Infection, Genetics and Evolution 61 (2018) 4–15

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Excavating chikungunya genome to design B and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection

Aruna Narula[a,1], Rajan Kumar Pandey[a,1], Nazia Khatoon[a], Amit Mishra[b], Vijay Kumar Prajapati[a,*]

[a] Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, Bandarsindri, Ajmer, Rajasthan 305817, India
[b] Cellular and Molecular Neurobiology Unit, Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan 342011, India

ARTICLE INFO

Keywords:
Immunoinformatics
Chikungunya virus
Subunit vaccine
T-cell epitope
B-cell epitope

ABSTRACT

Chikungunya infection has been a cause of countless deaths worldwide. Due to lack of permanent treatment and prevention of this disease, the mortality rate remains very high. Therefore, we followed an immunoinformatics approach for the development of multi-epitope subunit vaccine which is able to elucidate humoral, cell-mediated and innate immune responses inside the host body. Both structural and non-structural proteins of chikungunya virus were utilized for prediction of B-cell and T-cell binding epitopes along with interferon-γ (IFN-γ) inducing epitopes. The vaccine construct is composed of β-defensin as an adjuvant at the N-terminal followed by Cytotoxic T-Lymphocytes (CTL) and Helper T-Lymphocyte (HTL) epitopes. The same vaccine construct was also utilized for the prediction of B-cell binding epitopes and IFN-γ inducing epitopes. This was followed by the 3D model generation, refinement and validation of the vaccine construct. Later on, the interaction of modeled vaccine with the innate immune receptor (TLR-3) was explored by performing molecular docking and molecular dynamics simulation studies. Also to check the efficiency of expression of this vaccine construct in an expression vector, *in silico* cloning was performed at the final stage of vaccine development. Further, designed multi-epitope subunit vaccine necessitates experimental and clinical investigation to develop as an immunogenic vaccine candidate.

1. Introduction

Chikungunya is one of the mosquito-borne viral diseases which is characterized by the sudden onset of frequently occurring fever along with joint pain, muscle pain, headache, nausea, fatigue, and rashes. The first outbreak of this disease occurred in Southern Tanzania in the year 1952 (Khan et al., 2015). According to the data released by National Vector Borne Disease Control Programme (NVBDCP), 18,639 cases of chikungunya were reported globally in the year 2013 (Cecilia, 2014). Earlier, chikungunya virus outbreaks occurred majorly in Africa, Asia, Europe, and the Indian and Pacific Oceans but in late 2013, the first local transmission of chikungunya virus was reported in America in the Caribbean countries. It was followed by the spread of infection in 45 countries or territories throughout the Pan American region with > 1.7 million cases that have been reported by Pan American Health Organisation (PAHO) (Brathwaite Dick et al., 2012). Lately, in the year 2016, 31,000 chikungunya cases were reported to PAHO. Chikungunya virus (CHIKV) is an RNA virus belonging to the genus *Alphavirus* of family Togaviridae. There are two main species responsible for the transmission of chikungunya infection and its resulting epidemic namely *Aedes aegypti* and *Ae. Albopictus*; out of them *Ae. aegypti* is prominent in tropical and sub-tropical regions while *Ae. albopictus* occurrence has been reported in temperate as well as cold temperate regions. Despite such a huge proportion of the global population being affected by an epidemic that is caused by the infection of chikungunya virus, no specific treatment or commercial vaccine exists against this disease. Patients suffering from chikungunya are being given the only primary treatment to relieve symptoms. However, a group of researchers has discovered a virus-like particle vaccine against epidemic chikungunya virus for the protection of non-human primates (Akahata et al., 2010).

CHIKV can be localized in connective tissues such as epimysium, muscle, joints, skin fibroblast and central nervous system (CNS) but not in foetal or placental tissues of human body. Target organs at the early stage of CHIKV infection includes lymphoid tissues, liver, CNS, joints and muscles whereas the persistence of CHIKV at the later stage of infection can be observed in lymphoid organs, joints, liver, muscles and macrophages out of which macrophages act as the main site for the

* Corresponding author at: Department of Biochemistry, Central University of Rajasthan, NH-8, Bandarsindri, Ajmer, Rajasthan Pin- 305817, India.
E-mail address: vkprajapati@curaj.ac.in (V.K. Prajapati).
[1] Authors have equally contributed.

https://doi.org/10.1016/j.meegid.2018.03.007
Received 24 October 2017; Received in revised form 15 February 2018; Accepted 8 March 2018
Available online 10 March 2018
1567-1348/

persistence of CHIKV infection. CHIKV infection can be diagnosed by the characterization of an early viremia which can be observed at the onset of fever and it can increase up to 10^9 to 10^{12} RNA copies/mL. Also, this viremia can increase up to 12 days (Simon et al., 2011).

Chikungunya virus is having 11,805 nucleotides long genome (Solignat et al., 2009) encoding structural and non-structural proteins. Structural proteins include capsid protein, E3, E2, E1, P62, and 6K. Capsid protein is having protease activity and therefore gets auto-cleaved from the structural precursor protein after which it associates with ribosome for a short span of time followed by its binding to the viral RNA to form an icosahedral nucleocapsid. The resultant nucleocapsid then interacts with the cytoplasmic domain of membrane protein E2 thus leading to the formation of a mature virion particle. E2 is required for the attachment of the virus to the membrane receptors of the targeted cell while the function of E3 remains unexplored. Another structural protein 6K which is expressed on the viral membrane is involved in various functions such as viral glycoprotein processing, cell permeabilization, budding of viral particles and disruption of cellular calcium homeostasis in the virus. The fusion of viral particles with that of the host cell is assisted by E1 which is a class II viral fusion protein. Non-structural proteins of chikungunya virus include NSP1, NSP2, NSP3 and NSP4 out of which NSP1, NSP2, and NSP3 are cleaved from a transient non-structural polyprotein P123. The capping function of viral RNA is performed by NSP1 while NSP2 is having two functional domains (a) the N-terminal domain which is involved in the RNA polymerase complex, possess both RNA triphosphatase and RNA helicase activity and (b) the C-terminal domain which is involved in specific proteolytic cleavage of all four mature nonstructural proteins from the parental polyprotein. Inhibition of cellular transcription is reflected by degradation of PLOR2A which forms the catalytic subunit of RNA polymerase II complex (Akhrymuk et al., 2012). NSP3 is required for the biosynthesis of minus strand and subgenomic 26S mRNA and NSP4 serves as an RNA-dependent RNA polymerase.

Vaccines not only prevent the onset of a particular disease but also provide a gateway for its specific eradication along with the avoidance of toxicity associated with drug-based therapy. Traditional vaccines have been proved to be a most effective way to prevent the invasion of diseases inside the human body. These vaccines include administration of heat-killed or live attenuated virus/bacteria which reduce the safety regarding the utilization of such vaccines. In subunit vaccines, the safety profile can be enhanced as they do not consist of the whole micro-organism but contain only specific epitope regions of antigens present in them (Chander et al., 2017; Khatoon et al., 2017a). The immunogenicity of these antigens can be enhanced by the addition of adjuvants. Subunit vaccines allow the customization of immune responses by means of optimized delivery systems, adjustment of vaccine size, and generation of adaptive immune responses such as humoral, cell-mediated immune responses along with enhanced efficacy of vaccine (Moyle and Toth, 2013). Immunoinformatics is a bioinformatics-based approach which includes mapping of B-cell, T-cell, and IFN-γ epitopes. It is less time consuming as well as cost effective as compared to other classical laboratory-based methods for designing of vaccines. Various points should be kept in mind while designing a vaccine. More than one antigen should be selected and it should be capable of developing virulence in various stages of the disease development process. These antigens should be able to generate adaptive immune responses. In this study, we tried to develop a multi-epitope vaccine which is able to generate humoral, cell mediated and innate immune responses as it contains B-cell, T-cell and IFN-γ inducing epitopes. We retrieved the protein sequence of all structural and non- structural proteins of chikungunya virus in FASTA format from UniProt database. These sequences were utilized for the immunoinformatics analysis including Cytotoxic T-Lymphocytes (CTL), Helper T-Lymphocyte (HTL) epitopes and B-cell binding epitope predictions. Only those CTL and HTL binding epitopes were taken into consideration for designing the vaccine which was sharing common regions with B-cell binding epitopes. These epitopes were joined together by appropriate linkers to enhance the epitope presentation and efficient separation of epitopes by proteasomal degradation inside human body. Also an adjuvant (β-defensin) (Mohan et al., 2013) was added at the N-terminal of the above sequence of epitopes and linkers to enhance the immunogenicity of the vaccine. Epitopes selected for vaccine designing were subjected to the epitope conservancy analysis, which confirms its broader population coverage as shown by its identity with three different CHIV strains. The final sequence of vaccine was then further used to predict B-cell and IFN-γ binding epitopes to see the common regions within vaccine sequence which can induce both innate and adaptive immune response. Vaccine properties such as allergenicity, antigenicity, and physicochemical properties were also predicted. Secondary and tertiary structures were built using various servers and the final tertiary structure was refined to fill missing residues and to perform mild and aggressive relaxation and this refined structure was validated by Ramachandran plot. This structure was subjected to docking with TLR-3 to explore the binding interactions and the docked receptor ligand complex was analyzed for stable microscopic interactions by molecular dynamics simulation. TLR-3 was taken as the immune receptor of choice because its expression increases on immune cells with the entry of CHIKV inside host (Thangamani et al., 2010). Also to check the translation efficiency of vaccine model in an expression vector, *in silico* cloning was performed (Figure-1).

2. Methodology

2.1. Collection of chikungunya viral protein sequences for vaccine construction

In the first phase of the study, the genome of chikungunya virus (CHIKV) was extracted from literature in order to find out proteins that are being expressed in the virus (Solignat et al., 2009). Using UniProt (www.uniprot.org/) database, amino acid sequences of total nine proteins including both structural (UniProt ID- Q8JUX5) and nonstructural (UniProt ID- Q8JUX6) viral proteins of strain *S27-African prototype* were obtained for further analysis. The strain S27-African prototype has shown 99.9% similarity with the Indian strain of chikungunya, that's the reason we selected this strain for this study (Arankalle et al., 2007). Structural proteins of CHIKV include capsid protein, E1, E2, E3 and 6K whereas nonstructural proteins include NSP1, NSP2, NSP3, and NSP4. At the same time, TLR-3 agonist β-defensin sequence was also retrieved from the literature in FASTA format.

In the second phase of this study, we predicted HTL, CTL and B-cell epitopes by immunoinformatics analysis. Different parameters were used to select final epitopes for the vaccine construction. Later on, all the epitopes along with the TLR-3 agonist (adjuvant) were fused together along with suitable linker to generate the final vaccine sequence. At the end of this study, we also checked the efficiency of the final vaccine by performing molecular docking, molecular dynamics simulation.

2.2. CTL epitope prediction

CTL epitopes of all structural and non-structural proteins of CHIKV were predicted using NetCTL 1.2 server (http://www.cbs.dtu.dk/services/NetCTL/) (Ahluwalia et al., 2017; Ali et al., 2017). The prediction is based on three main components, prediction of the MHC-I binding peptide, proteasomal C-terminal cleavage, and TAP (Transporter Associated with Antigen Processing) transport efficiency. Predicted CTL epitopes were restricted to 12 MHC-I supertypes. MHC-I binding and proteasomal C-terminal cleavage are achieved by utilization of artificial neural networks whereas prediction of TAP transporter efficiency is done by the weight matrix. Scores of all three prediction methods are finally integrated and thresholds for this integrated score can be converted into sensitivity or specificity values. The threshold

value for epitope identification was set as 0.75 during prediction of CTL epitopes (Larsen et al., 2007).

2.3. HTL epitope prediction

T helper cell 15-mer epitopes for mouse MHC II alleles (IAb, IAd, IAs, IEb, IEd and IEs) were predicted for both structural and non-structural protein of CHIKV using Immune Epitope Database (IEDB) (http://www.iedb.org/) (Khatoon et al. 2017; Pandey et al., 2016a). The screening of epitopes predicted by IEDB is based on affinity for their receptor that can be inferred from the IC_{50} value assigned to each epitope (Duvvuri et al., 2014). Peptides with higher affinity should have IC_{50} values < 50 nM. The IC_{50} value < 500 nM indicates an intermediate affinity while a value < 5000 nM can be directly related to the low affinity of epitopes (Pandey et al., 2017a; Rana and Akhter, 2016). IEDB also assigns a percentile rank to each predicted epitope which is based on the comparison of peptide's score with five million random 15-mer taken from SWISS-PROT database. The value of percentile rank can be inversely related to the affinity of each epitope (Vita et al., 2015).

2.4. B-cell epitope prediction

Linear B-cell epitopes for all structural and non-structural proteins of CHIKV were predicted using BCPREDS server (http://ailab.ist.psu.edu/bcpred/) (Khatoon et al., 2017b; Pandey et al., 2017c). This server utilizes three prediction methods: (i) AAP method (Chen et al., 2007), (ii) BCPred (El-Manzalawy et al., 2008a, b) and (iii) FBCPred (El-Manzalawy et al., 2008a, b). The BCPred method of epitope prediction is based on subsequence kernel-based SVM classifier. The method predicts linear epitopes having < 80% sequence identity. The performance of BCPred is higher (AUC 0.758) as compared to AAP method (AUC 0.7) (Potocnakova et al., 2016). Epitopes that are generated can be screened on the basis of the scoring function assigned to them by BCPREDS. The cut-off value of score > 0.8 was chosen to select the peptides having maximum epitope-like properties (Barh et al., 2010).

2.5. Construction of multi-epitope vaccine sequence

From the above immunoinformatics analysis, a vaccine sequence was constructed including overlapping sequences of HTL and CTL epitopes with linear B-cell epitopes. They were further linked together by AAY and GPGPG linkers respectively (Hajighahramani et al., 2017). To make the final vaccine, at the N-terminal of above sequence, β-defensin (Ali et al., 2017; Mohan et al., 2013; Pandey et al., 2017b) was added as an adjuvant with EAAAK linker (Hajighahramani et al., 2017). Linkers were included for efficient separation of epitopes and to elevate the epitope presentation inside the host cell (Nezafat et al., 2014). β-defensin is responsible for elucidating innate as well as adaptive immune responses by inducing recruitment of naïve T-cells and immature dendritic cells through their interaction with the immune receptor such as TLRs and CCR6 (Chemokine Receptor 6) receptor at the site where infection has occurred (Mohan et al., 2013).

2.6. Conservation across antigen analysis

The designing of an epitope-based subunit vaccine necessitates the use of conserved epitopes to provide the wide-ranging protection against the multiple strains of pathogens. As in past, many chikungunya virus strains have been reported, it was necessary for us to determine the conservancy of vaccine epitopes to confirm their identity among the available chikungunya strains. The conservancy of CTL and HTL epitopes were determined among three different chikungunya strains namely Senegal37997, DRDE-06 and Wuerzburg-1, by using IEDB conservation across antigen analysis tool (Bui et al., 2007).

2.7. IFN-γ inducing and linear B-cell binding epitope prediction from final vaccine construct

IFNepitope server (http://crdd.osdd.net/raghava/ifnepitope/) was utilized for the prediction of IFN-γ inducing epitopes from the final vaccine construct including both HTL and CTL epitopes (Khatoon et al., 2017b; Pandey et al., 2016b). This web server has been developed on the basis of a dataset which is comprised of both IFN-γ inducing and non-inducing MHC-II binder (Hajighahramani et al., 2017). Linear B-cell binding epitopes were also predicted from final vaccine construct sequence by using BCPREDS.

2.8. Allergenicity and antigenicity prediction

The allergic or non-allergic nature of vaccine construct was predicted by the AlgPred server (http://www.imtech.res.in/raghava/algpred/). This server predicts allergens on the basis of the similarity of the already known epitope with any region within our protein of interest. In order to locate epitope in our protein of interest, the server possesses IgE epitope mapping feature. The server is also specialized to search MEME/MAST allergen motifs which help in discriminating between allergens and non-allergens. Blast search on allergen representative peptides (ARPs) and SVM module which is based on amino acid/dipeptide composition are also one of the allergenicity prediction approaches offered by this server (Saha and Raghava, 2006).

The antigenicity of final multi-epitope vaccine sequence was determined by ANTIGENpro (http://www.scratch.proteomics.ics.uci.edu/) (Pandey and Prajapati, 2018) which requires protein sequence as input. This server is not dependent upon the pathogen identity and alignment. The prediction pipeline follows a two-step approach including multiple representations of the primary input sequence and five algorithms. The SVM classifier is responsible for briefing the resultant predictions that finally predict whether a protein is antigenic or not (Magnan and Baldi, 2014). VaxiJen v2.0 (http://www.ddg-pharma.net/vaxijen/VaxiJen.html/) was also used for prediction of antigenic nature of vaccine. It predicts antigens independently of sequence alignment and is based on Auto Cross-Covariance (ACC) transformation of input amino acid sequence of a protein into uniform vectors of physiochemical properties (Doytchinova and Flower, 2007).

2.9. Physicochemical properties prediction

Physicochemical properties including molecular weight, pI, half-life, instability index, aliphatic index and GRAVY (Grand Average Hydropathy) were calculated for the final vaccine amino acid sequence by ExPASy-ProtParam tool (http://web.expasy.org/protparam/). Half-life of a protein can be described as the amount of time taken for its disappearance after it has been synthesized in the cell. Instability index was calculated to know the estimate of stability of proteins in a test tube. The aliphatic index gives an idea of the relative volume occupied by aliphatic side chains and GRAVY is the average of hydropathy values of amino acids in a protein. ProtParam calculates results by analyzing FASTA sequence of proteins and doesn't require any additional information (Wilkins et al., 1999).

2.10. Secondary structure prediction

The secondary structure of final multi-epitope subunit vaccine was predicted using PSIPRED (http://bioinf.cs.ucl.ac.uk/index.php?id=779). This method utilizes incorporation of two feed-forward neural networks which perform the investigation on result retrieved from PSI-BLAST (Position-Specific Iterated BLAST). PSIPRED 3.2 has an average Q_3 score of 81.6% (Jiang et al., 2016).

2.11. Tertiary structure prediction

Raptor X (http://raptorx.uchicago.edu/) is specialized for homology modeling of target proteins having < 30% sequence identity with solved crystal structure available on Protein Data Bank (PDB). This server was used for the modeling of the 3D structure of the final vaccine sequence. It utilizes an input sequence of the protein in FASTA format and predicts its secondary and tertiary structures, contacts, solvent accessibility, disordered regions and binding sites. The quality of predicted 3D-model can be verified by confidence scores assigned to the predicted model by RaptorX. These confidence scores include *P*-value for the relative global quality, GDT (global distance test) for the absolute global quality (Källberg et al., 2012).

2.12. Tertiary structure refinement

The modeled tertiary structure was further refined using GalaxyRefine (http://galaxy.seoklab.org/cgi-bin/subunit.cgi?type=REFINE) which is specialized for performing repeated structure perturbation by means of replacement of amino acids with highest probability rotamers and subsequent overall structural relaxation by molecular dynamics simulation at different time scales. It gives five models in output result, out of which, for model 1, the structural perturbation is restricted to clusters of side chains whereas, for model 2 to 5, more aggressive perturbations were applied to secondary structure elements and loops. It uses triaxial loop method in order to avoid breakage in model structures due to perturbation (Heo et al., 2013).

2.13. Tertiary structure validation

Validation of tertiary structure was done by the generation of Ramachandran plot using RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php). Ramachandran plot is used to predict the probability of particular amino acid to form secondary structure based on dihedral angles namely ϕ and Ψ (allowed and disallowed) of amino acids which are calculated based on the Van der Waals radius of side chain atoms. The quality of modeled tertiary structure was assessed on the basis of percentage of residues in the favored region, allowed region and outlier region (Hardwick et al., 2011). ProSA-Web server was also used for the structural validation purpose and respective *Z*-score was obtained as described elsewhere (Garima Yadav et al., 2017).

2.14. Discontinuous B cell epitope prediction

Discontinuous B cell epitopes were predicted for the final validated 3D structure of vaccine using ElliPro server (http://tools.iedb.org/ellipro/). It assigns each epitope with a score which can be referred to as PI (Protrusion Index) value averaged over residues of the epitope. This method approximates protein's 3D structure by some ellipsoids. A value of 0.9 PI indicates that 90% of protein's residues are being included whereas the remaining 10% are being lying outside of the ellipsoid. This PI value is dependent upon each residue's center of mass which is lying outside of the largest possible ellipsoid. Residues with a high score are said to be associated with high solvent accessibility. Discontinuous epitopes are predicted on the basis of PI value and their clustering is based on the distance R (distance between residue's center of mass) which is measured in Å. A higher value of R indicates a large number of predicted discontinuous epitopes (Ponomarenko et al., 2008).

2.15. Molecular docking of vaccine with immune receptor

Immune receptors are playing an important role to receive the antigenic molecules for the generation of an appropriate immune response (Ali et al., 2017; Pandey et al., 2015). During most of the viral infection Toll-like receptors-3 present on the sentinels of innate immunity, play a crucial role to enhance the antiviral response by its ability to sense double-stranded RNA intermediates. As the same condition originated in the case of chikungunya infection and intracellular TLR-3 bring out the early antiviral immune response by recognizing the pathogen-associated molecular pattern; TLR-3 was selected as a receptor molecule and molecular docking of vaccine protein was performed against it to check the binding affinity of the complex. Molecular docking of the vaccine with TLR3 (PDB ID: 2A0Z) receptor was performed using two servers namely PatchDock (http://bioinfo3d.cs.tau.ac.il/PatchDock/) and COACH (Yang et al., 2013) (https://zhanglab.ccmb.med.umich.edu/COACH/), to check the binding affinity, and interacting residue of protein-ligand complex, respectively, to confirm the development of immune response. PatchDock is an algorithm used for molecular docking of the receptor with the ligand molecule. In this case, both receptor and ligand were protein molecules. PatchDock divides the surface of both input molecules into patches in accordance with the shape of the surface. These patches then further correspond to specific patterns that can visually distinguish between puzzle pieces. After identification of these patches, their superimposition is achieved by using shape matching algorithms. PatchDock algorithm has majorly three stages: (a) molecular shape representation, (b) surface patch matching and (c) filtering and scoring (Schneidman-Duhovny et al., 2005). COACH server was used to identify the protein-ligand interacting residue present at the interface of docked complex, using two comparative methods namely TM-SITE and S-SITE.

2.16. Molecular dynamics simulation of receptor-ligand complex

Molecular dynamics simulation is a widely used approach to study structural properties and microscopic interactions between the receptor and ligand in the docked conformation (Khatoon et al. 2017; Pandey et al., 2016c). In order to confirm the complex stability, molecular dynamics simulation was performed using Gromacs v5.1. The GROMOS96 43A1 force field was used for the same and construction of protein topology was done *via* pdb2gmx. This topology file gives information about bonded and nonbonded parameters. System neutralization was achieved by the addition of appropriate concentration of sodium or chloride counter ions. In this case, the overall charge of the docked complex, as calculated by this software was approximate −3 and so 3 counter sodium ions were added for the purpose of system neutralization. Energy minimization was also performed to ensure that there are no steric clashes or inappropriate geometry in the system prior to simulation and the algorithm used for this process was steepest descent algorithm. System equilibration was achieved *via* NVT (isothermal- isochoric) and NPT ensemble for which leap-frog algorithm was used and the temperature was then brought up to 300 K and 1 bar pressure. After the equilibration process, the simulation was carried out for the receptor-ligand complex for 10 ns time scale to analyze the trajectories (Pandey et al., 2015).

2.17. In silico cloning

For the purpose of cloning and expression of designed subunit vaccine in an expression vector, reverse translation and codon optimization studies of the vaccine protein sequence was performed by using Java Codon adaptation tool (http://www.jcat.de/). This program is specialized for prediction of an optimized coding sequence for any input DNA or protein sequence and its result output includes optimized gene sequence along with their codon adaptive index (CAI) and percentage GC content (Pandey et al., 2018). CAI gives an idea about the influence of natural selection on codon usage bias (Morla et al., 2016). For the expression of the gene in an organism of interest, ideal CAI value should be 1.00 but a value > 0.8 is also considered as good. Percentage of GC content should lie in the range of 30–70%. Values lying outside of this range indicate adverse effects on the transcriptional and translational efficiencies. Secondly, the optimized codons sequence

was subjected to the NEBcutter v2.0 to check the presence of commercially available restriction sites present in it. As the optimized sequence has lack of aforementioned restriction sites, the sequence was reversed by using online tool (http://arep.med.harvard.edu/labgc/adnan/projects/Utilities/revcomp.html) to change the codon position in opposite direction. The codon sequence was reversed to match its complementarity with the pET28a vector replication cycle. Additionally, the optimized codons were justified for their ability to synthesize the same vaccine protein sequence by translation of optimized codons followed by the multiple sequence alignment between vaccine protein and translated protein sequence. Lastly, *Xho*I and *BamHI* restriction sites were added to the N-terminal and C-terminal of the reversed optimized cDNA sequence. The resulting cDNA sequence was further ligated into the pET-28a(+) vector using SnapGene tool to make a clone and ensure its bacterial expression.

3. Results

3.1. Collection of chikungunya viral protein sequences for vaccine construction

The genome of chikungunya virus consists of 11,805 nucleotides and genes encoding for structural and non-structural proteins are organized in the manner: 5′ cap-NSP1-NSP2-NSP3-NSP4-(junction region)-CP-E3-E2-6K-E1-poly(A)-3′ (Solignat et al., 2009). The amino acid sequence of these proteins was obtained in the FASTA format for the *structural (*UniProt ID- Q8JUX5) and nonstructural polyprotein (UniProt ID- Q8JUX6) of *strain S27-African prototype* for the lymphocytes T-cells epitopes prediction.

3.2. CTL epitope prediction

CTL epitopes of all structural and non-structural proteins that were predicted using NetCTL 1.2 server, were screened based on the obtained high score which indicates low sensitivity and high specificity for their adaptive immune receptor. For each protein, top 5 high scorer epitopes were selected to get a total of 35 final CTL epitopes (Supplementary Table 1). Although, no CTL epitopes were obtained for 6K and E3 structural proteins due to their short amino acid sequences.

3.3. HTL epitope prediction

MHC-II binding epitopes of structural and non-structural proteins for mouse alleles (IAb, IAd, IAs, IEb, IEd, and IEs) were predicted using IEDB (Garima Yadav et al., 2017). Top 10 epitopes were selected for each protein on the basis of their least percentile rank and IC_{50} value < 50 nM thereby indicating their higher affinity for the receptor molecule (Chandra et al., 2010). No HTL epitopes were predicted for 6K protein due to its short amino acid sequence. Epitopes of each protein that were having overlapping regions were merged together to get a total of 25 final HTL epitopes (Supplementary Table 2). These HTL epitopes along with final CTL epitopes were then further analyzed for overlapping regions with B-cell epitopes Fig. 1.

3.4. B cell epitope prediction

BCPREDS server was used for the prediction of linear B-cell binding epitopes for all structural and non-structural proteins (Chandra and Singh, 2012). However, no B-cell epitopes were obtained for 6K structural protein due to its short sequence limitation. A total of 68 epitopes having 0.8 and above score (Supplementary Table 3) were chosen for further analysis. These epitopes were then checked for their overlapping sequences with HTL and CTL epitopes and only those epitopes were finally considered for construction of vaccine sequence that was found to have some overlapping regions.

3.5. Multi-epitope vaccine construction

A total of 12 CTL and 14 HTL epitopes (Table 1) having overlapping sequences with B cell epitopes were fused together with the help of AAY and GPGPG linkers, respectively (Pandey et al., 2018). Also in order to increase the immunogenicity of the vaccine, an adjuvant (β-defensin having sequence GIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK) (Mohan et al., 2013) was added at the N-terminal with the help of EAAAK linker. A total of 494 amino acids were there in the final multi-epitope subunit vaccine construct (Fig. 2).

3.6. Conservation across antigen analysis

The conservation analysis of vaccine epitopes with three more strains protein reveals that most of the epitopes are conserved among all strains. If we consider the individual conservancy among three strains, a total of 73%, 73% and 69% epitopes have shown the 100% identity with the Senegal37997, DRDE-06 (Kumar et al., 2014) and Wuerzburg-1 strain, respectively. While, 11.5%, 15% and 19% residues having > 90% identity against the same strains, respectively. All these data showing the good identity and the conserved nature of vaccine protein epitopes (Supplementary table 4).

3.7. IFN-γ inducing and linear B-cell binding epitopes prediction from final vaccine construct

The final vaccine construct including both HTL and CTL epitopes was used for the prediction of IFN-γ inducing epitopes by IFNepitope server. IFN-γ is responsible for inducing a Th1 type of immune response in the host and therefore this vaccine constructs shows the ability to induce a cell-mediated immune response as it contains both IFN-γ inducing and T-cell binding epitopes. Epitopes having score 1 and above (Supplementary Table 5) were considered as final epitopes (Fig. 3A) because high scores can be directly related to the higher affinity of these epitopes for their receptor molecule. Also, linear B-cell binding epitopes were predicted using the same vaccine construct sequence by BCPREDS to explore the humoral immunity induced by the same and only those epitopes were selected which were having 0.8 and above score (Table 2).

3.8. Predictions of allergic and antigenic nature of vaccine construct

Non-allergic nature of vaccine construct was determined by using the AlgPred server. The prediction was performed by using SVM prediction method which is based on amino acid composition and the score obtained was −0.74 which is below the threshold value of −0.4. According to the server parameters, scores less than −0.4 indicates non-allergic nature of a protein. Antigenicity of vaccine construct was determined by using ANTIGENpro and VaxiJen v2.0. The predicted probability of antigenicity by ANTIGENpro was 0.90 which is considered as a good value whereas overall prediction score of the antigen by VaxiJen was 0.48 which is above of 0.4 *i.e.*, the minimum score an antigen should have. Both prediction results indicate the immunogenic nature of vaccine construct.

3.9. Physicochemical properties prediction

ExPASy ProtParam tool was used for the prediction of various physiochemical properties of the final vaccine construct. The molecular weight was calculated as 53.19 kDa which reflects good antigenic nature (antigens having a molecular weight < 5-10KDa are said to have poor immunogenic) (Berzofsky and Berkower, 1993) whereas the pI value of 8.82 indicates that the protein is slightly basic in nature. Estimated half-life was calculated as 30 h in mammalian reticulocytes, *in vitro* and > 20 h in yeast, & > 10 h in *E. coli, in vivo*. An instability index of 38.09 qualified the vaccine as a stable protein and other

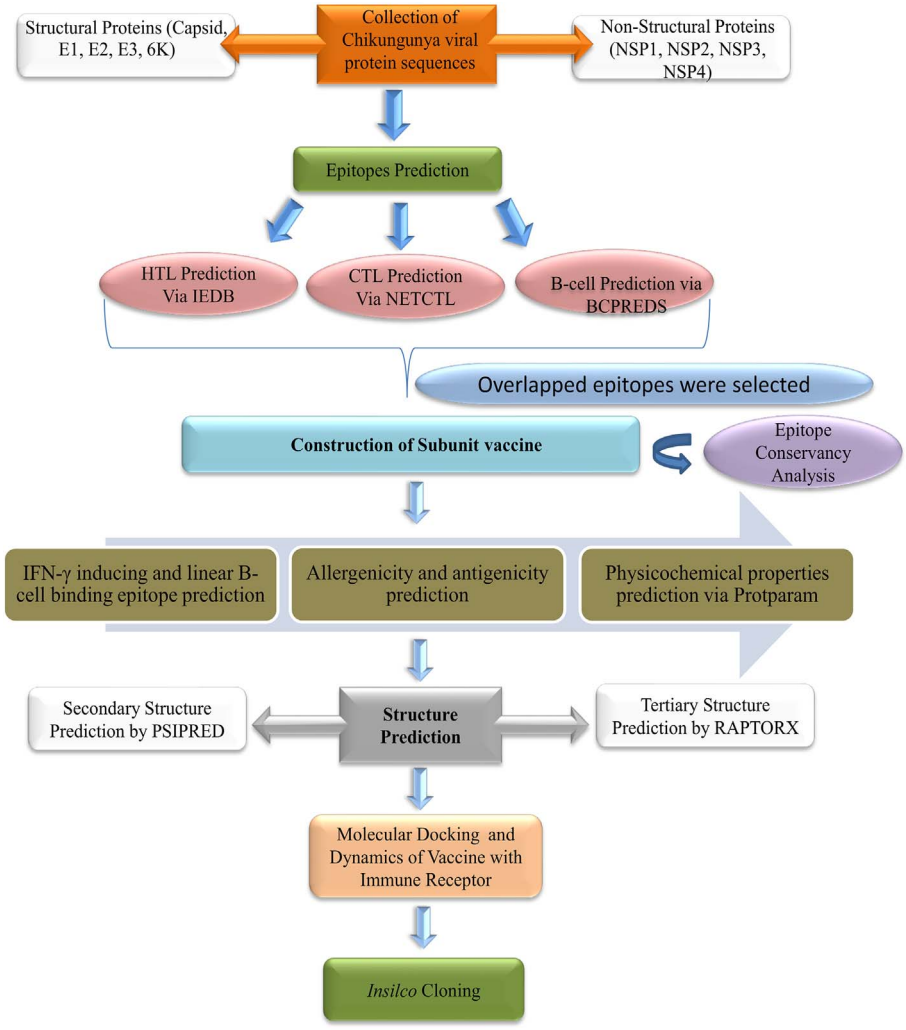

Fig. 1. Systemic workflow of subunit vaccine designing using structural and non-structural proteins of chikungunya virus.

Table 1
Predicted CTL and HTL epitopes among the chikungunya virus structural and non-structural proteins overlapping with B-cell epitopes of the same.

S. No.	Protein	CTL epitope	HTL epitope	B-cell epitope
1	NSP1	MSDRKYHCV	APTSLYHQAIKGVRLAYWV	EARDAEKEAEEEREAELTLE
		DVAIYQDVY	ANARAFSHLAIKLIEQ	KYHCVCPMRSAEDPERLANY
			APARRMMSDRKYHCVC	IYQDVYAVHAPTSLYHQAIK
			REAELTLEALPPLQA	IEQEIDPDSTILDIGSAPAR
2	NSP2	NTDEESYEL	RKFRSSRALKPPCVT	FRSSRALKPPCVTSNTEMFF
				KLHHIAMHGPALNTDEESYE
3	NSP3	VSDWVMSTV	CLCRYAMTPERVTRLR	AAVSDWVMSTVPVAPPRRRR
			RDTAMSLQAPPSTATELS	PVDDADASSPPKTVPCLCRY
			RFHQTAVDMAEIYTM	QAPPSTATELSHPPISFGAP
				EIYTMWPKQTEANEQVCLYA
4	NSP4	SSYQITDEY	DRRRALADEVVRWQR	GKPLAAGDEQDEDRRRALAD
		ITDEYDAYL		ITTENLTTYVTKLKGPKAAA
		ITTENLTTY		RNYPTVSSYQITDEYDAYLD
5	CP	KSDASKFTH	ISAVNKLTMRAVPQQKPR	KPRRNRKNKKQKQKQQAPQN
		FTHEKPEGY	GANEGARTALSVVTWNKDI	VVTWNKDIVTKITPEGAEEW
				KFTHEKPEGYYNWHHGAVQY
6	E1	SLDYITCEY	SQLQISFSTALASAE	VYNMDYPPFGAGRPGQFGDI
		KGDVYNMDY		REAEIEVEGNSQLQISFSTA
				ITCEYKTVIPSPYVKCCGTA
7	E2	VTWGNNEPY	VTNHKKWQYNSPLVPRNAEL	AEEIEVHMPPDTPDRTLLSQ
			LPCSTYVQSNAATAEE	PLVPRNAELGDRKGKIHIPF
				VTWGNNEPYKYWPQLSANGT

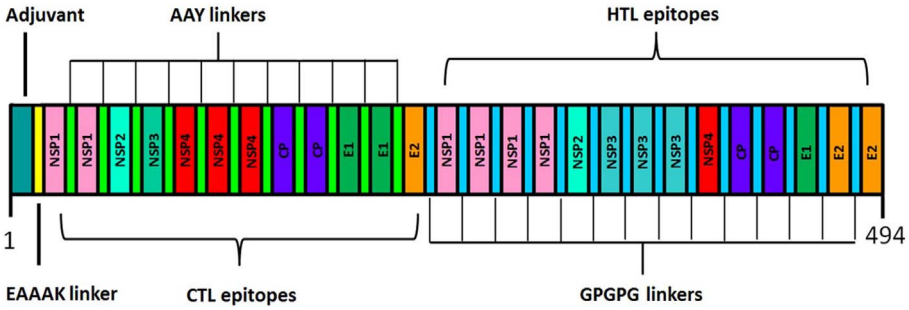

Fig. 2. Schematic arrangement of multi-epitope based subunit vaccine construct. Multi-epitope subunit vaccine sequence consisting of 494 amino acids has been represented where adjuvant was linked at N-terminal fused with a multi-epitope sequence with the help of EAAAK linker (yellow). CTL and HTL epitopes of all structural and non-structural proteins have been joined together by AAY linkers (parrot green) and GPGPG linkers (light blue), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

physicochemical properties were an aliphatic index (61.17) and GRAVY (−0.488). The high aliphatic index is an indication of high thermostability of a protein and a negative GRAVY value reflects hydrophilic nature of protein which is responsible for better interaction with the polar environment.

3.10. Secondary structure prediction

PSIPRED was used for the prediction of secondary structure of final subunit vaccine construct. 25.5% of the sequence consists of alpha helices, 17% consists of beta strands whereas the remaining 57.5% is composed of random coils. These results are an indication of possible inter-chain and intra-chain interactions that might be involved in the formation of protein's secondary structure. Also, they indicate the percentage of amino acid residues that favor the formation of different types of secondary structures based on their hydrophobic and hydrophilic nature.

3.11. Tertiary structure prediction

The tertiary structure of the multi-epitope based subunit vaccine construct was modeled using Raptor X (Fig. 3B). The best template for the building of model was 4BL8A. *P*-value (calculated by Raptor X) was 2.66×10^{-3}, and overall uGDT (un-normalized Global Distance Test) was equal to 129. The *p*-value for relative global quality indicates the deviation of predicted protein structure model from the already solved crystal structure that is available on PDB and GDT value gives a measure of the absolute global quality of predicted protein structural model. Out of 494 residues, 352 (72%) were modeled whereas 52 (10%) positions predicted were disordered. These values are indicative of an overall good tertiary model of the subunit vaccine maximum number of residues being modeled.

Table 2
Linear B-cell binding epitopes mapping for the sequence of final vaccine construct.

S. No.	Position	Epitope	Score
1	396	RAVPQQKPRGPGPGGANEGA	1
2	231	IKLIEQGPGPGAPARRMMSD	1
3	355	MAEIYTMGPGPGDRRRALAD	1
4	271	ALPPLQAGPGPGRKFRSSRA	1
5	184	TWGNNEPYGPGPGAPTSLYH	1
6	315	TRLRGPGPGRDTAMSLQAPP	1
7	209	VRLAYWVGPGPGANARAFSH	1
8	420	SVVTWNKDIGPGPGSQLQIS	1
9	445	ASAEGPGPGVTNHKKWQYNS	1
10	292	KPPCVTGPGPGCLCRYAMTP	1
11	467	VPRNAELGPGPGLPCSTYVQ	1
12	91	VMSTVAAYSSYQITDEYAAY	0.881
13	64	VAIYQDVYAAYNTDEESYEL	0.88

3.12. Tertiary structure refinement

Refinement of the vaccine model was done using GalaxyRefine server. Out of all refined models, model number 5 was found to be the best one on the basis of various parameters including GDT-HA (0.9375), RMSD (0.455), MolProbity (2.099), Clash score (14.0), Poor rotamers (0.7) and Rama favored (93.0). GDT-HA indicates the global quality of protein's model, RMSD gives a measure of the deviation of bond lengths and bond angles of the refined tertiary structure from the non-refined tertiary structure and so RMSD value should be minimum. MolProbily is a score given for relative local and global quality of the refined model, Clash score is a function of steric clashes in the refined structure and poor rotamers gives a number of amino acid residues that show less possible rotation in their side chains. The method followed by GalaxyRefine in the refinement of protein's tertiary structure involves

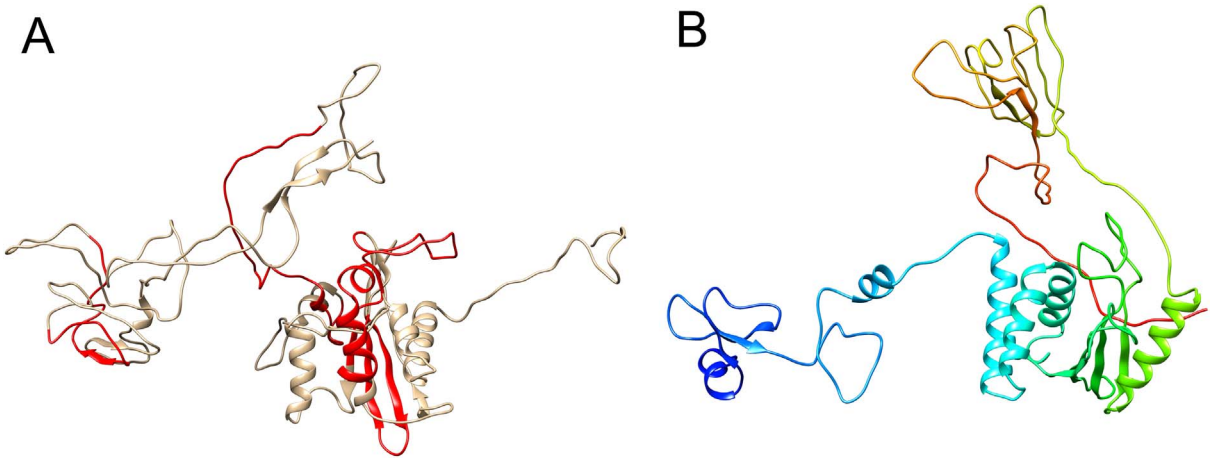

Fig. 3. IFN-γ inducing epitopes and multi-epitope vaccine model (A) Epitope sequences responsible for inducing IFN-γ have been depicted in red colour (B) Represents final 3D model of multi-epitope vaccine obtained after homology modeling, refinement of tertiary structure and its validation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

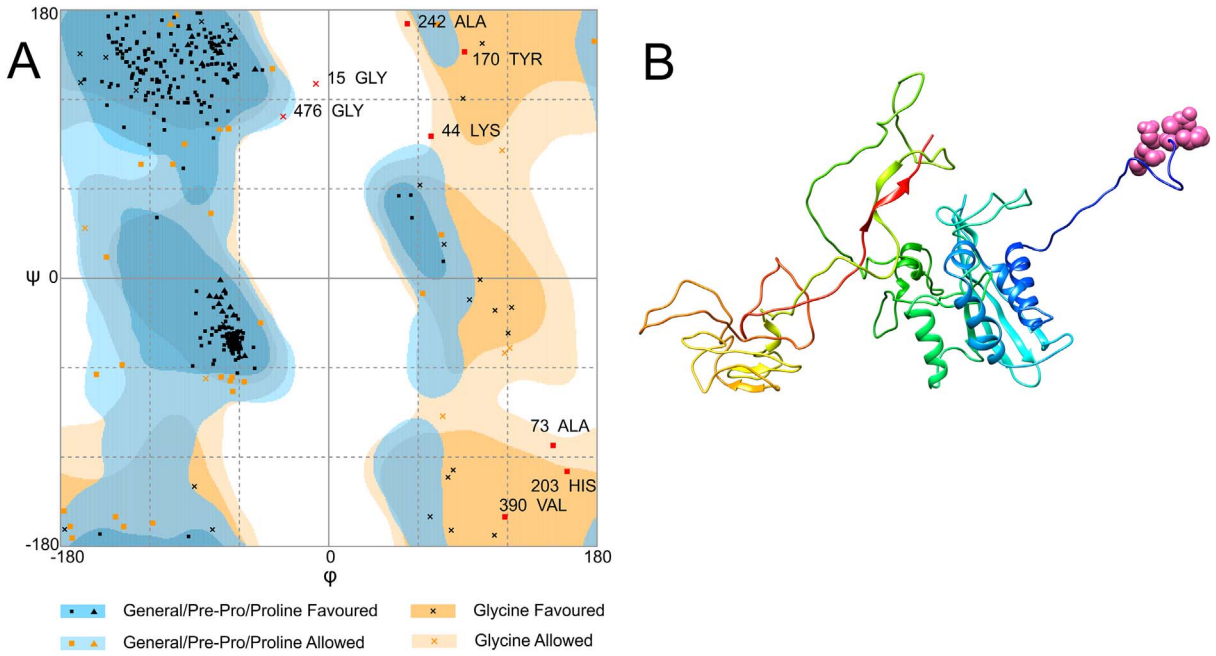

Fig. 4. Vaccine model validation and conformational B-cell epitope prediction (A) The refined 3D structure of multi-epitope vaccine was validated by the generation of Ramachandran plot where 94.7% residues were found to lie in the favored region whereas 4.3% and 1.1% residues were there in allowed and outlier region, respectively. (B) The conformational B-cell epitopes have been depicted as pink spheres. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

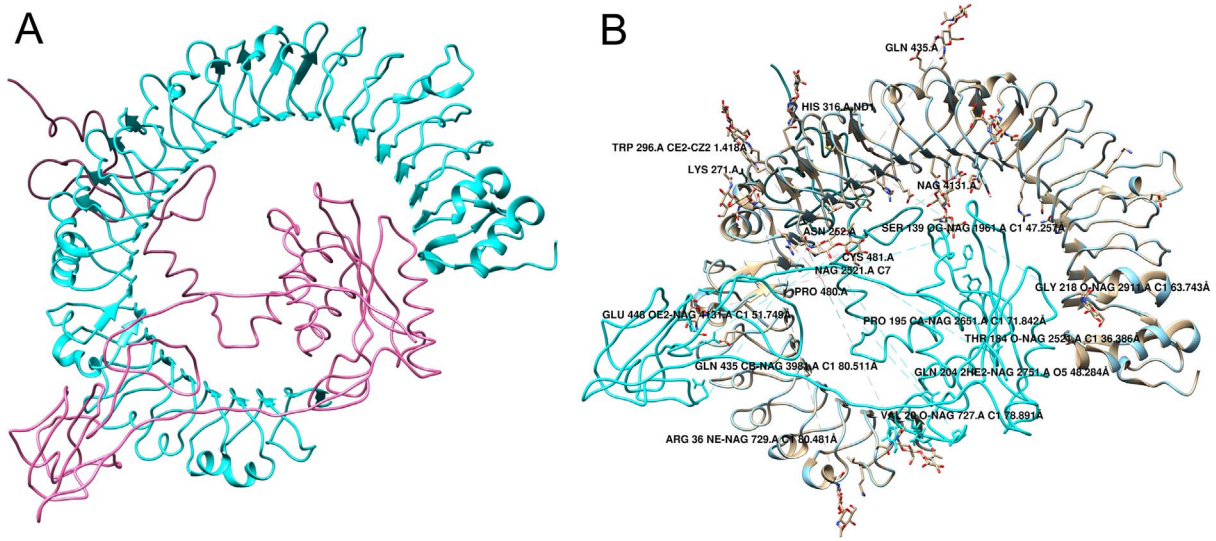

Fig. 5. Docked complex of TLR-3 (PDB ID: 2A0Z) with the multi-epitope vaccine. (A) Receptor (TLR-3) has been depicted in cyan colour whereas pink colour depicts the multi-epitope vaccine as a ligand in the docked complex obtained from molecular docking (B) representing the interacting residues present at the interface of vaccine protein and TLR-3 complex. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

replacement of amino acid residues with highest probable rotamers at the initial stage which is also known as repeated structural perturbation and this is further followed by mild and aggressive relaxation process involving molecular dynamics simulation at different time scales. Therefore, this model was taken as the final vaccine model for the further analysis as it was having best scores for above-mentioned parameters out of all refined models.

3.13. Tertiary structure validation

Validation of the refined 3D structure of multi-epitope based subunit vaccine was performed by generation of Ramachandran plot through RAMPAGE. Prior to the refinement of vaccine model, Ramachandran plot showed 88.5% residues in the favored region, 9.4% residues in allowed region and 2.1% residues in the outlier region. After refinement of the model, 94.7% residues were lying in the favored region whereas 4.3% and 1.1% residues were there in allowed and outlier region respectively (Fig. 4A). ProSA-Web server was also used for the structural validation and Z-score was found to be −2.51.

3.14. Discontinuous B-cell epitope prediction

Discontinuous B-cell binding epitopes were predicted using ElliPro. A total of 6 epitopes were predicted out of which top scorer (0.974) epitope was selected as the final epitope (Fig. 4B). The PI value (score given by ElliPro server) of 0.974 depicts that 97.4% residues are lying in the predicted ellipsoid region of the epitope and this particular epitope is having the highest solvent accessibility.

3.15. Molecular docking of vaccine with immune receptor (TLR-3)

The binding of refined vaccine model with TLR-3 (PDB ID- 2A0Z) was evaluated by performing molecular docking using PatchDock. While the residues participating in this binding complex and present at the interface of the protein-ligand complex were determined using COACH meta server approach. Among these two servers, the former requires an input of PDB format of both receptor and ligand molecules. RMSD value which was used by the server for final clustering was 4.0 for the purpose of protein-protein docking and the complex type was set as default. Out of all docking results, top 10 results were chosen for further analysis and docked complex number 5 (Fig. 5A) showed best interactions between receptor and ligand in the complex form. The geometric shape complexity score was 22,082, approximate interface area of the complex was 4101.8, atomic contact energy (ACE) was 263.70 and 3D transformations including 3 rotational angles and 3 translational parameters were 0.92, 1.01, 1.78, −100.67, 0.40 and − 18.98, respectively for the fifth docked complex. This docked complex was further subjected to COACH and the interacting residues were found at the interface of the docked complex (Fig. 5B). Those interacting residues were ASN252, LYS271, CYS481, HIS316, GLN435, PRO480 from the receptor (TLR-3) molecule, while GLN435, GLU448, GLN204, PRO195, GLY218, and THR184 from the designed vaccine candidate. This complex was further utilized for performing molecular dynamics simulation studies.

3.16. Molecular dynamics simulation of receptor-ligand complex

In order to explore stable microscopic interactions between receptor (TLR-3, PDB ID: 2A0Z) and ligand (multi-epitope subunit vaccine) in the docked complex, molecular dynamics simulation was performed using GROMACS v5.1. At the initial stage, energy minimization was performed and various parameters such as pressure, temperature, potential energy and volume were evaluated. The temperature progression curve (Fig. 6A) indicates that the system attained 300K of temperature very quickly and then remained constant for rest of the 100 ps equilibration period. Pressure progression curve (Fig. 6B) shows the fluctuation of pressure over the entire equilibration phase of 100 ps with an average value of 1 bar pressure. The output of MD simulation was obtained as Root Mean Square Deviation (RMSD) of vaccine backbone and Root Mean Square Fluctuation (RMSF) of all amino acid residues. RMSD of the receptor-ligand complex reflecting stability over a 10 ns timescale has been shown in the graph (Fig. 6C). The deviation started with 0.2 nm value but at the end of 10 ns time period, 0.7 nm of RMSD was obtained. RMS fluctuation depicts stability of vaccine residues during the simulation time period. Mild fluctuations in side chains of most amino acid residues are indicative of establishment of imperishable interactions with the immune receptor TLR-3 while high fluctuation is an indicator of the presence of highly flexible regions within the complex (Hajighahramani et al., 2017). Regions of higher flexibility have been shown in the graph (Fig. 6D) with the higher peak with an RMSF of approximately 0.25 nm.

3.17. In silico cloning

For the efficient cloning and expression of vaccine protein inside *E. coli* pET28a expression vector, reverse translation and codon optimization studies was performed and found that the CAI value and GC content for the optimized gene was 0.97 and 56.3, respectively. The CAI value for the vaccine lies in the range of 0.8–1.0 which is an ideal value and shows good expression whereas GC content of the sequence also falls in the ideal range (30–70%). The optimized codon sequence has shown the absence of commercially available restriction sites, therefore it was good for the cloning into pET28a vector. After evaluation on these parameters, *Xho*I and *BamHI* restriction sites were added at N and C terminal of the reverse optimized codon sequences, respectively. Ultimately, the clone of vaccine gene into pET28a(+) vector was obtained showing its ability to express in the microbial system (Fig. 7).

4. Discussion

Vaccines play a crucial role in providing protection against a particular disease to host organism, therefore it provides help in saving millions of lives annually across the globe. Vaccine development processes are generally very laborious, expensive and demand a long period of their accomplishment. There has been much encroachment in the area of computational biology that aids many types of research and helps in diminishing the expected time consumption. Therefore, vaccine designing can be achieved by using various parameters of immunoinformatics which is a bioinformatics-based approach. It involves identification of potential antigenic proteins of a pathogen and then using this information, this process further helps in the determination of various immune-dominant epitopes that are responsible for the development of both humoral and cell-mediated type of immune responses against the pathogen inside the host organism. Multi-epitope based subunit vaccines can, therefore, be designed by defining the B and T-cell epitopes regions on potential antigen proteins of a pathogen. The first vaccine which was designed following immunoinformatics approach was against *Neisseria meningitidis* and it was successfully developed thereafter (Adu-Bobie et al., 2003). Many researchers were then able to develop vaccines using the same methodology in subsequent years and some of these are effectual vaccines against *Streptococcus pneumoniae*, *Chlamydia pneumoniae*, *Staphylococcus aureus* (Adu-Bobie et al., 2003), *Rickettsia prowazekii* (Caro-Gomez et al., 2014), enterotoxigenic *Escherichia coli etc.* (Mehla and Ramana, 2016). There have been numerous attempts in the development of a vaccine against Chikungunya infection but no such attempt is reported to date where the development of a subunit vaccine has been succeeded following computational approach. In this study, we have tried to design a multi-epitope based subunit vaccine against CHIKV using immunoinformatics approach and the efficacy of this vaccine in terms of its receptor binding has been explored using molecular docking and molecular dynamics simulation studies. This vaccine consists of β-defensin as an adjuvant at the N-terminal trailed by a sequence of various CTL and HTL epitopes present on both structural and non-structural proteins of CHIKV. These epitopes along with the adjuvant have been fused together with the assistance of appropriate linkers.

Subunit vaccines lack a bit of immunogenicity as the only limited number of epitopes is involved but it can be enhanced with the help of some adjuvants therefore to design a vaccine, here, β-defensin was added at the N-terminal with the help of EAAAK linker. β-defensin is accountable for generating both innate and cell-mediated immunity as they recruit both immature dendritic cells and naïve T-cells at the location where infection has occurred and they achieve this by binding to their immune receptors such as TLRs and CCR6. The final multi-epitope based subunit vaccine construct constitutes of 494 amino acids and its antigenicity, allergenicity along with physicochemical properties was determined afterward. In order to validate the immunogenicity of designed vaccine protein among the other chikungunya virus strains, epitope conservancy analysis was also performed and found that most of the CTL and HTL epitopes are identical with other three strains. The secondary structure and tertiary structure predictions of the final vaccine construct sequence were performed using PSIPRED and RaptorX respectively. Secondary structure prediction was executed to determine the proportion of amino acids involved in the formation of alpha helices, β sheets, and random coils. Tertiary structure prediction through homology modeling was done to get a 3D model of the vaccine which can be used for the inspection of probable interactions with its specific human immune receptor. This vaccine model was further exposed to refinement by repeated structural perturbations and structural relaxation by GalaxyRefine. In order to check the binding affinity of vaccine model with TLR-3, molecular docking studies were performed using

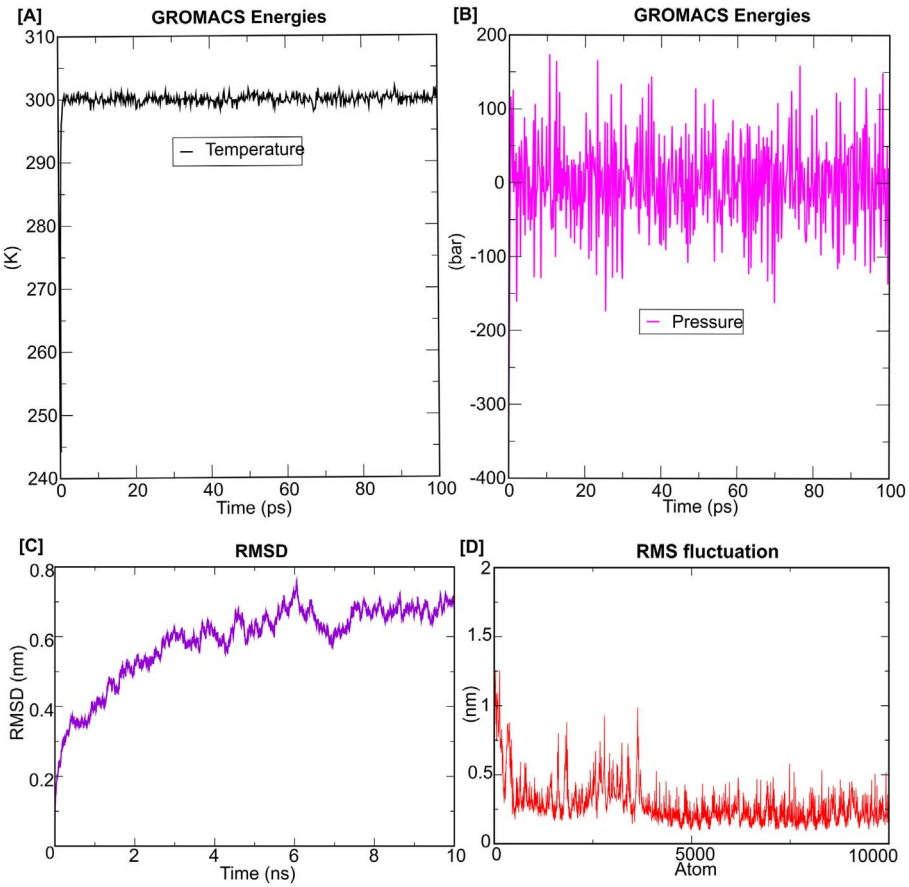

Fig. 6. Molecular dynamics simulation of receptor-ligand (TLR-3 & vaccine) complex, and relative components (A) Temperature progression curve of receptor-ligand (TLR-3-vaccine) complex indicating that temperature of the system attains a value of 300 K very quickly and remains near around the same value for rest of the equilibration period (100 ps). (B) Pressure progression curve of receptor-ligand (TLR-3-vaccine) complex indicating pressure fluctuation over 100 ps equilibration phase with an average value of pressure being 1 bar. (C) RMSD representation of the docked complex protein backbone consists of TLR-3 as a receptor and multi-epitope vaccine as a ligand. (D) RMSF representation of the docked complex protein side chains.

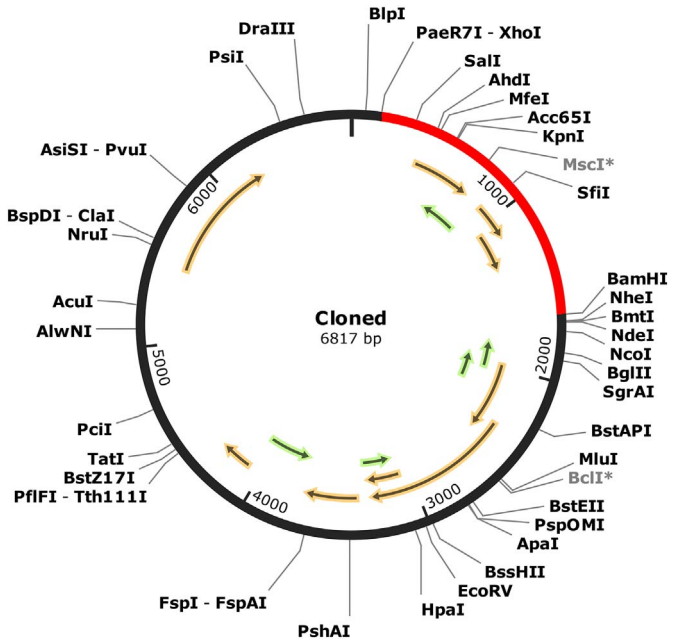

Fig. 7. *In silico* restriction cloning of optimized and reversed codon sequences showing single cDNA fragment insert (Red colored semi circle) into pET28a(+) vector (Black colored semi circle). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

PatchDock server and the interacting residues were determined by using COACH meta-server. This vaccine-TLR-3 complex was then tested for its stability by molecular dynamics simulation at 10 ns using Gromacs. Finally, to check the translational efficiency of the final vaccine construct in an expression system, reverse translation using Java Codon adaptation Tool were accomplished. The CAI and GC content were found to lie in the required range.

Dendritic cells are considered as the intermediate connector between the innate and acquired immunity and upon activation with pathogen-associated molecular pattern (PAMP), it differentiated into the matured phenotype and express the MHC conjugated antigenic epitopes. Some of the PAMP may have the ability to provide the appropriate direction to acquired immunity *via* TLRs present on the dendritic cells. The Toll-like receptor is broadly expressed in immune as well as non-immune cells. TLR3 has a dual location, in un-stimulated cells, it is located in the endoplasmic reticulum while upon stimulation with double-stranded RNA, it moves towards the endosome and encounter with the target ligand in a pH-dependent manner (Wang et al., 2010). Normally, TLR-3 stimulated by the viral pathogen-associated molecular pattern *i.e.*, dsRNA but additionally polyriboinosinic polyribocytidylic acid [poly(I:C)], has shown its ability to being used as critical stimuli for TLR-3 (Alexopoulou et al., 2001). The selective high expression of human CD141 + dendritic cells decides the utilization of TLR3 ligands as an adjuvant molecule for the vaccine based therapeutic options (Perales-Linares and Navas-Martin, 2013). Therefore, in the present study, utilization of [poly(I:C)] along with the final vaccine construct may stimulate the TLR-3 signaling followed by intracellular trafficking (Fujimoto et al., 2004). Ultimately, due to the high binding affinity of adjuvant, vaccine protein may bind to TLR-3 followed by internalized and processing. Therefore it will activate the adaptive limb of cellular immunity and memory cells will be formed to tackle the same pathogen adaptive immune response.

5. Conclusion

Chikungunya virus infection has emerged as a severe problem at a

worldwide level and reasons for a large number of deaths in tropical and sub-tropical regions across the globe. Due to its similarity in symptoms with dengue, this disease is sometimes misdiagnosed in the regions where dengue infection is prevalent. Despite such an alarming data, there is no permanent cure and prevention for this infection. Recently there have been numerous attempts in the development of a vaccine against CHIKV and only three have entered in the clinical trials phase I (Schwameis et al., 2016). In this study, we have tried to develop a multi-epitope based subunit vaccine through immunoinformatics approach. This method is advantageous over classical vaccine development methods including heat-killed or attenuated viral vaccines as it avoids the culturing of pathogenic viral strains and such vaccines are easy to synthesize and manufacture. We first retrieved the amino acid sequence of all structural and non-structural proteins of CHIKV and then predicted their potential B-cell, T-cell binding, and IFN-γ inducing epitopes. Finally, a vaccine construct was prepared including overlapping sequences of HTL, CTL along with an immunogenic adjuvant fused together by the appropriate linker sequences. Vaccine model was generated using various bioinformatics approaches and the interaction of this vaccine with its specific immune receptor TLR-3 was evaluated by performing molecular docking and molecular dynamics simulation. The *in silico* cloning studies of the vaccine was also performed thereafter to check its probable efficiency of being expressed in an expression vector.

Acknowledgment

AN and VKP are thankful to the Central University of Rajasthan for providing the computational facility. RKP is thankful for Department of Science and Technology (IF140901) for providing INSPIRE fellowship.

Author contribution

Protocol designed by - AN, RKP, VKP.
Methodology performed by - AN, RKP, AM, NK, VKP.
Manuscript was written by - AN, RKP, VKP.

Competing financial interests

The Authors have declared no competing interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2018.03.007.

References

Adu-Bobie, J., Capecchi, B., Serruto, D., Rappuoli, R., Pizza, M., 2003. Two years into reverse vaccinology. Vaccine 21, 605–610.
Ahluwalia, P.K., Pandey, R.K., Sehajpal, P.K., Prajapati, V.K., 2017. Perturbed microrna expression by Mycobacterium tuberculosis promotes macrophage polarization leading to pro-survival foam cell. Front. Immunol. 8, 107. http://dx.doi.org/10.3389/fimmu.2017.00107.
Akahata, W., Yang, Z.-Y., Andersen, H., Sun, S., Holdaway, H.A., Kong, W.-P., Lewis, M.G., Higgs, S., Rossmann, M.G., Rao, S., 2010. A virus-like particle vaccine for epidemic Chikungunya virus protects nonhuman primates against infection. Nat. Med. 16, 334–338.
Akhrymuk, I., Kulemzin, S.V., Frolova, E.I., 2012. Evasion of the innate immune response: the Old World alphavirus nsP2 protein induces rapid degradation of Rpb1, a catalytic subunit of RNA polymerase II. J. Virol. 86, 7180–7191.
Alexopoulou, L., Holt, A.C., Medzhitov, R., Flavell, R.A., 2001. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 413, 732–738.
Ali, M., Pandey, R.K., Khatoon, N., Narula, A., Mishra, A., Prajapati, V.K., 2017. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. Sci. Rep. 7, 9232. http://dx.doi.org/10.1038/s41598-017-09199-w.
Arankalle, V.A., Shrivastava, S., Cherian, S., Gunjikar, R.S., Walimbe, A.M., Jadhav, S.M., Sudeep, A.B., Mishra, A.C., 2007. Genetic divergence of Chikungunya viruses in India (1963–2006) with special reference to the 2005–2006 explosive epidemic. J. Gen. Virol. 88, 1967–1976.
Barh, D., Misra, A.N., Kumar, A., Azevedo, V., Kangueane, P., 2010. A novel strategy of epitope design in Neisseria gonorrhoeae. Bioinformation 5, 77–85.
Berzofsky, J.A., Berkower, I., 1993. Immunogenicity and antigen structure. Fundamental Immunol. 5, 631–684.
Brathwaite Dick, O., San Martín, J.L., Montoya, R.H., del Diego, J., Zambrano, B., Dayan, G.H., 2012. The history of dengue outbreaks in the Americas. Am. J. Trop. Med. Hyg. 87, 584–593.
Bui, H.H., Sidney, J., Li, W., Fusseder, N., Sette, A., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinform. 8, 361.
Caro-Gomez, E., Gazi, M., Goez, Y., Valbuena, G., 2014. Discovery of novel cross-protective rickettsia prowazekii T-cell antigens using a combined reverse vaccinology and in vivo screening approach. Vaccine 32, 4968–4976.
Cecilia, D., 2014. Current status of dengue and chikungunya in India. WHO. South. East. Asia. J. Public. Health. 3 (22–26), 2014. http://dx.doi.org/10.4103/2224-3151.115828.
Chander, S., Pandey, R.K., Penta, A., Choudhary, B.S., Sharma, M., Malik, R., Prajapati, V.K., Murugesan, S., 2017. 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. 20, 734–746.
Chandra, S., Singh, T.R., 2012. Linear B cell epitope prediction for epitope vaccine design against meningococcal disease and their computational validations through physicochemical properties. Netw. Model Anal. Health Inform. Bioinform. 1, 153–159.
Chandra, S., Singh, D., Singh, T.R., 2010. Prediction and characterization of T-cell epitopes for epitope vaccine design from outer membrane protein of Neisseria meningitidis serogroup B. Bioinformation 5, 155–161.
Chen, J., Liu, H., Yang, J., Chou, K.-C., 2007. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids 33, 423–428.
Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform. 8, 4. http://dx.doi.org/10.1186/1471-2105-8-4.
Duvvuri, V.R., Duvvuri, B., Alice, C., Wu, G.E., Gubbay, J.B., Wu, J., 2014. Preexisting CD4+ T-cell immunity in human population to avian influenza H7N9 virus: whole proteome-wide immunoinformatics analyses. PLoS One 9, e91273.
El-Manzalawy, Y., Dobbs, D., Honavar, V., 2008a. Predicting flexible length linear B-cell epitopes. Comput. Systems Bioinform. Comput. Syst. Bioinform. Conf. 7, 121–132.
El-Manzalawy, Y., Dobbs, D., Honavar, V., 2008b. Predicting linear B-cell epitopes using string kernels. J. Mol. Recogn. 21, 243–255.
Fujimoto, C., Nakagawa, Y., Ohara, K., Takahashi, H., 2004. Polyriboinosinic polyribocytidylic acid [poly (I: C)]/TLR3 signaling allows class I processing of exogenous protein and induction of HIV-specific CD8+ cytotoxic T lymphocytes. Int. Immunol. 16, 55–63.
Hajighahramani, N., Nezafat, N., Eslami, M., Negahdaripour, M., Rahmatabadi, S.S., Ghasemi, Y., 2017. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus. Infect. Genet. Evol 48, 83–94.
Hardwick, S.W., Chan, V.S., Broadhurst, R.W., Luisi, B.F., 2011. An RNA degradosome assembly in Caulobacter crescentus. Nucleic Acids Res. 39, 1449–1459.
Heo, L., Park, H., Seok, C., 2013. GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res. 41, W384–W388.
Jiang, Y., Oron, T.R., Clark, W.T., Bankapur, A.R., D'Andrea, D., Lepore, R., Funk, C.S., Kahanda, I., Verspoor, K.M., Ben-Hur, A., 2016. An expanded evaluation of protein function prediction methods shows an improvement in accuracy. Genome Biol. 17, 184. http://dx.doi.org/10.1186/s13059-016-1037-6.
Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., Xu, J., 2012. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. 7, 1511–1522.
Khan, S.A., Dutta, P., Topno, R., Borah, J., Chowdhury, P., Mahanta, J., 2015. Chikungunya outbreak in Garo Hills, Meghalaya: an epidemiological perspective. Indian J. Med. Res. 141, 591–597.
Khatoon, N., Pandey, R.K., Prajapati, V.K., 2017. Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. Sci. Rep. 7, 8285. http://dx.doi.org/10.1038/s41598-017-08842-w.
Kumar, A., Mamidi, P., Das, I., Nayak, T.K., Kumar, S., Chhatai, J., Chattopadhyay, S., Suryawanshi, A.R., Chattopadhyay, S., 2014. A novel 2006 Indian outbreak strain of Chikungunya virus exhibits different pattern of infection as compared to prototype strain. PLoS One 9, e85714.
Larsen, M.V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., Nielsen, M., 2007. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinform. 8, 424. http://dx.doi.org/10.1186/1471-2105-8-424.
Magnan, C.N., Baldi, P., 2014. SSpro/ACCpro 5: almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles, machine learning and structural similarity. Bioinformatics 30, 2592–2597.
Mehla, K., Ramana, J., 2016. Identification of epitope-based peptide vaccine candidates against enterotoxigenic Escherichia coli: a comparative genomics and immunoinformatics approach. Mol. BioSyst. 12, 890–901.
Mohan, T., Sharma, C., Bhat, A.A., Rao, D., 2013. Modulation of HIV peptide antigen specific cellular immune response by synthetic α-and β-defensin peptides. Vaccine 31, 1707–1716.
Morla, S., Makhija, A., Kumar, S., 2016. Synonymous codon usage pattern in glycoprotein gene of rabies virus. Gene 584, 1–6.
Moyle, P.M., Toth, I., 2013. Modern subunit vaccines: development, components, and research opportunities. ChemMedChem 8, 360–376.
Nezafat, N., Ghasemi, Y., Javadi, G., Khoshnoud, M.J., Omidinia, E., 2014. A novel multi-epitope peptide vaccine against cancer: an in silico approach. J. Theor. Biol. 349, 121–134.

Pandey, R.K., Prajapati, V.K., 2018. Exploring sand fly salivary proteins to design multi-epitope subunit vaccine to fight against visceral leishmaniasis. J. Cell. Biochem. http://dx.doi.org/10.1002/jcb.26719.

Pandey, R.K., Sharma, D., Bhatt, T.K., Sundar, S., Prajapati, V.K., 2015. Developing imidazole analogues as potential inhibitor for Leishmania donovani trypanothione reductase: virtual screening, molecular docking, dynamics and ADMET approach. J. Biomol. Struct. Dyn. 33, 2541–2553.

Pandey, R.K., Prajapati, P., Goyal, S., Grover, A., Prajapati, V.K., 2016a. Molecular modeling and virtual screening approach to discover potential Antileishmanial inhibitors against ornithine decarboxylase. Comb. Chem. High Throughput Screen. 19, 813–823.

Pandey, R.K., Sundar, S., Prajapati, V.K., 2016b. Differential expression of miRNA regulates T cell differentiation and plasticity during visceral Leishmaniasis infection. Front. Microbiol. 7, 206. http://dx.doi.org/10.3389/fmicb.2016.00206.

Pandey, R.K., Verma, P., Sharma, D., Bhatt, T.K., Sundar, S., Prajapati, V.K., 2016c. High-throughput virtual screening and quantum mechanics approach to develop imipramine analogues as leads against trypanothione reductase of leishmania. Biomed. Pharmacother. 83, 141–152.

Pandey, R.K., Kumbhar, B.V., Srivastava, S., Malik, R., Sundar, S., Kunwar, A., Prajapati, V.K., 2017a. Febrifugine analogues as Leishmania donovani trypanothione reductase inhibitors: binding energy analysis assisted by molecular docking, ADMET and molecular dynamics simulation. J. Biomol. Struct. Dyn. 35, 141–158.

Pandey, R.K., Kumbhar, B.V., Sundar, S., Kunwar, A., Prajapati, V.K., 2017b. 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. 37, 60–70.

Pandey, R.K., Narula, A., Naskar, M., Srivastava, S., Verma, P., Malik, R., Shah, P., Prajapati, V.K., 2017c. Exploring dual inhibitory role of febrifugine analogues against Plasmodium utilizing structure-based virtual screening and molecular dynamic simulation. J. Biomol. Struct. Dyn. 35, 791–804.

Pandey, R.K., Bhatt, T.K., Prajapati, V.K., 2018. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating Anopheles salivary protein. Sci. Rep. 8, 1125. http://dx.doi.org/10.1038/s41598-018-19456-1.

Perales-Linares, R., Navas-Martin, S., 2013. Toll-like receptor 3 in viral pathogenesis: friend or foe? Immunology 140, 153–167.

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 Bioinform. 9, 514.

Potocnakova, L., Bhide, M., Pulzova, L.B., 2016. An introduction to B-cell epitope mapping and in silico epitope prediction. J Immunol Res 2016, 6760830. http://dx.doi.org/10.1155/2016/6760830.

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

Saha, S., Raghava, G., 2006. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res. 34, W202–W209.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. 33, W363–367.

Schwameis, M., Buchtele, N., Wadowski, P.P., Schoergenhofer, C., Jilma, B., 2016. Chikungunya vaccines in development. Hum. Vaccin. Immunother. 12, 716–731.

Simon, F., Javelle, E., Oliver, M., Leparc-Goffart, I., Marimoutou, C., 2011. Chikungunya virus infection. Curr. Infect. Dis. Rep. 13, 218.

Solignat, M., Gay, B., Higgs, S., Briant, L., Devaux, C., 2009. Replication cycle of chikungunya: a re-emerging arbovirus. Virology 393, 183–197.

Thangamani, S., Higgs, S., Ziegler, S., Vanlandingham, D., Tesh, R., Wikel, S., 2010. Host immune response to mosquito-transmitted chikungunya virus differs from that elicited by needle inoculated virus. PLoS One 5, e12137.

Vita, R., Overton, J.A., Greenbaum, J.A., Ponomarenko, J., Clark, J.D., Cantrell, J.R., Wheeler, D.K., Gabbard, J.L., Hix, D., Sette, A., 2015. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 43, D405–D412.

Wang, Y., Liu, L., Davies, D.R., Segal, D.M., 2010. Dimerization of Toll-like receptor 3 (TLR3) is required for ligand binding. J. Biol. Chem. 285, 36836–36841.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. 112, 531–552.

Yadav, Garima, Rao, Richa, Raj, Utkarsh, Varadwaj, P.K., 2017. Computational modeling and analysis of prominent T-cell epitopes for assisting in designing vaccine of ZIKA virus. J. Appl. Pharm. Sci. 7, 116–122.

Yang, J., Roy, A., Zhang, Y., 2013. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29, 2588–2595.