up to the 299.78 K very quickly within 2 ps, remaining stable thereafter for remaining time period. Further, NPT ensemble was performed to equilibrate the pressure of the system and it was found that the pressure of the system increases very quickly and remains stable at an average pressure of 0.47 bar (Supporting Information Figure 3B). Molecular dynamics simulation of vaccine protein and TLR3 complex for the time period of 10 ns results in the RMSD of 9.2 nm, whereas the RMSF was approximately 0.25 nm up to 10 000 atoms and then a sudden increase was observed up to 1.0 nm for atom 10-12.5 K followed by RMSF decrease up to 0.25 nm (Figure 5C,D).

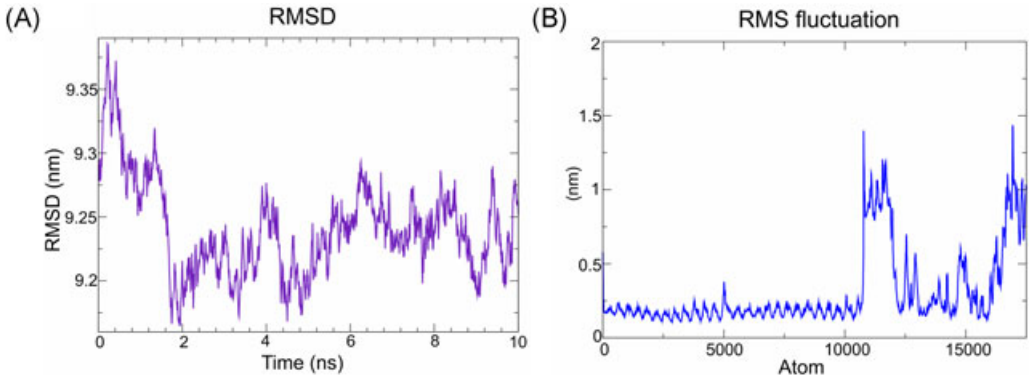

FIGURE 5 Molecular dynamics simulation and its parametric evaluation for the docked complex (TLR-3 and vaccine). (A) RMSD obtained for the backbone of docked complex consisting of TLR-3 as a receptor and multiepitope vaccine as a ligand (B) RMSF plot obtained for the side chain residues of docked complex protein showing the stable nature of the complex

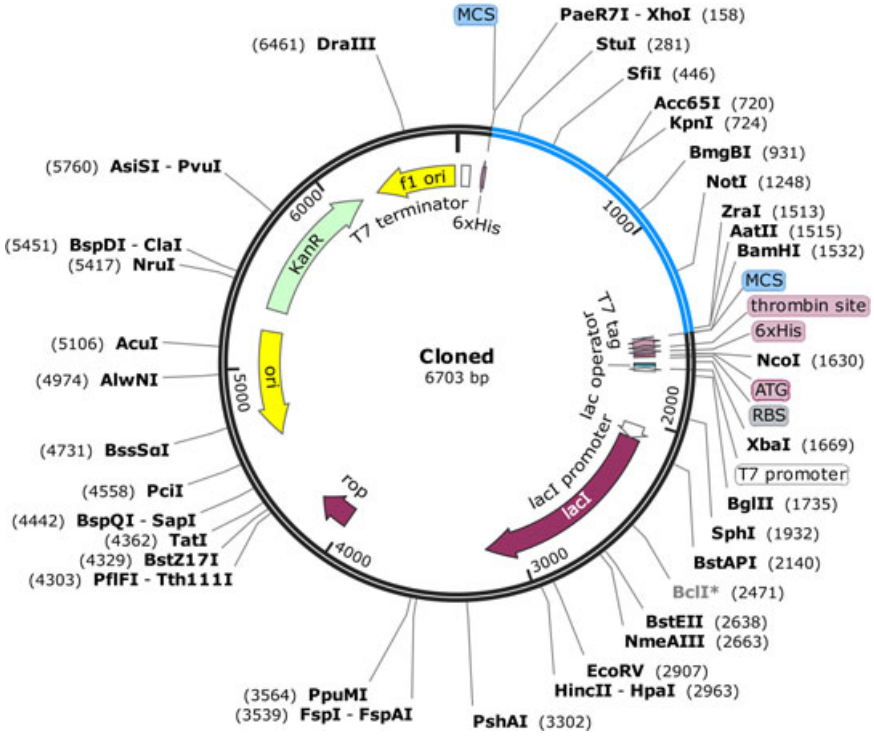

FIGURE 6 In silico cloning to express final vaccine construct representing pET-28a(+) vector (black semicircle) and optimized codon of final vaccine construct (blue semicircle)

5.6 | In silico cloning

The results obtained from the reverse translation and codon optimization of the final vaccine construct have shown the CAI value of 1 for our optimized nucleotide sequence representing high-level protein expression in *E coli*. The ideal CAI value should be 1, but it may vary from 0.8-1.0. The CAI values of less than 0.8 shows the poor expression of the target gene. The ideal percentage range of the GC content should be in between 30% and 70%. Any regions lying outside of this range will unfavorably impact expression proficiency.[53] The percentage of GC content obtained for our vaccine construct was 62.01%, which is quite good, showing good expression efficiency in *E coli*. Finally, to clone the vaccine construct in *E coli* pET-28a(+) vectors, XhoI and BamHI restriction sites were introduced at the N- and C-terminals of the optimized codon, respectively. Later on, the optimized codons with restriction sites were inserted into the vector by restriction cloning. Finally, a cloned construct was generated having a sequence length of 6703 base pairs (Figure 6).

6 | CONCLUSION

ZKV has shown rapid movement from the obscurity to a Public Health Emergency of Global Concern. In spite of such a huge global emergence of ZKV infection, we still don't have chemotherapeutic or vaccine options. The idea behind the scanning of the viral genome to find immunogenic epitopes leads to an elicited immune response without any reversal of viral pathogenesis. The recent advancements in the technological and bioinformatics fields allow us to apply this algorithm using computer-based approaches. Therefore, this study is a step ahead in the path of effective vaccine search against ZKV infection. We used the proteomics sequence data of ZKV polyprotein and subjected it to various freely accessible online servers in a sequential manner to obtain the HTL, CTL, BCL, and IFN-γ inducing multiepitopes to make the final vaccine construct. Further, physiochemical characterization of the vaccine construct was performed to predict their allergenic behavior, antigenicity, stability and molecular weight. Further, homology modeling was performed to obtain the 3D structure of the final vaccine construct. The binding affinity and complex stability of the vaccine construct and TLR3 were checked by performing molecular docking and dynamics simulation. At last, in silico cloning was performed for the expression of the final vaccine construct. This study uses a large number of immunoinformatic approaches to find the vaccine candidate to fight against ZKV infection.

ACKNOWLEDGMENTS

RK Pandey is thankful to the Department of Science and Technology for providing INSPIRE fellowship. VK Prajapati is thankful to the Central University of Rajasthan for providing computational facility.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

AUTHOR CONTRIBUTIONS

Protocol was designed by RKP, AM, and VKP. Methodology was performed by RKP, RO, and VKP. Manuscript was written by RKP, AM, and VKP.

ORCID

Amit Mishra http://orcid.org/0000-0001-9401-4400
Vijay Kumar Prajapati http://orcid.org/0000-0001-6510-0596

REFERENCES

1. Wilder-Smith A, Gubler DJ, Weaver SC, Monath TP, Heymann DL, Scott TW. Epidemic arboviral diseases: priorities for research and public health. *Lancet Infect Dis*. 2016;17:e101-e106.
2. Cao-Lormeau VM, Roche C, Teissier A, et al. Zika virus, French Polynesia, South Pacific, 2013. *Emerg Infect Dis*. 2014;20:1085-1086.
3. Petersen LR, Jamieson DJ, Powers AM, Honein MA. Zika Virus. *N Engl J Med*. 2016b;374:1552-1563.
4. Diallo D, Sall AA, Diagne CT, et al. Zika virus emergence in mosquitoes in southeastern Senegal, 2011. *PLoS One*. 2014;9:e109442.
5. Garcia R, Marchette NJ, Rudnick A. Isolation of Zika virus from *Aedes aegypti* mosquitoes in Malaysia. *Am J Trop Med Hyg*. 1969;18:411-415.
6. Petersen E, Wilson ME, Touch S, et al. Rapid spread of Zika Virus in The Americas—implications for public health preparedness for mass gatherings at the 2016 Brazil olympic games. *Int J Infect Dis*. 2016a;44:11-15.
7. Shi Y, Gao GF. Structural biology of the Zika Virus. *Trends Biochem Sci*. 2017;42:443-456.
8. Prasad VM, Miller AS, Klose T, et al. Structure of the immature Zika virus at 9 A resolution. *Nat Struct Mol Biol*. 2017;24:184-186.
9. Avirutnan P, Fuchs A, Hauhart RE, et al. Antagonism of the complement component C4 by flavivirus nonstructural protein NS1. *J Exp Med*. 2010;207:793-806.
10. Khatoon N, Pandey RK, Prajapati VK. Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep*. 2017;7:8285.
11. Li G, Poulsen M, Fenyvuesvolgyi C, et al. Characterization of cytopathic factors through genome-wide analysis of the Zika viral proteins in fission yeast. *Proc Natl Acad Sci U S A*. 2017;114:E376-E385.
12. Pandey RK, Bhatt TK, Prajapati VK. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for Malaria by investigating anopheles salivary protein. *Sci Rep*. 2018;8:1125.
13. Pandey RK, Prajapati VK. Exploring sand fly salivary proteins to design multi-epitope subunit vaccine to fight against visceral leishmaniasis. *J Cell Biochem*. 2018. https://doi.org/10.1002/jcb.26719
14. Usman Mirza M, Rafique S, Ali A, et al. Towards peptide vaccines against Zika virus: immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. *Sci Rep*. 2016;6:37313.
15. Wong SSY, Poon RWS, Wong SCY. Zika virus infection-the next wave after dengue?. *J Formos Med Assoc*. 2016;115:226-242.
16. Nielsen M, Lund O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics*. 2009;10:296.
17. Nielsen M, Lundegaard C, Lund O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics*. 2007;8:238.
18. Wang P, Sidney J, Kim Y, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics*. 2010;11:568.
19. Mohan T, Mitra D, Rao DN. Nasal delivery of PLG microparticle encapsulated defensin peptides adjuvanted gp41 antigen confers strong and long-lasting immunoprotective response against HIV-1. *Immunol Res*. 2014;58:139-153.
20. Mohan T, Sharma C, Bhat AA, Rao DN. Modulation of HIV peptide antigen specific cellular immune response by synthetic alpha- and beta-defensin peptides. *Vaccine*. 2013;31:1707-1716.
21. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25:1605-1612.
22. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct*. 2013;8:30.
23. Saha S, Raghava GPS. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res*. 2006;34:W202-W209.
24. Pandey RK, Sundar S, Prajapati VK. Differential expression of miRNA regulates T Cell differentiation and plasticity during visceral leishmaniasis infection. *Front Microbiol*. 2016b;7:206.
25. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8:4.
26. Wilkins MR, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*. 1999;112:531-552.
27. Källberg M, Wang H, Wang S, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc*. 2012;7:1511-1522.
28. Lovell SC, Davis IW, Arendall WB, 3rd, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*. 2003;50:437-450.
29. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35:W407-W410.
30. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16:404-405.
31. Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*. 2013;41:W384-W388.
32. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*. 2005;33:W363-W367.

33. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4: 435-447.
34. Shukla R, Chetri PB, Sonkar A, Pakharukova MY, Mordvinov VA, Tripathi T. Structure-based screening and molecular dynamics simulations offer novel natural compounds as potential inhibitors of Mycobacterium tuberculosis isocitrate lyase. *J Biomol Struct Dyn*. 2017b:1-13. https://doi.org/10.1080/07391102.2017.1341337
35. Ahluwalia PK, Pandey RK, Sehajpal PK, Prajapati VK. Perturbed microRNA expression by mycobacterium tuberculosis promotes macrophage polarization leading to pro-survival foam cell. *Front Immunol*. 2017;8:107.
36. Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng*. 2001;14:529-532.
37. Pandey RK, Narula A, Naskar M, et al. Exploring dual inhibitory role of febrifugine analogues against Plasmodium utilizing structure-based virtual screening and molecular dynamic simulation. *J Biomol Struct Dyn*. 2017c;35:791-804.
38. Bergmann CC, Yao Q, Ho C-K, Buckwold SL. Flanking residues alter antigenicity and immunogenicity of multi-unit CTL epitopes. *J Immunol*. 1996;157:3242-3249.
39. Livingston B, Crimi C, Newman M, et al. A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes. *J Immunol*. 2002;168:5499-5506.
40. Pandey RK, Prajapati P, Goyal S, Grover A, Prajapati VK. Molecular modeling and virtual screening approach to discover potential antileishmanial inhibitors against ornithine decarboxylase. *Comb Chem High Throughput Screen*. 2016a;19:813-823.
41. Pandey RK, Kumbhar BV, Sundar S, Kunwar A, Prajapati VK. Structure-based virtual screening, molecular docking, ADMET and molecular simulations to develop benzoxaborole analogs as potential inhibitor against Leishmania donovani trypanothione reductase. *J Recept Signal Transduct Res*. 2017b;37:60-70.
42. Dong C, Flavell RA. Cell fate decision: T-helper 1 and 2 subsets in immune responses. *Arthritis Res*. 2000;2:179-188.
43. Gasteiger E, Hoogland C, Gattiker A, Se Duvaud, Wilkins MR, Appel RD, Bairoch A, et al. *Protein Identification and Analysis Tools on the ExPASy Server*. Berlin, Germany: Springer; 2005.
44. Liu H, Irvine DJ. Guiding principles in the design of molecular bioconjugates for vaccine applications. *Bioconjug Chem*. 2015; 26:791-801.
45. Bachmair A, Finley D, Varshavsky A. In vivo half-life of a protein is a function of its amino-terminal residue. *Science*. 1986;234:179-186.
46. Guruprasad K, Reddy BVB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng*. 1990;4:155-161.
47. Ikai A. Thermostability and aliphatic index of globular proteins. *J Biochem*. 1980;88:1895-1898.
48. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157:105-132.
49. Pandey RK, Verma P, Sharma D, Bhatt TK, Sundar S, Prajapati VK. High-throughput virtual screening and quantum mechanics approach to develop imipramine analogues as leads against trypanothione reductase of leishmania. *Biomed Pharmacother*. 2016c;83:141-152.
50. Pandey RK, Kumbhar BV, Srivastava S, et al. Febrifugine analogues as Leishmania donovani trypanothione reductase inhibitors: binding energy analysis assisted by molecular docking, ADMET and molecular dynamics simulation. *J Biomol Struct Dyn*. 2017a;35:141-158.
51. Shukla R, Shukla H, Kalita P, et al. Identification of potential inhibitors of *Fasciola gigantica* thioredoxin1: computational screening, molecular dynamics simulation, and binding free energy studies. *J Biomol Struct Dyn*. 2017a;1-16:2147-2162. https://doi.org/10.1080/07391102.2017.1344141
52. Chander S, Pandey RK, Penta A, et al. Molecular docking and molecular dynamics simulation based approach to explore the dual inhibitor against HIV-1 reverse transcriptase and Integrase. *Comb Chem High Throughput Screen*. 2017;20:734-746.
53. Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci Rep*. 2017;7:9232.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Pandey RK, Ojha R, Mishra A, Prajapati VK. Designing B- and T-cell multiepitope-based subunit vaccine using immunoinformatics approach to control Zika virus infection. *J Cell Biochem*. 2018;119:7631-7642. https://doi.org/10.1002/jcb.27110