Received: 6 August 2018 | Accepted: 10 September 2018

DOI: 10.1002/jcp.27531

ORIGINAL RESEARCH ARTICLE

WILEY Journal of Cellular Physiology

Combinatorial screening algorithm to engineer multiepitope subunit vaccine targeting human T-lymphotropic virus-1 infection

Rajan K. Pandey[1]* | Rupal Ojha[1]* | Nina Chatterjee[1] | Nitesh Upadhyay[1] | Amit Mishra[2] | Vijay K. 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, Rajasthan, India

Correspondence
Vijay K. Prajapati, Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, NH-8, Bandarsindri, Kishangarh, 305817 Ajmer, Rajasthan, India.
Email: vkprajapati@curaj.ac.in

Abstract

Human T-lymphotropic virus (HTLV), the first human retrovirus has been discovered which is known to cause the age-old assassinating disease HTLV-1 associated myelopathy. Cancer caused by this virus is adult T cell leukemia/lymphoma which targets 10–20 million throughout the world. The effect of this virus extends to the fact that it causes chronic disease to the spinal cord resulting in loss of sensation and further causes blood cancer. So, to overcome the complications, we designed a subunit vaccine by the assimilation of B-cell, cytotoxic T-lymphocyte , and helper T-lymphocyte epitopes. The epitopes were joined together along with adjuvant and linkers and a vaccine was fabricated which was further subjected to 3D modeling. The physiochemical properties, allergenicity, and antigenicity were evaluated. Molecular docking and dynamics were performed with the obtained 3D model against toll like receptor (TLR-3) immune receptor. Lastly, in silico cloning was performed to ensure the expression of the designed vaccine in pET28a(+) expression vector. The future prospects of the study entailed the in vitro and in vivo experimental analysis for evaluating the immune response of the designed vaccine construct.

KEYWORDS

adjuvant, epitopes, human T-lymphotropic virus (HTLV-1), pET28a, subunit vaccine, TLR-3 agonist

1 | INTRODUCTION

Human T-lymphotropic virus (HTLV) belongs to the group of retroviruses which are infamous for causing adult T-cell leukemia/lymphoma (ATL). Among four different serotypes of HTLVs, HTLV-1 is responsible for the HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP) which leads to the gradual demyelinating inflammatory neurologic infection (Khabbaz et al., 1993). HTLV-1 has shown its affinity for the T-lymphocytes and thus primarily affects the CD4 T-lymphocytes, on the contrary HTLV-1I has the tendency to infect the CD8 T-lymphocytes. The transmission of both HTLV-1 and II happens through intravenous drug injection, blood transfusion, breast milk feeding, sexual contact and further, can be incorporated into the vascular system (Wolfe et al., 2005). The HTLV-1 and HTLV-2 both share approximately 70% of the genome as well as nucleotide homology. When we differentiate the HTLV-1 from HTLV-2 on the basis of pathogenesis, it has been reported that the HTLV-1 is pathogenic and causes ATL and severe neurological disorders in comparison to HTLV-2, which is principally nonpathogenic and very rare cases of HTLV-2 infection among the population have been stated. We have designed the subunit vaccine against the HTLV-1 because of the higher mortality and morbidity rate among the

*Rajan K. Pandey and Rupal Ojha contributed equally to this study.

population and that is why the study is more focused on the HTLV-1 infection (Jones et al., 2006). Geographically this disease is distributed in Southwestern Japan, South and Central America Caribbean regions, Brazil, Caribbean community living in the United Kingdom, parts of Asian subcontinent and East and North Africa (Gessain & Cassar, 2012). The HTLV-1 infection eventually develops into fatal leukemia and alone targets nearly 0.5 million people, worldwide. Sometime, it also develops an infirming myelopathy, uveitis, inflammatory disorders or other infectious dermatitis (Proietti, Carneiro-Proietti, Catalan-Soares, & Murphy, 2005). The clinical manifestation of this disease is associated with the sphincter, upper motor neurons and mild sensory dysfunction. The histopathological findings depict the degeneration of axons, thoracic, and spinal cord atrophy which involves the perivascular demyelination. Immunological hallmark for this disease involves the rapid and spontaneous proliferation of the peripheral blood lymphocytes within the patients (Greten et al., 1998). The morbidity and mortality are caused primarily due to the HAM/TSP or ATL. The ATL has a latency period of 30–50 years, and is gradually fatal having a median survival period of 2 years (Yasunaga & Matsuoka, 2007).

Even after the severe assault of HTLV-1 mediated pathogenesis, by far, no treatment intervention is available to cure the acute or chronic attach of HTLV-1 infection. In 2012, Mogamulizumab, a defucosylated humanized anti-CCR4 IgG1 monoclonal antibody was registered in Japan as therapeutic options to treat the ATL. But its Phase-II trial results are pending for approval in the United State of America (USA; Yamauchi et al., 2014). Eventually, chemotherapy is the only available way to tackle this severe diseased condition. Patients suffering from HTLV-1 pathogenesis have shown a good initial response to the CHOP scheme consisting of cyclophosphamide, doxorubicin (hydroxydaunomycin), vincristine (Oncovin), and prednisone. Some other regimens used for the disease prevention constitute of zidovudine plus interferon alfa (IFN-α), arsenic trioxide plus IFN-α, topoisomerase inhibitors, or NF-κB inhibitor.

With the aim to contribute to the path of cancer vaccine development, we tried to design a multiepitope subunit vaccine which may have the ability to protect from HTLV-1 infection. The designed vaccine only consists of the immunogenic B-cell and T-cell epitopes instead of any viral live component and hence reduces the risk factors of pathogenic reversal. In this study, the exploration of antigenic structural and regulatory proteins of HTLV-1 oncovirus was done for the prediction of immunogenic binders. Initially, a literature search was performed to find the structural and regulatory proteins of HTLV-1 and found that its structural protein is encoded by the gag, pol, and env genes. The former proteins involved in the virion assembly and infection and composed of three proteins namely nucleocapsid, matrix, and capsid. Pol gene encode for the three viral proteins, reverse transcriptase, integrase, and RNase-H and play a key role in the viral life cycle. The *env* protein assists the virus entry into a new cell for the transmission of infection. It is composed of two subunits, a transmembrane and a surface protein (Le Blanc et al., 2001). HTLV genome also encodes for the regulatory and accessory proteins that are basic for the viral industriousness and disease progression. For example, Px gene actively participates in the open-reading-frame transcription and hence encodes for five regulatory proteins namely p12, p13, p30, rex, and Tax (Boxus et al., 2008). This way, structural and regulatory proteins of HTLV-1 was selected and subjected to predict the B-cell, cytotoxic T-lymphocyte (CTL), and helper T-lymphocyte (HTL) epitopes. These predicted epitopes were then conjugated together along with suitable adjuvants and linkers to construct the subunit vaccine (Khatoon, Pandey, & Prajapati, 2017). Furthermore, the vaccine protein sequence was subjected to the physiochemical characterization followed by the allergenicity and antigenicity prediction. The 3D model of vaccine protein was modeled and refined and further, the tertiary structure was subjected to the docking process. The molecular docking was performed between the toll like receptor (TLR-3) receptor and the vaccine construct (ligand). Then, the obtained docked complex was subjected for molecular dynamics simulation to check their interaction and stability. Lastly, the in-silico cloning was done within pET28a(+) expression vector for the futuristic expression of vaccine protein

2 | MATERIALS AND METHODOLOGY

2.1 | Retrieval of viral protein sequences

The literature survey was performed to sort the viral virulent proteins followed by their amino acids sequence retrieval from National Centre for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/protein). Virulent proteins include three structural and five regulatory proteins of HTLV-1. The structural proteins included gag, pol, and env whereas the regulatory proteins include tax, rex, p12, p13, and p30. A TLR-3 agonist mainly β-defensin was used as an adjuvant to enhance the subunit vaccine persistence and immunogenicity. All the virulent proteins were used for the epitope prediction to construct the subunit vaccine which may have the ability to prevent viral carcinoma.

2.2 | B-cell epitope prediction

As a part of the humoral limb of the immune system, B-cell epitopes for all the eight proteins were predicted using ABCpred server (http://crdd.osdd.net/raghava/abcpred/; Saha & Raghava, 2006). Human B cell receptor can recognize the epitope with a length of 10–23 amino acids followed by the generation of plasma and memory B-cells. The plasma B-cell is responsible for the antibody production during primary infection and the memory B-cell generates antibody during secondary infection (Pandey, Ali, Ojha, Bhatt, & Prajapati, 2018; Pandey, Ojha, Mishra, & Prajapati, 2018). Likewise, the server obtains the B-cell epitopes from B cell epitopes database; it removes the identical and nonimmunogenic epitopes. Thus, only the unique and continuous B-cell epitopes for the given sequence was retained and presented as output. The network yields a prediction accuracy of 65.93% after a five-fold cross-validation.

2.3 | Cytotoxic T-cell (CTL) epitope prediction

CTLs are used for killing the cells presenting MHC-I conjugated viral peptide by releasing granzymes and perforins. It thereby helps in reducing the viral titer and cellular infection in our body. The CTL epitopes were predicted for the structural and regulatory proteins of HTLV-1 using NetCTL 1.2 server (http://www.cbs.dtu.dk/services/NetCTL/; Larsen et al., 2007). This server performs the CTL epitope prediction in 12 MHC-I supertypes restricted manner, where the weight matrix is used for TAP efficiency prediction and artificial neural network performs MHCI binding and TAP cleavage. Three different supertypes A2, A3, and B7 were selected because these supertypes covers around 88.9% of the world's population (Pandey, Ojha, Aathmanathan, Krishnan, & Prajapati, 2018; Sette & Sidney, 1999) and used to predict the CTL epitopes for each protein and the one with highest combined score was selected from each supertype to make the subunit vaccine construct.

2.4 | Helper T-cell (HTL) epitope prediction

Helper T-cell are important to establish the whole-body immune response against a pathogen. The naïve T cell differentiates into different T cell subtypes like Th1, Th2, Treg, and Th17 and so on. Among them, Th1 releases IFN-γ, interleukin (IL)-2, and IL-12 and is responsible to activate macrophages and CTL cells and ultimately helps in pathogen clearance. Therefore, HTL epitopes were predicted for both for the structural and regulatory proteins of HTLV-1 using the IEDB online server (http://tools.iedb.org/mhcii/; Wang et al., 2008) and sorted epitopes were used for the vaccine designing to enhance the cell-mediated immune response. The resultant HTL epitopes were screened based on their lowest percentile rank and IC50 values which depicts that these sequences possess the highest affinity for T-cell receptor. All those HTL epitopes having a percentile rank of ≤1.0 were further subjected to perform IFN-γ and IL-4 epitope prediction.

2.5 | Designing the multiepitope subunit vaccine

The obtained B-cell, CTL, and HTL epitopes for the eight HTLV-1 proteins were now merged together with the help of linkers to construct the subunit vaccine candidate. The B-cell epitopes were joined with KK linkers, whereas CTL and the HTL were linked with AAY and GPGPG linkers, respectively (Rana & Akhter, 2016; Yang et al., 2015). Linkers are indispensable and play a very crucial role in providing stability to the protein complex. The adjuvant β-defensin was joined at the N-terminal of the vaccine construct with EAAAK linker (Pandey, Bhatt, & Prajapati, 2018). The adjuvant functions like an agonist of TLR-3 immune receptor and may enhance vaccine entry followed by antigen processing by antigen presenting cells.

2.6 | Antigenicity and allergenicity prediction of the designed vaccine construct

The antigenicity and allergenicity of the designed subunit vaccine were verified using the VaxiJen v.2 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen_help.html) and AllerTOP v.2.0 servers (http://www.ddg-pharmfac.net/AllerTOP). VaxiJen is alignment independent server which works on auto cross co variance transformation of protein sequence into a uniform vector of amino acid properties. This server was designed to check antigenicity of vaccine candidate based on the physiochemical character of the protein (Doytchinova & Flower, 2007). The AllerTOP sever applies Quality Surface Activity Relationship methods to different length peptides. The input was classified according to the k-nearest species algorithm (Dimitrov, Flower, & Doytchinova, 2013).

2.7 | Physiochemical properties analysis

The EXPASY ProtPram server was used for determining the physiochemical properties of designed subunit vaccine (https://web.expasy.org/cgi-bin/protparam/protparam; Gasteiger et al., 2005). This server helps to deduce the various properties of the vaccine candidate which helps in determining the mode of administration, the target site, the form in which it could be administered and half-life of the vaccine construct in the human body. The output includes predicted molecular weight (kDa), theoretical pI, half-life, grand average of hydropathy (GRAVY), aliphatic index and the extinction coefficient.

2.8 | Tertiary structure prediction

The local and global arrangement of amino acids which involves interaction from its side chain residues forms the tertiary structure of a protein. A stable protein must have the lowest energy and entropy. The tertiary structure of the vaccine was predicted using the RaptorX structure prediction tool (http://raptorx.uchicago.edu/StructurePrediction/; Källberg et al., 2012), which uses template-based prediction method. After submitting the subunit vaccine sequence as input, the solvent accessibility, global distance test (GDT) as well as the local and global quality was predicted for the tertiary structure model.

2.9 | Refinement of the obtained 3D model and validation

The refinement of the obtained structure is necessary for improving template-based structure which relies on the information retrieved from the template. GalaxyRefine (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) was used for refinement of the 3D model which works best in improving the quality of the local structure (Ko, Park, Heo, & Seok, 2012). The server rebuilds the side chain first and

then repacks, followed by the structure relaxation using molecular dynamics simulations (Heo, Park, & Seok, 2013). The 3D model output of RaptorX was given as the input and the output was obtained as five refined models. Validation of the obtained structure was done using RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) which plots Ramachandran plot of the phi verses psi dihedral angle of each amino acid present in proteins (Furnham et al., 2006).

2.10 | Molecular docking between the vaccine protein and the immune receptor

Molecular docking was done using ClusPro web tool (https://cluspro.org/tut_dock.php) which is a web-based program for the automatic proteins docking. The used algorithm searches billions of the ligand-receptor interactions and retains the ones with lowest energy values. The TLR-3 immune receptor (receptor) and vaccine construct (ligand) were subjected for performing the docking and the interaction between them was checked. Obtained clusters were divided based on the energy of all the docked complexes and top 1,000 structures with least energy were clustered based on their interface root means square deviation values. The one with lowest Gibbs free energy was selected for further analysis (Kozakov et al., 2017). Again, docking was performed for the same using PatchDock server (http://bioinfo3d.cs.tau.ac.il/PatchDock/) for the comparative validation. From this server, we got 10 models and these models were ranked based on geometry, and electrostatic complementarity of the protein surface. These were submitted to FireDock (Fast Interaction Refinement in Molecular Docking) server for further refinement. Top ten models were received as an output, among which the final model was chosen for molecular dynamics simulation study.

2.11 | Molecular structure dynamics of receptor-ligand complex

Gromacs v5.1.2 (Hess, Kutzner, Van Der Spoel, & Lindahl, 2008; Pandey, Kumbhar, Sundar, Kunwar, & Prajapati, 2017; Pandey, Sharma, Bhatt, Sundar, & Prajapati, 2015) standalone software, was used to carry out the molecular dynamics simulation studies to check the stability of TLR3-vaccine docked model. The parameters of the system are set according to the human physiological parameters and the protonation states of the side-chains residues such as Asp, Glu, Arg, Lys, and His were adjusted accordingly (R. Shukla, Shukla, Sonkar, Pandey, & Tripathi, 2018). The system was solvated in a cubic box using the SPC water model, with a buffer distance of 1 nm set between the protein and edge of the box. Energy minimization was carried out using the steepest descent method (H. Shukla, Shukla, Sonkar, & Tripathi, 2017). Following energy minimization, the system was equilibrated under NVT conditions to heat the system up to 300 K, using the V-rescale thermostat. NPT equilibration was carried out using the Berendsen barostat to bring the pressure of the system up to 1 bar. Position restraints applied to the solute were released and molecular dynamics simulations were carried out for the time duration of 20 ns under NPT conditions using the Parrinello-Rahman barostat. The trajectory was analyzed to obtain the backbone root mean square deviation (RMSD) and side chain RMSF values for the complex (H. Shukla, Shukla, Sonkar, Pandey, & Tripathi, 2017).

2.12 | Codon optimization and in silico cloning

To prevent the vaccine construct from getting mutated in the prokaryotic system, we need to optimize the codons in accordance to the host, as the codon usage of the host can be different from where the protein of our interest came from. For this process, Java Adaptation Tool (JCat) (http://www.jcat.de/; Grote et al., 2005) was used which work on codon adaptation index values (CAI). The CAI value was obtained by the server itself by summing all the codon usage of all the genes of that organism. The amino acid sequence of our vaccine was given as input. The selected organism for expressing our vaccine construct was *Escherichia coli* strain K12. Further rho-independent termination, ribosomal binding sites and the cleavage sites of restriction endonucleases were avoided. The output of JCat comprised of CAI value and GC content of DNA sequence which can assure the quality of our codon optimization. The output sequence was then reversed using reverse or compliment DNA sequence server (http://arep.med.harvard.edu/labgc/adnan/projects/Utilities/revcomp.html). The reversed sequence was used to clone the adapted vaccine construct sequence in *E. coli* pET-28a(+) vector. *Xho*I and *Bam*HI restriction sites were selected for creating nicks and selecting the required gene sequence and inserting it into the vector using SnapGene tool for expressing vaccine into *E. coli*. The entire flow of the work has been shown in the flow chart (Figure 1).

3 | RESULTS AND DISCUSSION

As chemotherapy is the only option to prevent the HTLV-1 mediated carcinogenesis and none of the vaccine candidates is registered against this severe diseased condition. This study was designed to construct the subunit vaccine candidate as reported elsewhere (Ali, Pandey, Khatoon, Narula, & Mishra, 2017; Chatterjee, Ojha, Khatoon, & Prajapati, 2018; Ojha, Khatoon, & Prajapati, 2018; Pandey et al.,

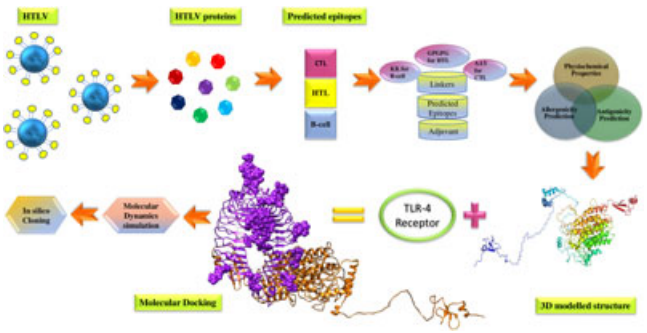

FIGURE 1 Simplified diagrammatic representation of the designed work followed for the construction of the multiepitope subunit vaccine [Color figure can be viewed at wileyonlinelibrary.com]

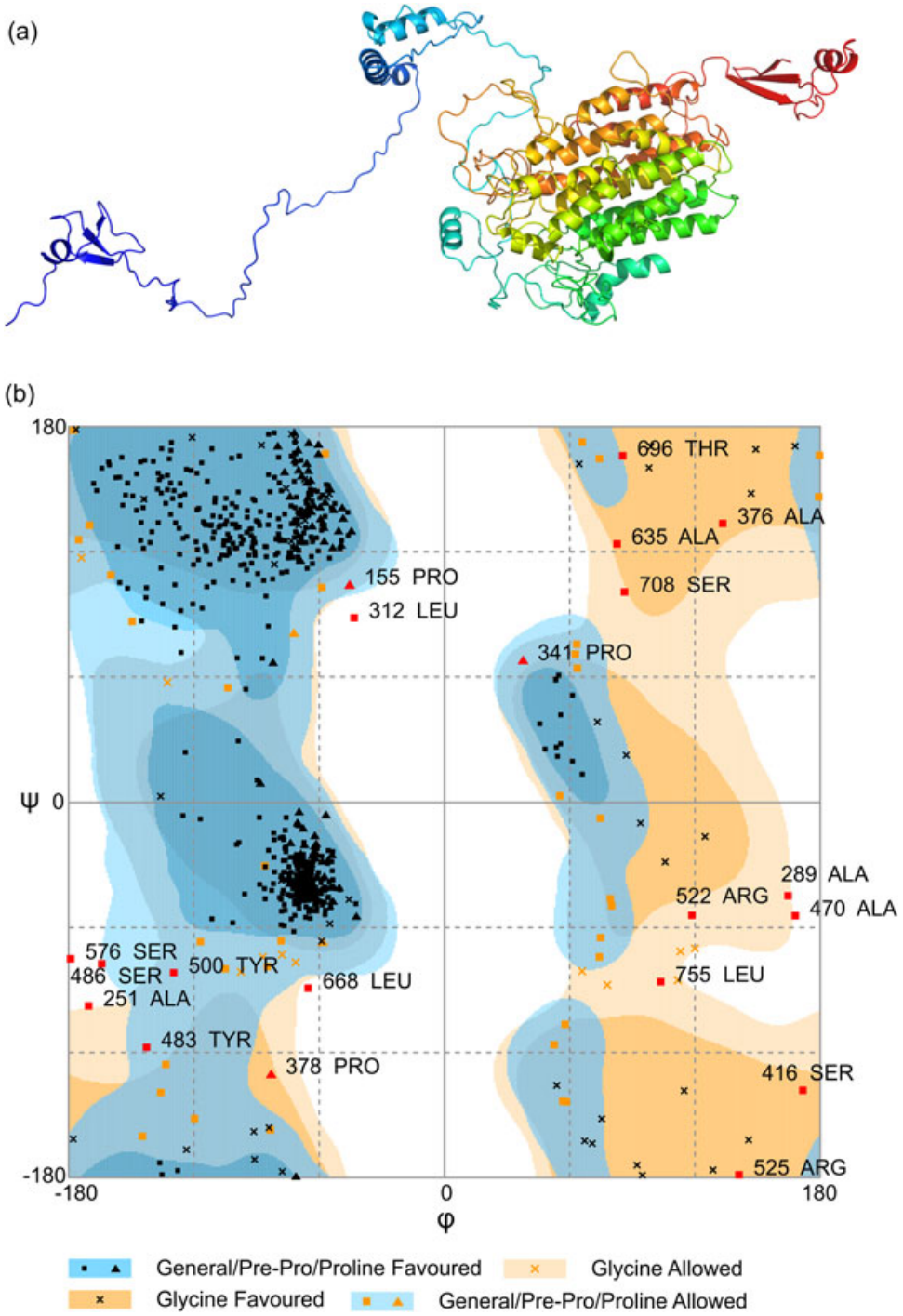

FIGURE 2 (a) 3D modeled structure of the designed vaccine construct. The structure obtained using RaptoxX server followed by refinement to improve the number of residues in favored region using GalaxyRefine server. (b) Validation of the 3D refined structure of subunit vaccine with the help of Ramachandran plot through RAMPAGE. In the modeled vaccine construct, 91.6% residues were lying in the favored region whereas 6% and 2.4% residues were lying in allowed and outlier region, respectively [Color figure can be viewed at wileyonlinelibrary.com]

2018; Pandey & Prajapati, 2018). A literature survey was done and eight immunogenic proteins including three structural proteins and five nonstructural proteins were sorted to complete this study. The structural proteins include gag, pol, and env whereas nonstructural proteins include tax, rex, p12, p13, and p30. All these nonstructural proteins have regulatory functions.

3.1 | Prediction of linear B-cell epitope

The antigen interacts with B-cell followed by B-cell activation and release antibody to develop a humoral immune response against the pathogen. The B-cell epitopes were predicted for each structural and regulatory proteins of HTLV-1 using ABCpred

server and top scoring two epitopes corresponds each protein was selected to be a part of subunit vaccine construct. Where higher the epitope score higher will be the probability of being a B-cell epitope. All the B-cell epitopes were of a 10mer length and total 16 B-cell epitopes were selected and used for the vaccine construction (Supporting Information Table-1).

3.2 | Prediction of CTL epitope

CTL epitopes for all the structural and regulatory proteins were predicted using the NetCTL 1.2 server. Each epitope was obtained with combined score where a high score indicates low sensitivity and high specificity. Each protein was subjected to CTL epitope prediction against A2, A3, and B7 supertype. This way total 122, 133, and 62 number of epitopes were predicted for A2-, A3-, and B7 supertypes, respectively. While for the vaccine designing top one epitope correspond each supertype was selected based on its highest score. Therefore, total 24 CTL epitopes were selected and used for the vaccine designing (Supporting Information Table 2).

3.3 | Prediction of HTL epitope

HTL develops cell-mediated immune response and release different cytokine and interleukin against the pathogenic peptides. The HTL epitopes of 15mer length against reference allele sets were predicted. Among them, only those epitopes having lowest percentile rank and IC_{50} value of ≤500 were sorted and used for the next set of evaluation. Overall 14 epitopes were selected and used for vaccine designing. This way variable number of epitopes was obtained for the structural and regulatory proteins of HTLV-1. IFN epitope and the IL4pred server were used to predict the ability of HTL epitopes as IFN-γ positive and IL-4 non-inducer. This way total 14 HTL epitopes with IFN-γ positive and IL-4 non-inducer epitope was selected and subjected to vaccine designing (Supporting Information Table 3). Only single epitopes were obtained for both the P13 and REX proteins, passing on the above parameters.

3.4 | Multiepitope vaccine construction

A multiepitope subunit vaccine was designed by the combination of humoral and cell-mediated immune response specific epitopes. Therefore, 16 B-cell epitopes, 14 CTL and 24 HTL epitopes were fused together with the help of KK, AAY, and GPGPG linkers, respectively. Also to increase the immunogenicity of vaccine protein, an adjuvant (β-defensin having sequence GIINTLQKYYCRVRGGR-CAVLSCLPKEEQIGKCSTRGRKCCRRKK; Mohan, Sharma, Bhat, & Rao, 2013) was added at the N-terminal with the help of EAAAK linker. A subunit vaccine construct of 808 amino acids residue was obtained consisting of B-cell, CTL, and HTL epitopes along with linkers and adjuvant (Supporting Information Figure 1).

3.5 | Antigenic and allergic nature prediction

Antigenicity of vaccine construct was determined by using VaxiJen v2.0 server. The predicted probability of antigenicity by VaxiJen was 0.694 which is above the threshold value of 0.4. The prediction results indicate that the vaccine construct can be considered as a probable antigen. The allergenicity of multisubunit vaccine was determined by Allertop found to be probable nonallergen.

3.6 | Physicochemical properties

ExPASy-ProtParam tool was used for the prediction of various physiochemical properties of the final multiepitope subunit vaccine construct. The molecular weight was calculated as 88 kDa which reflects good antigenic nature whereas the pI value of 10.38 indicates that the protein is slightly basic in nature. Estimated half-life was calculated as 30 hr in mammalian reticulocytes, in vitro and >20 hr in yeast, and >10 hr in *E. coli*, in vivo. An instability index of 34.97 qualified the vaccine as a stable protein and other physico-chemical properties were an aliphatic index (86.86) and GRAVY (−0.047). The high aliphatic index is indicative of high thermo stability of a protein and a GRAVY value reflects hydrophilic nature which is responsible for better interaction with the polar environment of the human body.

3.7 | Tertiary structure prediction and refinement

The 3D model for the designed subunit vaccine was obtained using RaptoxX server followed by refinement to improve the number of residues in the favored region using GalaxyRefine server. Out of all the refined models, Model-1 was found to be the best one on the basis of various parameters including GDT-HA (0.9143), RMSD (0.507), MolProbity (2.650), Clash score (32.9), Poor rotamers (1.4), and Rama favored (90.3). This model was taken as the final vaccine model for the further analysis (Figure 2a).

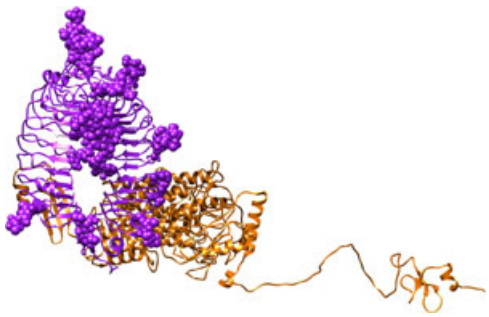

FIGURE 3 The binding of refined vaccine model with TLR-3 immune receptor (PDB ID- 2A0Z) was achieved by the ClusPro server. The selected cluster has shown the best interaction with the maximum negative lowest energy of −1230 kJ mol^{-1}. The receptor (TLR-3) has been represented in purple color whereas the interacting vaccine construct has been displayed in golden color. PDB: protein data bank; TLR-3: toll like receptor [Color figure can be viewed at wileyonlinelibrary.com]

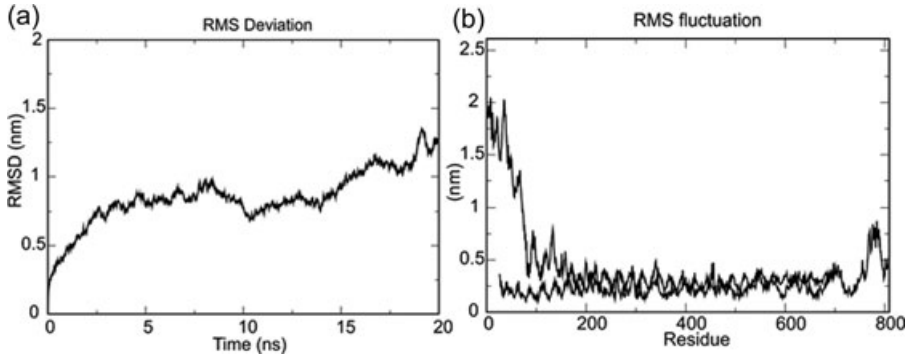

FIGURE 4 Molecular dynamic simulation. The stability of the docked complex was validated with the help of molecular dynamics simulation. The RMSD output plot for the protein backbone was ~0.75. Initially, the deviation of 0.25 starts at 0 ns which gradually increases till 2 ns then after becomes stable up to 20 ns with minor fluctuation. While the side chain fluctuation was found to be an average of 0.25 nm. RMSD: root mean square deviation

3.8 | Tertiary structure validation

Validation of the refined 3D structure of subunit vaccine was performed by generation of Ramachandran plot through RAMPAGE. Before the refinement of vaccine model, Ramachandran plot showed 89.6% residues in favored region, 6.3% residues in allowed region and 4% residues in the outlier region. After refinement of the model vaccine model 91.6% residues were lying in the favored region whereas 6% and 2.4% residues were there in allowed and outlier region, respectively (Figure 2b).

3.9 | Molecular docking of the designed vaccine construct and receptor (TLR-3)

The binding affinity of refined vaccine model with the TLR-3 immune receptor (PDB ID- 2A0Z) was evaluated by performing molecular docking study using ClusPro and PatchDock servers. The ClusPro server gives 30 results outputs as 30 model, and we were selected the Model-2 showing best result based on the maximum negative lowest energy of −1230 (Figure 3). The PatchDock server requires PDB file of vaccine protein as input which was used by the server for

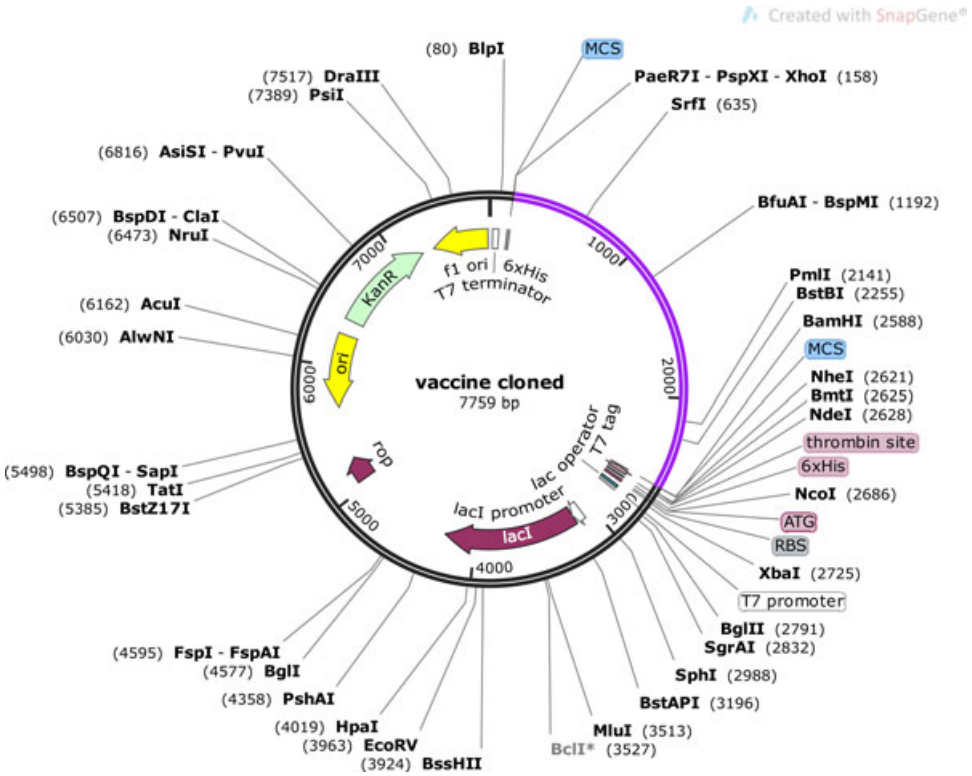

FIGURE 5 The in silico cloning of the designed vaccine construct into expression vector pET28a(+). In circle the purple colored part representing the codons corresponds to designed vaccine whereas the remaining black part illustrates the expression vector pET28a [Color figure can be viewed at wileyonlinelibrary.com]

final clustering purpose. Out of all docked complex resulted from PatchDock server, top 10 results were chosen for further analysis and FireDock docked complex number eight showed best interactions between receptor and ligand in the complexed form based on the maximum negative global energy.

3.10 | Molecular dynamics simulation

Molecular dynamics simulation study was performed to validate the complex stability as obtained by the molecular docking of the subunit vaccine and TLR-3 immune receptor. The RMSD output plot for the protein backbone was determined and found to be ~0.75. Initially, the deviation of 0.25 starts at 0 ns which gradually increases till 2 ns then after becomes stable up to 20 ns with minor fluctuation. While the side chain fluctuation was found to be an average of 0.25 nm. All these results validate the stable nature of docked complex (Figure 4a, b and Supporting Information Figure 2a,b).

3.11 | In silico cloning

To estimate the expression of designed vaccine construct into selected *E. coli* strain, in silico cloning was performed into pET28a(+) expression vector. First, the procedure of codon optimization (codon usage) was performed with the help of Java Codon Adaptation tool. The obtained CAI value and average GC content of the improved sequence were 0.93% and 55.56%, respectively. The reversed sequence was used to clone the adapted vaccine construct sequence in *E. coli* pET-28a(+) vector. *Xho*I and BHamI restriction sites were selected for digestion and selecting the required gene sequence and inserting it into the vector using SnapGene tool for expressing the subunit vaccine into *E. coli*. From the output, it can be concluded that the obtained optimized sequence has the potential to express in the *E. coli* system successfully (Figure 5).

4 | CONCLUSIONS

HTLV infection is eventually increasing as a life challenging issue which is spreading rapidly throughout the globe and have huge chances of incrementing in the upcoming years. Reportedly the statistics of mortality are sufficiently high enough to draw our attention towards healing the sufferers. There are no such existing medicines which have a very high success rate of curing the disease. The existing chemotherapies have huge side effects and few of the medicines did not even pass the Phase-II trials. Thus, there is an urgent need to develop a multiepitope subunit vaccine using immunoinformatic approaches. To enhance the humoral and cell-mediated immunity, the B-cell, CTL, and HTL epitope sequences are linked together by using proper linkers. The generated vaccine model was further studied for allergenicity and antigenicity. Furthermore, the receptor (TLR-3) and ligand (multiepitope vaccine) interactions were checked through molecular docking and molecular dynamics simulation. Ultimately, in silico cloning was done to ensure expression and translation efficiency.

ACKNOWLEDGMENT

R.K. Pandey is thankful to DST India for providing INSPIRE fellowship. R. Ojha is thankful to Central University of Rajasthan for providing UGC fellowship. V.K. Prajapati is thankful to the Central University of Rajasthan for providing computational facility.

CONFLICTS OF INTEREST

The Authors have declared no conflicts of interest.

AUTHOR CONTRIBUTIONS

Protocol was designed by R.K. Pandey, N. Upadhyay, and V.K. Prajapati. Methodology was performed by R.K. Pandey, N. Upadhyay. Manuscript was written by R.K. Pandey, R. Ojha, N. Chatterjee, N. Upadhyay, A. Mishra, and V.K. Prajapati.

ORCID

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

REFERENCES

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. *Scientific Reports*, *7*(1), 9232. https://doi.org/10.1038/s41598-017-09199-w

Boxus, M., Twizere, J. -C., Legros, S., Dewulf, J. -F., Kettmann, R., & Willems, L. (2008). The HTLV-1 tax interactome. *Retrovirol.*, *5*(1), 76.

Chatterjee, N., Ojha, R., Khatoon, N., & Prajapati, V. K. (2018). Scrutinizing *Mycobacterium tuberculosis* membrane and secretory proteins to formulate multiepitope subunit vaccine against pulmonary tuberculosis by utilizing immunoinformatic approaches. *International Journal of Biological Macromolecules*, *118*, 180–188.

Dimitrov, I., Flower, D. R., & Doytchinova, I. (2013). AllerTOP—A server for in silico prediction of allergens. *BMC Bioinformatics*, *14*(Suppl 6), S4. https://doi.org/10.1186/1471-2105-14-S6-S4

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*(1), 4. https://doi.org/10.1186/1471-2105-8-4

Furnham, N., Doré, A. S., Chirgadze, D. Y., de Bakker, P. I. W., Depristo, M. A., & Blundell, T. L. (2006). Knowledge-based real-space explorations for low-resolution structure determination. *Structure*, *14*(8), 1313–1320. https://doi.org/10.1016/j.str.2006.06.014

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server, *The proteomics protocols handbook* (571–607) Springer.

Gessain, A., & Cassar, O. (2012). Epidemiological aspects and world distribution of HTLV-1 infection. *Frontiers in Microbiology*, *3*, 388.

Greten, T. F., Slansky, J. E., Kubota, R., Soldan, S. S., Jaffee, E. M., Leist, T. P., … Schneck, J. P. (1998). Direct visualization of antigen-specific

T cells: HTLV-1 Tax11–19-specific CD8+ T cells are activated in peripheral blood and accumulate in cerebrospinal fluid from HAM/TSP patients. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7568–7573.

Grote, A., Hiller, K., Scheer, M., Munch, R., Nortemann, B., Hempel, D. C., & Jahn, D. (2005). JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research*, *33*(Web Server issue), W526–W531. https://doi.org/10.1093/nar/gki376

Heo, L., Park, H., & Seok, C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Research*, *41*(Web Server issue), W384–W388. https://doi.org/10.1093/nar/gkt458

Hess, B., Kutzner, C., Van Der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, *4*(3), 435–447.

Jones, K. S., Fugo, K., Petrow-Sadowski, C., Huang, Y., Bertolette, D. C., Lisinski, I., … Ruscetti, F. W. (2006). Human T-cell leukemia virus type 1 (HTLV-1) and HTLV-2 use different receptor complexes to enter T cells. *Journal of Virology*, *80*(17), 8291–8302. https://doi.org/10.1128/jvi.00389-06

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nature Protocols*, *7*(8), 1511–1522.

Khabbaz, R. F., Fukuda, K., Kaplan, J. E., Bianco, C., Blattner, W. A., Busch, M., … Jackson, C. (1993). Centers for Disease Control and Prevention and US Public Health Service Working Group for recommendations for counseling persons infected with human T-lymphotrophic virus, types I and II. *Morbidity and Mortality Weekly Report: Recommendations and Reports*, *42*, 1–13.

Khatoon, N., Pandey, R. K., & Prajapati, V. K. (2017). Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Scientific Reports*, *7*(1), 8285.

Ko, J., Park, H., Heo, L., & Seok, C. (2012). GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Research*, *40*(W1), W294–W297.

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., … Vajda, S. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols*, *12*(2), 255–278. https://doi.org/10.1038/nprot.2016.169

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 Bioinformatics*, *8*, 424. https://doi.org/10.1186/1471-2105-8-424

Le Blanc, I., Grange, M. P., Delamarre, L., Rosenberg, A. R., Blot, V., Pique, C., & Dokhélar, M. C. (2001). HTLV-1 structural proteins. *Virus Research*, *78*(1-2), 5–16.

Mohan, T., Sharma, C., Bhat, A. A., & Rao, D. N. (2013). Modulation of HIV peptide antigen specific cellular immune response by synthetic α-and β-defensin peptides. *Vaccine*, *31*(13), 1707–1716.

Ojha, R., Khatoon, N., & Prajapati, V. K. (2018). Conglomeration of novel *Culex quinquefasciatus* salivary proteins to contrive multi-epitope subunit vaccine against infections caused by blood imbibing transmitter. *International Journal of Biological Macromolecules*, *118*, 834–843. https://doi.org/10.1016/j.ijbiomac.2018.06.112

Pandey, R. K., Ali, M., Ojha, R., Bhatt, T. K., & Prajapati, V. K. (2018). Development of multi-epitope driven subunit vaccine in secretory and membrane protein of *Plasmodium falciparum* to convey protection against malaria infection. *Vaccine*, *36*(30), 4555–4565. https://doi.org/10.1016/j.vaccine.2018.05.082

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. *Scientific Reports*, *8*(1), 1125.

Pandey, R. K., Kumbhar, B. V., Sundar, S., Kunwar, A., & Prajapati, V. K. (2017). Structure-based virtual screening, molecular docking, ADMET and molecular simulations to develop benzoxaborole analogs as potential inhibitor against *Leishmania donovani* trypanothione reductase. *Journal of Receptor and Signal Transduction Research*, *37*(1), 60–70. https://doi.org/10.3109/10799893.2016.1171344

Pandey, R. K., Ojha, R., Aathmanathan, V. S., Krishnan, M., & Prajapati, V. K. (2018). Immunoinformatics approaches to design a novel multi-epitope subunit vaccine against HIV infection. *Vaccine*, *36*(17), 2262–2272.

Pandey, R. K., Ojha, R., Mishra, A., & Prajapati, V. K. (2018). Designing B- and T-cell multi-epitope based subunit vaccine using immunoinformatics approach to control Zika virus infection. *Journal of Cellular Biochemistry*, https://doi.org/10.1002/jcb.27110

Pandey, R. K., & Prajapati, V. K. (2018). Exploring sand fly salivary proteins to design multi-epitope subunit vaccine to fight against visceral leishmaniasis. *Journal of Cellular Biochemistry*, https://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. *Journal of Biomolecular Structure and Dynamics*, *33*(12), 2541–2553. https://doi.org/10.1080/07391102.2015.1085904

Proietti, F. A., Carneiro-Proietti, A. B. F., Catalan-Soares, B. C., & Murphy, E. L. (2005). Global epidemiology of HTLV-1 infection and associated diseases. *Oncogene*, *24*(39), 6058–6068.

Rana, A., & Akhter, Y. (2016). A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. *Immunobiology*, *221*(4), 544–557. https://doi.org/10.1016/j.imbio.2015.12.004

Saha, S., & Raghava, G. P. S. (2006). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins*, *65*(1), 40–48. https://doi.org/10.1002/prot.21078

Sette, A., & Sidney, J. (1999). Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics*, *50*(3-4), 201–212.

Shukla, H., Shukla, R., Sonkar, A., Pandey, T., & Tripathi, T. (2017). Distant Phe345 mutation compromises the stability and activity of *Mycobacterium tuberculosis* isocitrate lyase by modulating its structural flexibility. *Scientific Reports*, *7*(1), 1058. https://doi.org/10.1038/s41598-017-01235-z

Shukla, H., Shukla, R., Sonkar, A., & Tripathi, T. (2017). Alterations in conformational topology and interaction dynamics caused by L418A mutation leads to activity loss of *Mycobacterium tuberculosis* isocitrate lyase. *Biochemical and Biophysical Research Communications*, *490*(2), 276–282. https://doi.org/10.1016/j.bbrc.2017.06.036

Shukla, R., Shukla, H., Sonkar, A., Pandey, T., & Tripathi, T. (2018). Structure-based screening and molecular dynamics simulations offer novel natural compounds as potential inhibitors of *Mycobacterium tuberculosis* isocitrate lyase. *Journal of Biomolecular Structure and Dynamics*, *36*(8), 2045–2057. https://doi.org/10.1080/07391102.2017.1341337

Wang, P., Sidney, J., Dow, C., Mothé, B., Sette, A., & Peters, B. (2008). A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Computational Biology*, *4*(4), e1000048. https://doi.org/10.1371/journal.pcbi.1000048

Wolfe, N. D., Heneine, W., Carr, J. K., Garcia, A. D., Shanmugam, V., Tamoufe, U., … Switzer, W. M. (2005). Emergence of unique primate

T-lymphotropic viruses among central African bushmeat hunters. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(22), 7994–7999.

Yamauchi, J., Coler-Reilly, A., Sato, T., Araya, N., Yagishita, N., Ando, H., … Yamano, Y. (2014). Mogamulizumab, an anti-CCR4 antibody, targets human T-lymphotropic virus type 1–infected CD8+ and CD4+ T cells to treat associated myelopathy. *Journal of Infectious Diseases*, *211*(2), 238–248.

Yang, Y., Sun, W., Guo, J., Zhao, G., Sun, S., Yu, H., … Zhou, Y. (2015). In silico design of a DNA-based HIV-1 multi-epitope vaccine for Chinese populations. *Human Vaccines & Immunotherapeutics*, *11*(3), 795–805. https://doi.org/10.1080/21645515.2015.1012017

Yasunaga, J.-I., & Matsuoka, M. (2007). Human T-cell leukemia virus type I induces adult T-cell leukemia: From clinical aspects to molecular mechanisms. *Cancer Control*, *14*(2), 133–140.

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, Chatterjee N, Upadhyay N, Mishra A, Prajapati VK. Combinatorial screening algorithm to engineer multiepitope subunit vaccine targeting human T-lymphotropic virus-1 infection. *J Cell Physiol*. 2019;234:8717–8726. https://doi.org/10.1002/jcp.27531