Microbial Pathogenesis 137 (2019) 103791

Contents lists available at ScienceDirect

Microbial Pathogenesis

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

Exploring T & B-cell epitopes and designing multi-epitope subunit vaccine targeting integration step of HIV-1 lifecycle using immunoinformatics approach

Faruq Abdulla[a], Utpal Kumar Adhikari[b], M. Kamal Uddin[a,*]

[a] Department of Statistics, Islamic University, Kushtia-7003, Bangladesh
[b] School of Medicine, Western Sydney University, Campbelltown, NSW-2560, Australia

ARTICLE INFO

Keywords:
HIV-1
Integrase protein
Conserved domain
T-cell epitopes
B-cell epitopes
Subunit vaccine

ABSTRACT

Till now, AIDS, caused by the human immunodeficiency virus (HIV) is still a severe health problem worldwide. It weakens the immune system by targeting the T-helper cells. Specifically, the severity of the pandemic HIV-1 makes the emergence of an enduring effective vaccine against HIV-1. Therefore, we have applied a series of immunoinformatics approaches to the four conserved domains of HIV-1 integrase (IN) proteins to design an effective multi-epitope based subunit vaccine which might induce a competent immunity against HIV-1. Therefore, we have selected three peptide fragments that contained all overlapping epitopes (35 $CD4^+$, 8 $CD8^+$ T-cell epitopes, and 3 B-cell epitopes) where the epitopes had a high conservancy score. The cumulative population coverage for combined $CD8^+$ and $CD4^+$ T-cell epitopes and their respective HLA-alleles were found as 98.03% in the world which is also followed by East Asia (96.24%), South Asia (96.31%), North Africa (96.14%), North America (98.99%), and Europe (98.80%). The proposed vaccine composed by an adjuvant (β-defensin) at the N-terminal site of the vaccine constructs and three peptide fragments where the adjuvant was fused by EAAAK linker and the peptide fragments were fused by GPGPG linker. The designed final vaccine construct (length: 159 amino acid) was found to be antigenic and non-allergic, which indicates its safety. The vaccine construct was found as good antigenic, stable, higher thermostable, and hydrophilic in nature. The codon adaptation and *in silico* cloning ensured the high expression rate of the vaccine constructs in *E. coli* K12 with CAI value of 1.0. Finally, the binding affinity of the vaccine constructs with the immune receptor TLR3 was confirmed by the lowest energy score of −1026.8 evaluated by molecular docking. However, the proposed *in silico* vaccine construct needs experimental validation for assuring the safety and immunogenicity profile which will ensure an active immunity against HIV-1.

1. Introduction

Human immunodeficiency virus (HIV) has been recognized as a causative agent of acquired immune deficiency syndrome (AIDS) and remains an epidemic threat to the global public health. Hitherto, around 76.1 million people have been infected, 35 million people have died, and 36.7 million people are living with HIV. HIV is most severe in the sub-Saharan African countries where one in 25 individuals is infected with this virus [1]. Among two different HIV types, such as HIV-1 and HIV-2, HIV-1 is the major pandemic and pathogenic agent of AIDS. The HIV genome consists of three structural (*Env*, *Gag*, and *Pol*) and six regulatory (*Tat*, *Rev*, *Nef*, *Vif*, *Vpr*, and *Vpu*) genes and these nine genes encode a total of 16 viral proteins [2]. Successful HIV replication requires three enzymes encoded by structural genes and one of them integrase integrates the viral DNA into the host DNA. Therefore, it is a rational target for the discovery of a vaccine against HIV [3].

The most recent treatment for AIDS/HIV involves highly active antiretroviral therapy (HAART). The HAART can extend the lifespan and quality of life by slowing the progression of HIV. It can reduce the mortality from AIDS by 60%, but does not eradicate the virus in the infected individuals and it is also very expensive [4]. This severe situation necessitates the need for an effective, safe and cost effective preventive vaccine, as vaccination is the most effective preventive administration against infectious diseases. Due to years of clinical research, only three of the major HIV vaccines, such as AIDSVAX, MRKAd5, and RV144 clinical trials have been successful. Under three, AIDSVAX can induce neutralizing antibody activity to elicit only the homologous viruses, MRKAd5 can only induce cytotoxic T-cell

* Corresponding author.
E-mail address: drkamal@stat.iu.ac.bd (M.K. Uddin).

https://doi.org/10.1016/j.micpath.2019.103791
Received 21 January 2019; Received in revised form 8 October 2019; Accepted 8 October 2019
Available online 10 October 2019
0882-4010/

responses, but the viral load after HIV infection could not be reduced, and RV144 can induce both $CD8^+$ T-cell and neutralizing antibody responses, but cannot induce the $CD4^+$ T-cell responses and unable to reduce viral load after HIV infection [5]. So, the lack of an effective vaccine prompted us to rethink about designing an effective vaccine against HIV. In modern times, the advanced bioinformatics approach called immunoinformatics has rapidly been used in the field of vaccinology with less expensive and saving-time [6]. In recent years, the multi-epitope subunit vaccine has received more attention worldwide [7]. Moreover, the several advantages such as simplicity of production, specificity of the target, safety, success in preclinical models, and a high degree of accuracy increase the novelty of the multi-epitope subunit vaccine [8]. Pandey *et* al. (2018) designed a multi-epitope subunit vaccine that used both antigenic and non-antigenic epitopes against HIV, using conserved regions of the structural and regulatory proteins as input to the immunoinformatics approach [9]. Therefore, our study harnessed the immunoinformatics tool on the conserved domains of the integrase protein of HIV-1 for designing an efficient multi-epitope subunit vaccine against HIV-1 using potential antigenic T-cell ($CD8^+$ & $CD4^+$) and B-cell epitopes. The designed vaccine contained a suitable adjuvant and subsequently the peptide fragments containing the immunogenic cellular and humoral immune response specific epitopes along with suitable linkers. The vaccine epitopes were analyzed for the antigenicity and conservancy and followed by binding affinity with their corresponding HLA-alleles, and the population coverage. The antigenicity and allergenicity were predicted of the vaccine construct and followed by the primary and secondary structural features determination. The prediction of the tertiary structure of the designed vaccine was performed and then the structure was refined and validated. Disulfide engineering was performed to increase the stability of the vaccine construct, and the expression rate of the designed vaccine in *E. coli* K12 was checked by performing codon adaptation and *in silico* cloning. The vaccine-receptor complex was predicted and followed by its binding affinity prediction through molecular docking.

2. Methods and materials

A flow chart representing the whole working procedure of epitope-based vaccine target identification and design of a vaccine construct against HIV-1 has been illustrated in Fig. 1.

2.1. Retrieving of HIV-1 IN protein sequences and identification of conserved domains

The IN protein sequences of HIV-1 were retrieved from the RCSB PDB (https://www.rcsb.org/). The retrieved protein sequences were checked for their HIV-1 IN property through BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) and Pfam (https://pfam.xfam.org/). The conserved domain analysis helps us to identify the essential functional parts of the target protein sequences. The NCBI CDD-BLAST (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) was used to search the conserved domains.

2.2. Prediction of T-cell ($CD8^+$ & $CD4^+$) epitopes and their MHC binding HLA-alleles

The T-cell epitopes are nothing but the small peptide fragment obeying the rules to control and dominate the immune system for developing a precise immune response. So, the reliable prediction of the T-cell epitope is one of the most crucial characteristics of epitope-based peptide vaccine design. The $CD8^+$ T-cell epitopes were portended using the NetCTL v1.2 server (http://www.cbs.dtu.dk/services/NetCTL/) with the threshold parameter 0.5, and the remaining parameters were fixed as default. NetCTL can predict the $CD8^+$ T-cell epitopes restricted to 12 MHC class I supertypes [10], but the specific MHC class I binding HLA alleles prediction for the epitope is not available in NetCTL server. So, the MHC class I binding HLA alleles corresponding to $CD8^+$ T-cell epitopes were predicted using IEDB MHC class I binding tool (http://tools.iedb.org/mhci/), where human and ANN (Artificial Neural Network) were selected as MHC source species and prediction method, respectively [11]. However, the cutoff IC_{50} value was set at 500 nM. The $CD4^+$ T-cell epitopes and their corresponding MHC class II alleles were predicted through IEDB MHC II binding tool (http://tools.iedb.org/mhcii/) using the NN-align method, and the threshold IC_{50} value was set at 500 nM [12].

2.3. Immunological properties and conservancy analysis

Immunological properties such as antigenicity and immunogenicity are important factors for vaccine design. In this study, the antigenicity was predicted using VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) with the threshold set at 0.4 and immunogenicity was observed using IEDB MHC I Immunogenicity tool (http://tools.iedb.org/immunogenicity/) [13,14]. The conservancy analysis of an epitope is essential in an epitope-based vaccine design because of their broader protection athwart to several strains of a species or even several species. So, conservancy of the $CD8^+$ and $CD4^+$ T-cell epitopes was predicted through the IEDB conservation across antigens tool (http://tools.iedb.org/conservancy/) with default parameters [15]. The allergenicity of the designed vaccine was predicted using the AlgPred server (http://crdd.osdd.net/raghava/algpred/index.html) [16]. The allergenicity of the vaccine construct was cross-checked through AllerTOP v2.0 (http://www.ddg-pharmfac.net/AllerTOP/) [17] and AllergenFP v1.0 (http://www.ddg-pharmfac.net/AllergenFP/) [18].

2.4. Population coverage prediction for $CD8^+$ and $CD4^+$ T-cell epitopes and their alleles

The T-cell epitope-based vaccine should be designed by maximizing the population coverage by the minimum number of T-cell epitopes with minimum variability in different ethnic groups. So, the prediction of the population coverage is an important factor for epitope-based vaccine design. In this study, IEDB population coverage tool (http://tools.iedb.org/population/) was introduced to predict the population coverage for different ethnic groups using the finally predicted $CD8^+$ and $CD4^+$ T-cell epitopes and their corresponding MHC restricted alleles [19].

2.5. Molecular docking simulation of the HLA allele-peptide interaction

The interaction performance between predicted epitopes and their binding alleles were evaluated through molecular docking using the AutoDock tools and AutoDockVina software [20,21]. To do this, the crystal structure of the most frequent alleles: HLA-A*02:06 (PDB ID: 3OXR), HLA-A*03:01 (PDB ID: 3RL1), HLA-C*07:02 (PDB ID: 5VGE), HLA-C*12:03 (PDB ID: 1EFX), HLA-DRB1*01:01 (PDB ID: 2FSE), HLA-DRB1*15:01 (PDB ID: 1YMM), HLA-DRB1*03:01 (PDB ID: 1A6A), and HLA-DRB1*04:01 (PDB ID: 5JLZ) were retrieved from protein data bank (PDB) as PDB format [22–29]. Since the retrieved crystal structures were in the complex form of protein and ligand, hence the Discovery Studio v16.0.0.400 was used to separate the protein and ligand from the complex structures. The separated protein and ligand were then converted as PDBQT format using AutoDock tools. Finally, the docking simulation was done through AutoDockVina software, and the obtained results were used as a benchmark for comparing the docking results of the predicted epitopes using similar parameters. The 3D modelling of the predicted epitopes were performed using PEPFOLD v3.5 [30]. The conversion of the resulted PDBQT format files to PDB format was done using Open Babel v2.4.1 and then visualized by PyMOL software [31].

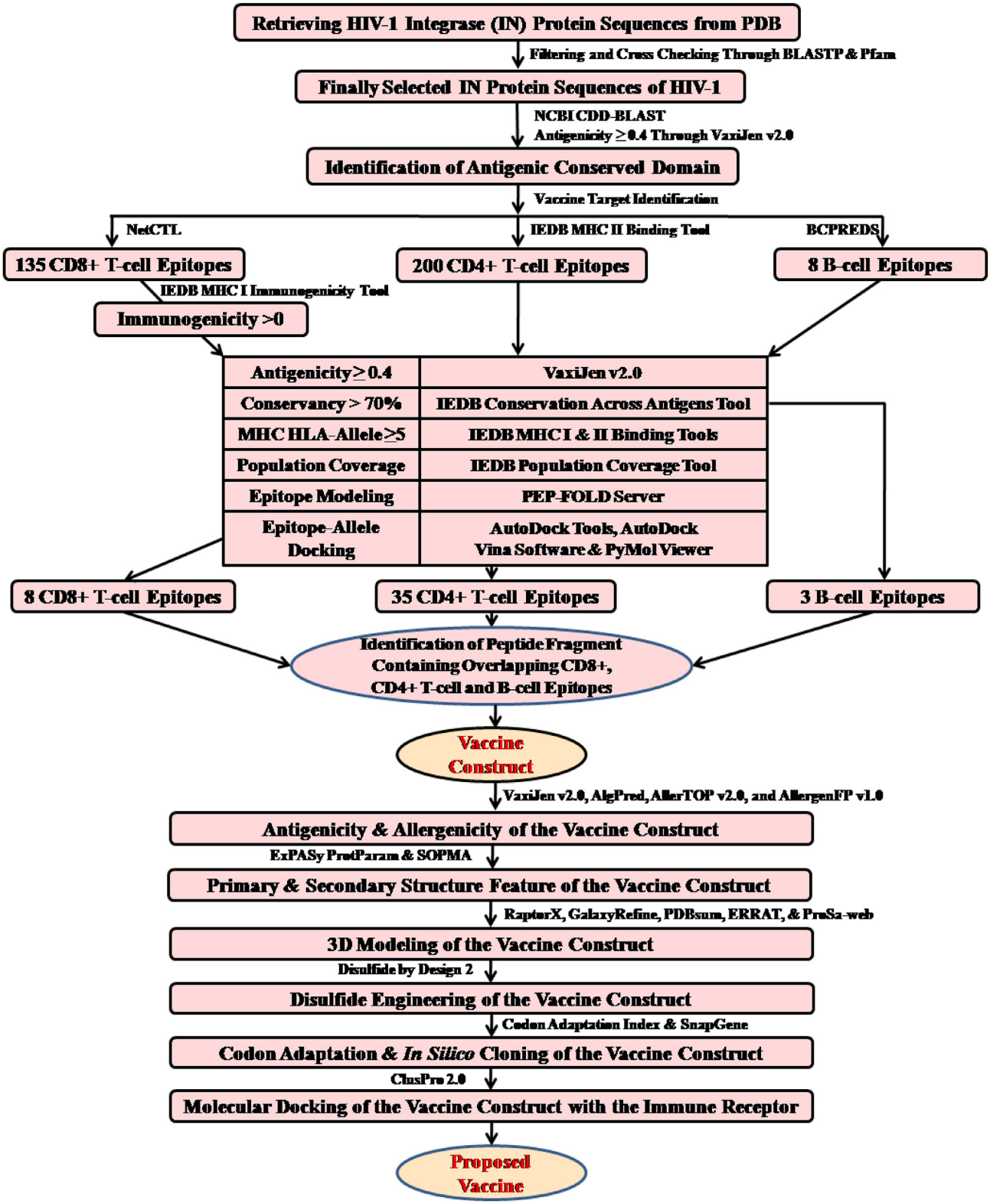

Fig. 1. Working-flowchart of the multi-epitope subunit vaccine construct against HIV-1.

2.6. B-cell epitopes prediction

The B-cell epitopes are the antigenic parts of the protein sequences that are bound to the receptors of the B-cells for stimulating the B-cell response. The B-cell epitope prediction is an essential characteristic of a multi-epitope subunit vaccine because an ideal multi-epitope subunit vaccine contains T & B-cell epitopes. The linear B-cell epitopes were predicted through the BCPREDS server (http://ailab.ist.psu.edu/bcpred/predict.html). Herein, we used the BCPred method with the default length of 20 residues and 90% specificity threshold value [32].

2.7. Construction of multi-epitope subunit vaccine

An efficient multi-epitope subunit vaccine should contain a strong

immunostimulatory adjuvant and T_C, T_H, and B-cell epitopes along with suitable linkers [33]. An adjuvant is an immunological agent which is added to the vaccine in order to help in stimulating and enhancing the adaptive immune system. The defensins comprise a critical component that provides an innate host response against microbial infection and recruit naïve T-cell and immature dendritic cells at the site of infection through the CCR6 receptor to provide an adaptive immune response and exhibit anti-HIV-1 activities and immunomodulatory capabilities [34]. The β-defensin-2 was added as an adjuvant at the N-terminal of the vaccine construct by EAAAK linker. In the vaccine construct, the T_C, T_H, and B-cell epitopes enriched peptide fragments were fused by GPGPG linker [9].

2.8. Primary and secondary structural feature assessment

The physicochemical properties and secondary structural features are highly associated with the functionality of the protein sequences. The ExPASy ProtParam (https://web.expasy.org/protparam/) and SOPMA server (https://npsa-prabi.ibcp.fr/NPSA/npsa_sopma.html) were used with default parameters to compute various physicochemical properties and secondary structural features, respectively [35,36].

2.9. Tertiary structure prediction, refinement, and validation of the vaccine construct

The tertiary structure of the final vaccine construct was generated through a template-based protein modelling server RaptorX (http://raptorx.uchicago.edu/StructurePrediction/predict/) [37]. The quality of the predicted tertiary structure depends on the degree of the alikeness of the target and available template structure [38]. Hence, there was a necessity to improve the precision of the predicted model using refinement technique. The refinement of the predicted model was done through the GalaxyRefine server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE), which refine the tertiary structure by the process: rebuilding-side-chains-repacking as well as by relaxing the overall structure using molecular dynamic simulation [39]. Validation is an important step for structural modelling which verifies the accuracy of the predicted refined 3D model using Ramachandran plot analysis, different quality factors, and errors plot analysis. Herein, the PDBsum server (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html) was used to depict the Ramachandran maps [40]. We used ProSA-web server (https://prosa.services.came.sbg.ac.at/prosa.php) that provides the overall quality Z-score, X-ray and NMR plot, and local quality plot for the input structure [41]. The statistics of the non-bonded atom-atom interactions of the input structure was analyzed through ERRAT server (http://servicesn.mbi.ucla.edu/ERRAT/) [42].

2.10. Disulfide engineering of the designed vaccine construct

Disulfide bonds can provide increased stability of proteins by reducing unfolded conformational entropy and raising the free energy of the denatured state along with the examination of protein interactions and molecular dynamics [43]. Disulfide engineering is an advanced technique to design new disulfide bonds in the protein structure. Disulfide engineering of the predicted 3D model of vaccine construct was done by Disulfide by Design 2 (DbD2) v2.12 (http://cptweb.cpt.wayne.edu/DbD2/) with default parameters. DbD2 run with a 3D structure of the vaccine construct and predict pairs of residues that have the characteristics to form disulfide bond if the individual residues are mutated to cysteines [44].

2.11. Codon adaptation and in silico cloning of the vaccine construct

Codon adaptation is a process expressing the foreign genes in the host with a high expression rate when the codon usage in the host and foreign organism differ from each other. Codon optimization was done through Java Codon Adaptation Tool (JCat) (http://www.jcat.de/) for expressing the vaccine construct in *E. coli* K12 because whose expression system differ from the human expression system [9,45]. Expression rate parameters, the codon adaptation index (CAI), and percentage of GC-content codon should be in range of 0.8–1.0 and 30–70%, respectively, where CAI of 1.0 is considered as best score [38]. In JCat, additional three options were checked to avoid rho-independent transcription terminators, prokaryotic ribosome binding sites, and cleavage sites of restriction enzymes. In the optimized nucleotide sequence, the absence of the commercially available restriction enzyme cutting sites (XhoI and BamHI) was confirmed. Afterward, the restriction cloning module of SnapGene tool was used to clone the optimized adapted codon sequence corresponding to the vaccine construct into the *E. coli* pET28a(+) vector [38].

2.12. Molecular docking of final vaccine construct with immune receptor TLR3

It is necessary to check the binding affinity of the vaccine construct with the immune receptor for inducing an appropriate immune response by the vaccine construct. Molecular docking is a computational technique to predict the binding affinity of the vaccine construct with an immune receptor as well as their interacted complex form [46]. Herein, the toll-like receptor 3 (TLR3) was used as the immune receptor because the HIV-1 mediated activation of dendritic cells is dependent on TLR3 [9]. Molecular docking of vaccine constructs with TLR3 (PDB ID: 2A0Z) was performed through automated online server ClusPro 2.0 (https://cluspro.bu.edu/login.php) [47].

3. Results

3.1. Retrieving HIV-1 IN protein sequences and identification of conserved domains

The PDB database was used to obtain our targeted HIV-1 IN Proteins. After searching with parameters HIV-1 and IN protein, a total of 356 protein chains were found, but after rigorous screening of each individual chain, 127 chains were found to be incompatible with our objectives and discarded, and the remaining 229 chains were assembled in a single FASTA file. The presence of those protein chains in the HIV-1 and IN type was confirmed by cross-checking through BLASTP and Pfam. Among those 229 chains, there were 198 identical chains which were discarded, and only 31 chains were selected as final experimental data set for conducting this study. The CDD-BLAST was used to search the conserved domains, and the searching results were enlisted in Table S1. The predicted result showed that the HIV-1 IN proteins contained four conserved domains which are rve, IN_DBD_C, Integrase_Zn, and Tra5; all of them mediates the integration of a DNA copy of the viral genome to the host chromosome. The antigenicity analysis showed that the identified conserved domains are antigenic with the predicted antigenicity score of 0.61 (rve), 0.56 (IN_DBD_C), 0.45 (Integrase_Zn), and 0.55 (Tra5). The length of the selected rve, IN_DBD_C, Integrase_Zn, and Tra5 conserved domains are 102, 48, 38, and 88, respectively. Finally, the identified conserved domains were used as input data for our immunoinformatics study.

3.2. Prediction and assessment of $CD8^+$ & $CD4^+$ T-cell epitopes and their MHC binding HLA-alleles

The identified four conserved domains, rve, IN_DBD_C, Integrase_Zn, and Tra5 were subjected to NetCTL v1.2 tool, and initially we found a total of 135 $CD8^+$ T-cell epitopes (51 in rve, 29 in IN_DBD_C, 15 in Integrase_Zn, and 40 in Tra5) and after discarding the identical epitopes with slighter combined scores, we attained 79 epitopes (28 in rve, 17 in IN_DBD_C, 10 in Integrase_Zn, and 24 in Tra5)

Table 1
$CD8^+$, $CD4^+$ T-cell epitopes and B-cell epitopes enriched peptide fragments.

Domain	$CD8^+$ T-cell epitope enriched peptide fragment	No. of $CD8^+$ T-cell epitopes	$CD4^+$ T-cell epitope enriched peptide fragment	No. of $CD4^+$ T-cell epitopes	B-cell epitope enriched peptide fragment	No. of B-cell epitopes
rve	^{18}VAVHVASGYIEAEVIPA^{34}	3	^{16}ILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVK^{54}	16	^{18}VAVHVASGYIEAEVIPAETG^{37}	1
	^{37}GQETAYFLL^{45}	1				
	^{46}KLAGRWPVK^{54}	1				
IN_DBD_C	^{20}KLLWKGEGA^{28}	1	^{17}GPAKLLWKGEGAVVIQDNSDIKVVPRRKAKII^{48}	9	^{29}VVIQDNSDIKVVPRRKAKII^{48}	1
Integrase_Zn	^{7}KYHSNWRAM^{15}	1	^{5}HEKYHSNWRAMASDFNLPPVVAKE^{28}	10	^{11}NWRAMASDFNLPPVVAKEIV^{30}	1
	^{16}ASDFNLPPV^{24}	1				

Table 2
Peptide fragments containing selected all overlapping $CD8^+$, $CD4^+$ T-cell epitopes and B-cell epitopes.

Domain	Peptide Fragment	$CD8^+$ T-cell epitopes	$CD4^+$ T-cell epitopes	B-cell epitopes
rve	^{16}ILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVK^{54}	^{18}VAVHVASGY^{26} ^{21}HVASGYIEA^{29} ^{26}YIEAEVIPA^{34} ^{37}GQETAYFLL^{45} ^{46}KLAGRWPVK^{54}	^{16}ILVAVHVASGYIEAE^{30} ^{17}LVAVHVASGYIEAEV^{31} ^{18}VAVHVASGYIEAEVI^{32} ^{19}AVHVASGYIEAEVIP^{33} ^{20}VHVASGYIEAEVIPA^{34} ^{21}HVASGYIEAEVIPAE^{35} ^{22}VASGYIEAEVIPAET^{36} ^{23}ASGYIEAEVIPAETG^{37} ^{24}SGYIEAEVIPAETGQ^{38} ^{25}GYIEAEVIPAETGQE^{39} ^{33}PAETGQETAYFLLKL^{47} ^{34}AETGQETAYFLLKLA^{48} ^{35}ETGQETAYFLLKLAG^{49} ^{36}TGQETAYFLLKLAGR^{50} ^{37}GQETAYFLLKLAGRW^{51} ^{40}TAYFLLKLAGRWPVK^{54}	^{18}VAVHVASGYIEAEVIPAETG^{37}
IN_DBD_C	^{17}GPAKLLWKGEGAVVIQDNSDIKVVPRRKAKII^{48}	^{20}KLLWKGEGA^{28}	^{17}GPAKLLWKGEGAVVI^{31} ^{18}PAKLLWKGEGAVVIQ^{32} ^{19}AKLLWKGEGAVVIQD^{33} ^{20}KLLWKGEGAVVIQDN^{34} ^{23}WKGEGAVVIQDNSDI^{37} ^{24}KGEGAVVIQDNSDIK^{38} ^{25}GEGAVVIQDNSDIKV^{39} ^{33}DNSDIKVVPRRKAKI^{47} ^{34}NSDIKVVPRRKAKII^{48}	^{29}VVIQDNSDIKVVPRRKAKII^{48}
Integrase_Zn	^{5}HEKYHSNWRAMASDFNLPPVVAKEIV^{30}	^{7}KYHSNWRAM^{15} ^{16}ASDFNLPPV^{24}	^{5}HEKYHSNWRAMASDF^{19} ^{6}EKYHSNWRAMASDFN^{20} ^{7}KYHSNWRAMASDFNL^{21} ^{8}YHSNWRAMASDFNLP^{22} ^{9}HSNWRAMASDFNLPP^{23} ^{10}SNWRAMASDFNLPPV^{24} ^{11}NWRAMASDFNLPPVV^{25} ^{12}WRAMASDFNLPPVVA^{26} ^{13}RAMASDFNLPPVVAK^{27} ^{14}AMASDFNLPPVVAKE^{28}	^{11}NWRAMASDFNLPPVVAKEIV^{30}

(Table S2). Thereafter, these 79 epitopes were taken under consideration for the prediction of the antigenicity, immunogenicity, conservancy, and MHC-I binding HLA-alleles (Table S2). Afterward, only eight $CD8^+$ T-cell epitopes (5 in rve, 1 in IN_DBD_C, and 2 in Integrase_Zn) were selected as potential vaccine candidates according to the antigenicity value of ≥ 0.4, positive immunogenicity, $> 70\%$ conservancy, and ≥ 5 MHC-I HLA-alleles (Table 2). Interestingly, the epitope ^{18}VAVHVASGY^{26} has 100% conservancy and ^{21}HVASGYIEA^{29}, ^{26}YIEAEVIPA^{34}, and ^{46}KLAGRWPVK^{54} epitopes have $> 99\%$ conservancy score. However, the predicted $CD8^+$ T-cell epitopes from Tra5 conserved domain did not fulfill our criteria, so we did not use any epitopes for the final vaccine construct design. Again the identified conserved domain sequences were subjected to IEDB MHC-II binding tool, and initially, we identified 200 $CD4^+$ T-cell epitopes (78 in rve, 33 in IN_DBD_C, 24 in Integrase_Zn, and 65 in Tra5) (Table S3). Later on, the antigenicity and conservancy of the identified $CD4^+$ T-cell epitopes were identified and shown in Table S3. Now under the conditions: antigenicity score ≥ 0.4, $> 70\%$ conservancy, and ≥ 5 MHC-II HLA-alleles, only 35 $CD4^+$ T-cell epitopes (16 in rve, 9 in IN_DBD_C, 10 in Integrase_Zn) were considered as potential $CD4^+$ T-cell epitopes for vaccine construction (Table 2). Excitingly, the eight $CD4^+$ T-cell epitopes have $> 99\%$ conservancy score. None of the predicted $CD4^+$ epitopes from Tra5 conserved domain fulfilled our expected criteria.

3.3. Population distribution analysis of the predicted MHC-I & MHC-II HLA-alleles

Since the different MHC-I and MHC-II HLA alleles are exposed at different frequency (population/individual) level in different ethnicities in the world, hence the frequency distribution analysis or the analysis of how much individuals will be covered by the respective HLA alleles (MHC-I & MHC-II) of the predicted T-cell epitopes is a vital part of designing an effective vaccine. A separate, as well as a combined population coverage analysis, was performed for the finally selected $CD8^+$ & $CD4^+$ T-cell epitopes and their respective HLA alleles, and the analysis results are shown in Table S4 after sorting the population coverage

Table 3
The cumulative population coverage (%) in 16 largest geographical areas and in the whole world.

Population/area	Population Coverage (%)		
	$CD8^+$ T-cell Epitopes and Their Respective HLA-Alleles	$CD4^+$ T-cell Epitopes and Their Respective HLA-Alleles	Both $CD8^+$ & $CD4^+$ T-cell Epitopes and Their Respective HLA-Alleles
East Asia	81.91	79.21	96.24
Northeast Asia	88.67	56.89	95.12
South Asia	85.37	74.78	96.31
Southeast Asia	87.49	56.13	94.51
Southwest Asia	80.18	44.42	88.98
East Africa	81.17	62.42	92.92
West Africa	81.83	64.34	93.52
Central Africa	79.09	58.20	91.26
North Africa	84.44	75.20	96.14
South Africa	81.91	32.10	87.72
North America	91.02	88.72	98.99
Central America	4.14	53.39	55.32
South America	81.09	62.48	92.91
Australia	78.39	25.91	84.03
Europe	91.50	85.87	98.80
Oceania	72.53	53.30	87.17
World	89.11	81.94	98.03

in different countries, areas, and ethnic groups. The population coverage analysis results showed excellent coverage for many different populations/area including ≥90% coverage in 131 countries, areas, and ethnic groups. However, the combined analysis exhibited the maximum coverage of population in Ireland Northern (99.81%), closely followed by Ireland South (99.68%), England (99.60%), Germany (99.45%), Austria (99.25%), France (99.21%), Sweden (99.08%), United States (99.02%), North America (98.99%), Finland (98.87%), Europe (98.80%), Italy and Poland (98.75%), Czech Republic (98.73%), World (98.03%), and so on (Table S4). Herein, we mentioned 16 different geographical areas capturing all about the world that have a great percentage of cumulative population coverage (Table 3).

3.4. Analysis of the molecular docking simulation of the HLA allele-peptide interaction

To check the binding affinity of the $CD8^+$ T-cell vaccine candidates to the binding grooves of their corresponding allele proteins, the HLA-A*02:06 was used as it is common frequent in five epitopes or peptides (HVASGYIEA, YIEAEVIPA, GQETAYFLL, KLLWKGEGA, and ASDFNLPPV) and HLA-A*03:01, HLA-C*07:02, and HLA-C*12:03 were used corresponding to KLAGRWPVK, KYHSNWRAM, and VAVHVASGY, respectively. The peptide-allele interaction affinities were predicted by AutoDockVina, and the results exhibited good binding affinity compared to the control scores (Table S5). Specifically, the epitopes KLLWKGEGA, KLAGRWPVK, and KYHSNWRAM have binding scores of −7.5, −7.0, and −8.0, which are very close to their respective control scores of −7.6, −6.8, and −8.1, respectively. The peptide-allele interaction complexes were illustrated in Fig. 2. Conversely, the binding affinity of 31 $CD4^+$ T-cell vaccine candidates were checked through docking analysis of these epitopes with their common frequent allele HLA-DRB1*01:01. The HLA-DRB1*15:01 was used to check the binding affinity of two epitopes AVHVASGYIEAEVIP and DNSDIKVVPRRKAKI. The HLA-DRB1*03:01 and HLA-DRB1*04:01 were used to check the binding affinity of KGEGAVVIQDNSDIK and NWRAMASDFNLPPVV, respectively. The binding energy results confirmed their good binding affinity compared to the control scores (Table S5). The epitope EKYHSNWRAMASDFN has binding energy same as the control score of −6.6, epitopes HVASGYIEAEVIPAE, ASGYIEAEVIPAETG, SGYIEAEVIPAETGQ, and TAYFLLKLAGRWPVK have binding energy of −6.7 and epitopes PAETGQETAYFLLKL and KYHSNWRAMASDFNL have score of −6.5 where their control score was of −6.6. The interaction complexes were illustrated in Fig. 3.

3.5. Assessment of the predicted B-cell epitopes

The HIV-1 IN antigenic conserved domains were subjected to the BCPREDS server to predict B-cell epitopes, and we found a total of 8 epitopes (2 in rve, 2 in IN_DBD_C, 1 in Integrase_Zn, and 3 in Tra5), where three epitopes have score 1.0, and other five epitopes have score close to 1.0 (Table S6). The antigenicity and conservancy of those predicted B-cell epitopes were predicted (Table S6), and under the restrictions: antigenicity value ≥ 0.4 and conservancy > 70%, we have selected only three epitopes as potential vaccine candidates (1 in rve, 1 in IN_DBD_C, and 1 in Integrase_Zn) (Table 3) where the epitope ^{18}VAVHVASGYIEAEVIPAETG^{37} has > 99% conservancy score. The predicted B-cell epitopes from Tra5 conserved domain could not attain our conditions.

3.6. Construction of multi-epitope subunit vaccine

The finally selected 46 robust vaccine candidates are overlapped with each other. Hence, the peptide fragments containing overlapping epitopes were identified separately for $CD8^+$, $CD4^+$, and B-cell epitopes (Table 1). Later on, for each domain, a peptide fragment containing it's all overlapping epitopes was identified (Table 2). Finally, the peptide fragments were linked with the help of GPGPG linker and the β-defensin-2 of length 37 was added as an adjuvant at the N-terminal site of the fused chain of length 107 with the help of EAAAK linker. Finally, the both sites of the construction were enclosed by EAAAK linker and a multi-epitope subunit vaccine model of length 159 was constructed. A schematic diagram of the multi-epitope vaccine construct is shown in Fig. 4.

3.7. Immunological assessment of the vaccine construct

The VaxiJen server was used to check the antigenic property of the vaccine construct. The prediction result revealed that the vaccine construct has good antigenic behavior with the antigenicity score of 0.6813. The non-allergic behaviour of the vaccine construct was checked through the AlgPred server using different algorithms. The prediction result showed that the vaccine construct does not contain experimentally proven IgE epitope, there is no hit for the MEME/MAST motif and allergen representative peptides and also no hit in BLAST search. The allergenicity of the vaccine construct was cross-checked through AllerTOP v2.0 and AllergenFP v1.0 and found that the vaccine construct has non-allergic nature. So there is strong evidence for the non-allergic nature of our vaccine construct.

3.8. Physicochemical properties and secondary structure analysis of the vaccine construct

The various physicochemical parameters of the vaccine construct were calculated through ExPASy ProtParam server. The molecular weight of the vaccine construct was found to be 16.81 kDa, which indicates the good antigenic nature of the vaccine construct. The theoretical isoelectric point (pI) is valuable to provide buffer system for vaccine purification which was computed to be 9.37 showing slightly basic in nature of the construct, and the total numbers of negative and positive charge residues were found to be 14 and 23, respectively. The extinction-coefficient was 20970 $M^{-1}cm^{-1}$ at 280 nm measured in water, assuming all cysteine residues are reduced. The estimated half-life was found 1 h in mammalian reticulocytes (*in vitro*), while 30 min in yeast (*in vivo*), and > 10 h in *E. coli* (*in vivo*). The instability index classified the vaccine construct as stable with the instability index value of 38.16. The aliphatic index was assessed to be 85.97 showing the designed vaccine is more thermostable because the high aliphatic index

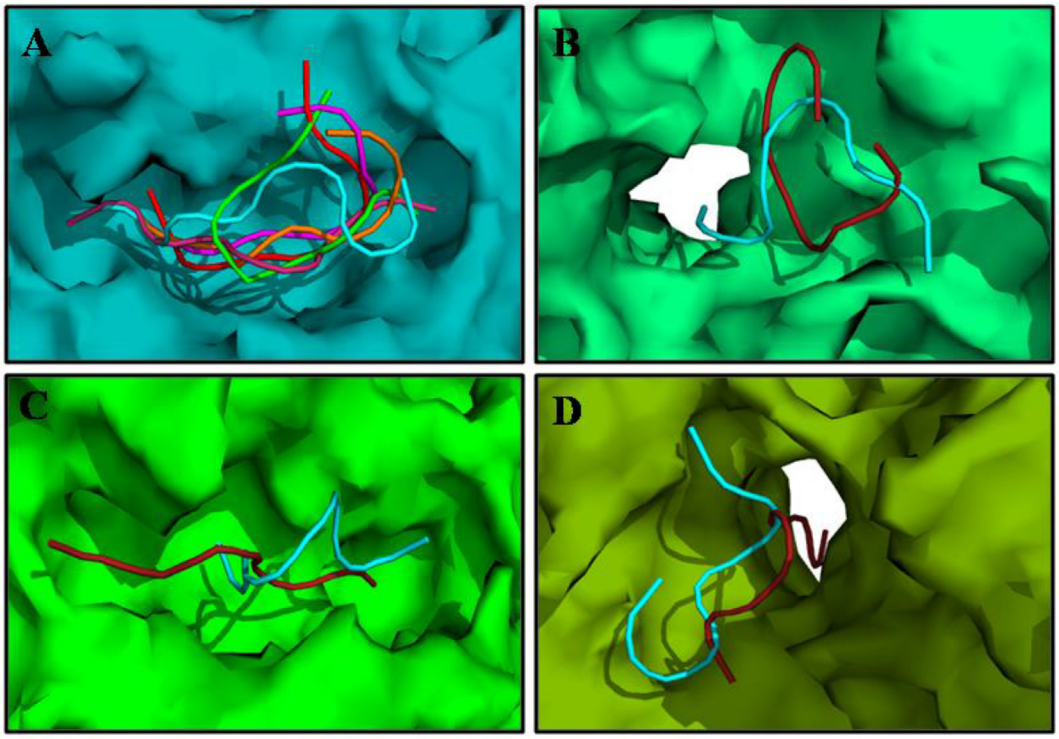

Fig. 2. Peptide-allele docked complexes of the predicted $CD8^+$ T-cell vaccine candidates with their respective frequent alleles. All the peptides and alleles are shown in cartoon structures and surface structures, respectively. (**A**) Presenting the binding interaction of the epitopes HVASGYIEA (Magenta color), YIEAEVIPA (Orange color), GQETAYFLL (Red color), KLLWKGEGA (Hot pink color), and ASDFNLPPV (Green color) to the binding grooves of their common frequent allele HLA-A*02:06 (Cyan color). (**B-D**) Presenting the binding interaction of the epitopes KLAGRWPVK (Firebrick color), KYHSNWRAM (Firebrick color), and VAVHVASGY (Firebrick color) to the binding grooves of their respective frequent alleles HLA-A*03:01 (Lime green color), HLA-C*07:02 (Green color), and HLA-C*12:03 (Green split pea color), respectively. In all the Figures, the cyan color represents the experimental control ligand.

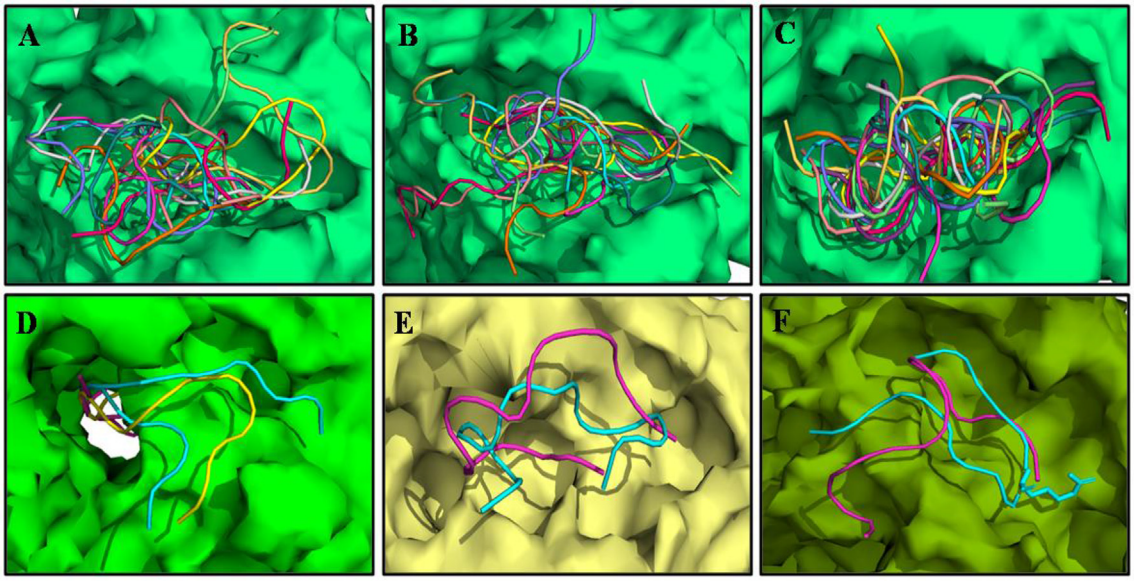

Fig. 3. Peptide-allele docked complexes of the predicted $CD4^+$ T-cell vaccine candidates with their respective frequent alleles. All the peptides and alleles are shown in cartoon structures and surface structures, respectively. (**A-C**) Presenting the binding interaction of 31 epitopes to the binding grooves of their common frequent allele HLA-DRB1*01:01 (Lime green color) where Figure A, B, and C are exhibiting a total of 10, 10, and 11 epitopes, respectively. (**D**) Presenting the binding interaction of the epitopes AVHVASGYIEAEVIP (Magenta color) and DNSDIKVVPRRKAKI (Yellow color) to the binding grooves of their common frequent allele HLA-DRB1*15:01 (Green color). (**E-F**) Presenting the binding interaction of the epitopes KGEGAVVIQDNSDIK (Magenta color) and NWRAMASDFNLPPVV (Magenta color) to the binding grooves of their respective frequent alleles HLA-DRB1*03:01 (Pale yellow color) and HLA-DRB1*04:01 (Split pea color), respectively. In all the Figures, the cyan color represents the experimental control ligand.

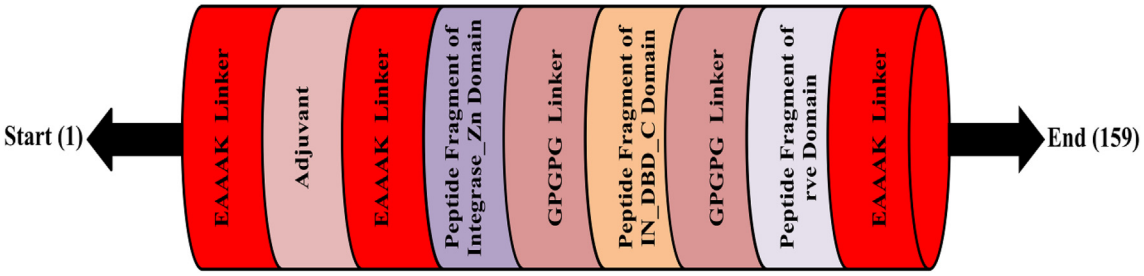

Fig. 4. Schematic diagram of the multi-epitope subunit vaccine construct of length 159 amino acids. An adjuvant was added at the N-terminal site of the construct with the help of EAAAK linker and followed by three peptide fragments fused by GPGPG linker.

is associated with more thermostability. The estimated grand average of hydropathicity (GRAVY) value declared the vaccine construct is hydrophilic in nature with a score of −0.111 that has good interaction with water molecule. The secondary structural features of the vaccine construct were assessed from SOPMA server using the amino acid sequence of length 159 of the construct. The predicted results showed that among 159 amino acids in the final vaccine construct, 54 (33.96%), 34 (21.38%), 19 (11.95%), and 52 (32.70%) amino acids are involved in α-helix, extended strand, β-turn, and random coil, respectively. A probability score graph of occurrence of helix, strand, turn, and coil at each amino acid position in the secondary structure of the final vaccine construct is shown in Fig. S1.

3.9. Analysis of the predicted tertiary structure of vaccine construct through refinement and validation

The tertiary structure of the vaccine construct was predicted using RaptorX server. The vaccine construct was predicted to have a single domain and modelled using 5U1CA as the best template with the score of 112. The p-value for the predicted model was found to be 7.99×10^{-08}; where lower p-value is the indication of the better quality of the predicted 3D model of the vaccine construct. The refinement was done through GalaxyRefine server. GalaxyRefine provided five models with various quality assessment parameters (Table S7). The model 5 was taken into consideration as the best 3D structure of the vaccine construct for further analysis with GDT-HA (0.9230), RMSD (0.512), MolProbity (1.612), clash score (6.1), poor rotamers (1.7), and Rama-favored (97.5).

The Ramachandran plot analysis validated the predicted refined model with 94.4%, 4.0%, 0.8%, and 0.8% residues in most Rama-favored regions, additional allowed regions, generously allowed regions, and disallowed regions, respectively (Fig. 5(A)). The overall quality factor was 89.47% produced by ERRAT which indicates the good quality of the predicted vaccine model and the potential error plot produced by ERRAT was shown in Fig. S2. ProSa-web calculated the Z-score of −3.03 which is in out of range that commonly found in the case of native proteins for comparable size but it is supported by experimentally validated structures (Fig. 5(B)). The structural visualization of the predicted refined 3D model of the vaccine construct was shown in Fig. 6.

3.10. Disulfide engineering, codon adaptation, and in silico cloning of the vaccine construct

To increase the stability of the vaccine construct, disulfide engineering was done through Disulfide by Design 2 (DbD2), and a total of 6 pairs of residues were predicted for the probable formation of disulfide bonds (Table S8). But, after considering the energy score less than 2.2 and the Chi3 value between −87 and +97, there is no selection of residue pairs for making disulfide bond. The adaptation of the codon usage of vaccine constructs as per *E. coli* K12 was done using JCat server. The JCat provided the optimized codon sequence was of 477 nucleotides. The CAI of 1.0 and percentage of the GC-content codon of 52.41% of the optimized codon sequence ensured the high expression rate of vaccine construct in *E. coli* K12. The improved DNA was translated into our vaccine protein appropriately with GC-content codon of 50.73% of *E. coli* K12. Later on, the adapted codon sequence was inserted into the *E. coli* pET28a(+) vector and obtained a cloned vaccine of 5812 base pairs (Fig. S3).

3.11. Molecular docking of final vaccine construct with immune receptor TLR3

The vaccine-receptor docking was performed through ClusPro v2.0 for evaluating the complex form of the vaccine construct with the immune receptor TLR3 along with their binding affinity. The ClusPro 2.0 server predicted a total of 30 docked complex forms with their respective centre score and lowest energy scores (Table S9). The model namely model.000.05 was selected as best-docked complex between vaccine and receptor based on the lowest binding energy. The centre score and lowest energy of the best docked-complex were found to be −1026.8. The 3D structure of the best docked-complex is illustrated in Fig. 7.

4. Discussion

The human immunodeficiency virus (HIV) is a T-helper cell killer pandemic virus responsible for the acquired immune deficiency syndrome (AIDS) and is threatening the human life from its first identification. HIV is classified into two categories, and out of these, HIV-1 is the widespread and leading cause of AIDS. In the successful HIV-1 life cycle, the integrase enzyme is the logical target for vaccine design [3]. A protein domain is defined as the independently evolved, functional, and structural conserved part of a protein sequence, which is functional and structural representative of the respective protein. Therefore, the conserved domain can be the novel target for vaccine design. The lack of understanding the immunity against the virulent pathogen HIV makes the vaccine development against this deadliest pathogen more challenging work to the researchers. However, in modern times, immunotherapy is the leading powerful and efficient strategy for the prevention of infectious diseases. A powerful branch of bioinformatics called immunoinformatics approach has been used to develop vaccine by analyzing immunological data using low cost within short-time [48]. There are many multi-epitope vaccine design studies involving various viruses like Dengue virus [38], Hepatitis B virus [49], Hepatitis C virus [50,51], Ebola virus [52], Chikungunya virus [53], Avian influenza A (H7N9) virus [54], Zika virus [55], Classical swine fever virus [56], Nipah virus [57], and Norovirus [58]. The researcher Olga and his coworkers studied *Leishmania infantum* eukaryotic Initiation Factor (LieIF) protein by *in silico* approaches using antigenicity prediction algorithms and predicted possible H-2 class I and II alleles presented LieIF epitopes. Later, they synthesized five promising LieIF peptides and tested their immunogenicity in vitro as well as their immunomodulatory properties in an *in vitro* system of bone marrow-derived dendritic cells [59]. On the

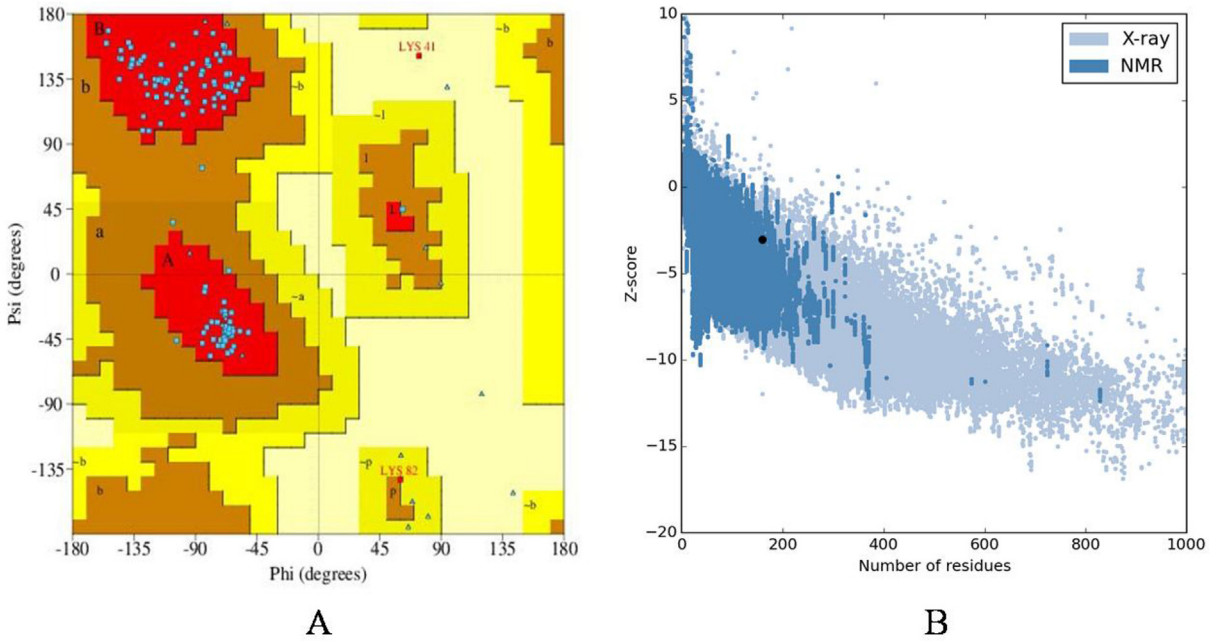

Fig. 5. Validation of the predicted 3D model of the final vaccine construct. (**A**) Ramachandran plot of the refined model showing 94.4%, 4.0%, 0.8%, and 0.8% residues were found in most Rama-favored regions, additional allowed regions, generously allowed regions, and disallowed regions respectively. (**B**) The overall quality Z-score plot from ProSa server showing Z-score of −3.03.

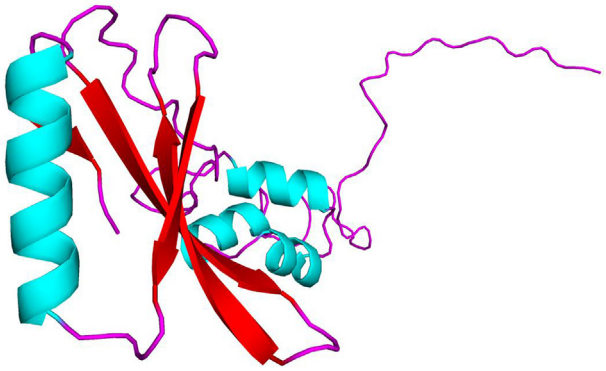

Fig. 6. Refined 3D structure of the vaccine construct where the cyan, red, and magenta colors represents the helix, sheet, and loop, respectively.

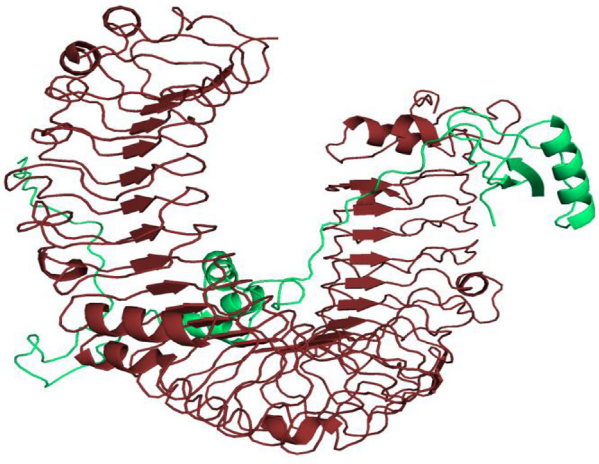

Fig. 7. Vaccine-receptor docked complex obtained after docking between the multi-epitope vaccine construct (Limegreen color) and the immune receptor TLR3 (Firebrick color) in ClusPro 2.0 server.

other hand, Andres *et* al. used immunoinformatics tools to identify class I and II T-cell epitopes that are highly conserved in seven representative strains of influenza A virus (IAV) in U.S. Pigs, and that are expected to be commercially available Pigs bind predominant leukocyte antigen (SLA) alleles. Finally, they constructed a multi-epitope plasmid DNA vaccine and performed *in vivo* test in Pigs [60].

The peptide-based vaccine provides some advantages such as wide HLA-restrictions, an absence of infectious potential and self-reactivity, ability to exclude deleterious sequences from full-length antigens or other pathogen-derived molecules, improved immunogenicity, improved stability, improved solubility, easy to produce, easy to analyze for quality control [61,62]. Although there are some drawbacks such as MHC restriction, lack of helper activity, weak presentation of antigen by antigen-presenting cell (APC) especially in small peptide vaccine, peptide-containing multiple epitopes for inducing antigen-specific-$CD8^+$ & $CD4^+$ T-cell overcome these drawbacks [62,63]. In present times, the multi-epitope subunit vaccine is preferable than the traditional vaccine for several advantages including safety, higher stability, less allergic, autoimmune responses, and more convenient production [64]. This study used immunoinformatics approaches for designing an efficient multi-epitope subunit vaccine which consists of peptide fragments containing overlapping multiple T-cell (T_C & T_H) epitopes and B-cell epitopes that are immunodominant and can produce specific immune responses along with proper adjuvant and linkers [33]. In case of multi-epitope subunit vaccine, considering peptide fragments containing overlapping multiple $CD8^+$, $CD4^+$, and B-cell epitopes is one of good approach to keep small length of the vaccine construct whenever it contains large number of overlapping epitopes [65,66]. Besides the advantages, there is an important drawback of low immunogenicity of the multi-epitope subunit vaccine, but the addition of a proper adjuvant to the vaccine construct can overcome this drawback by playing a significant role in stimulation of immune system as well as vaccine delivery [67–69]. In this study, the β-defensin-2 was used as an adjuvant, which can inhibit HIV-1 replication [70]. The linkers EAAAK and GPGPG have been widely used in most of the previous studies on the design of multi-epitope vaccines for viruses and in many other immunoinformatics studies [9,38,43,71]. Therefore, we selected these linkers to bring together our selected vaccine candidates.

The T-cells having capability of inducing multiple cytokines are highly associated with protection against HIV [72]. The vaccine design based on T-cell epitopes is a distinctive method for inducing vast repertoires of robust immune responses of the host cell to the infected cell, dealing more effectively with genetic variation both in pathogens and humans [73–75]. The necessity of T-cell epitope-based vaccine is promising because cytotoxic T lymphocytes (CTLs or $CD8^+$ T-cell) control the viral infection during asymptomatic HIV phase. Therefore, it interrupts the HIV progression by detecting and killing the malignant cells through the recognition of the epitopes presented by MHC-I molecule on the cell surface [9,61,76,77]. On the other hand, helper T lymphocytes (HTLs or $CD4^+$ T-cell) provide agnate help for the propagation of the potent humoral and cellular responses by promoting optimal expansion of $CD8^+$ T-cells, maintaining an effective $CD8^+$ T-cell [78,79], and hence HIV-1 clinical trial suggested that the $CD4^+$ T-cells are potential vaccine candidate to induce vigorous immune responses [73]. The CTLs monitor the cell surface major histocompatibility (MHC) class I molecules which can alert the immune system to virally infected cells and present peptides derived from intracellular viral proteins for eliminating the infected cells [80]. MHC class II molecules monitored by HTLs can stimulate cellular and humoral immunity by extracellularly derived peptides to antigen-presenting cells such as dendritic cells, B-cells, etc. [12]. B-cell plays a key role in the humoral immunity of the adaptive immune system by producing antibodies and for the clearance of blood pathogen in an antibody-dependent manner [81]. The subunit vaccine containing the antigenic B-cell epitopes can enhance the humoral and cellular immunity [43].

Antigenicity is the capability of an antigen to produce the immune response by binding to T-cells or B-cells, where the strong binding ability to T-cells indicates the high antigenicity. On the other hand, immunogenicity is the measure of an antigen to produce an immune response when an antigen causes an immune response but does not bind to T-cells. So, antigenicity and immunogenicity are important for vaccine design, and these are considered as the center of the vaccine competency. It would be expected that, in case of epitope-based vaccine design, the use of highly conserved epitopes can provide broader protection across multiple strains or even species [15]. The antigenicity, immunogenicity and conservancy were used as the epitope selection parameter and the epitopes interacted with ≥ 5 MHC HLA-alleles are most potential vaccine candidate [62,63,74]. A vaccine will be more effective if it covers maximum population in any geographical region especially in the region where the associated disease frequently occurred; hence the population coverage is a key factor for the development of the peptide-based vaccine [74]. The exactly or approximately 100% population coverage of the HLA alleles of the predicted epitopes lead to consider these epitopes as the best presumable vaccine candidates [63,82]. Sub-Saharan Africa, Eastern Europe, Central Asia, North America, Latin America, South Asia, Southeast Asia, and Oceania are the most HIV-1 affected geographical region [9,83]. In this study, the population coverage analysis showed an excellent coverage of the population in those regions (Table S4 and Table 3).

Allergy involves a series of complex reaction by the immune system due to the foreign antigenic substance and results in sneezing, wheezing, skin rash, and swelling of the mucous membrane [71]. A vaccine with non-allergic behavior will be completely safe for human life [71]. The various physicochemical parameters of the designed vaccine concluded the vaccine as good antigenic, basic, stable, high thermostable, and hydrophilic [35,38,43,71]. The p-value is the important model's quality parameter of the RaptorX server, where the lower p-value declared the better quality of the predicted 3D model of the vaccine construct [37,43]. The quality of the predicted tertiary structure of the final vaccine construct from a template-based protein modeling tool depends on the degree of the alikeness of the target and available template structure. Hence, the precision of the predicted model was improved using refinement technique. Protein structure validation is a computational process of evaluating the reliability of the predicted refined 3D model using different quality factors such as Ramachandran plot, different quality factors, and errors plots analysis. In the Ramachandran plot analysis, a good quality model would probably exceed 90% in the most favored regions [84]. The analysis of the Ramachandran plot, different quality factors, and errors plots of the 3D model of the vaccine construct shows good quality of the predicted refined 3D model of the final vaccine construct [40–43]. Since the function and structure of the vaccine protein are directly related to its stability, so increasing of stability is essential during the vaccine design. However, several factors affect the protein stability, and disulfide bond is one of them which can help us to increase stability by increasing the number of disulfide bonds [85,86]. We selected *E. coli* K12 as a host to check the expression rate of the final vaccine protein, so the codon optimization was performed to improve the transcriptional and translational efficiency using codon adaptation index and total GC content of DNA sequence analysis [87]. There is an essential requirement of solubility of the vaccine protein in host *E. coli* K12 for many biochemical and functional investigations. The analysis results prove that the final vaccine protein shows an admissible solubility [43]. Finally, it is necessary to know the immune response of HIV-1 mediated dendritic cell-dependent TLR3 against the vaccine protein [43]. To make sure, the molecular docking was performed to check the binding affinity between immune receptor and vaccine, and the analysis result shows the good binding affinity between TLR3 and vaccine protein [9], which initially ensure the acceptability of our designed vaccine construct.

5. Conclusion

The human immunodeficiency virus continues its severity as the potential agent for the morbidity and mortality of AIDS. The lack of permanent prevention against HIV-1 makes it more severe. Therefore, this study designed a multi-epitope subunit vaccine against HIV-1 by applying the novel immunoinformatics approach to the conserved domains of the integrase proteins of HIV-1. In this study, the high immunogenic T-cell (CTL and HTL) and B-cell epitopes were predicted as potential vaccine candidates. The predicted epitopes were characterized by the high antigenicity, conservancy, good binding affinity with their corresponding HLA-alleles, and excellent population coverage. The designed vaccine contained a suitable adjuvant β-defensin-2 for increasing immunogenicity and followed by peptide fragments containing T (HTL & CTL), and B-cell epitopes for inducing cellular and humoral immunity, respectively. The vaccine construct showed high antigenicity with no allergic function. The designed vaccine was predicted as good antigenic, basic, stable, higher thermostable, and hydrophilic. The high expression rate of the designed vaccine in *E. coli* K12 was confirmed by performing codon adaptation and *in silico* cloning. The predicted complex of the vaccine construct with the immune receptor TLR3 showed good binding affinity. Finally, this study requires the experimental validation to confirm its anti-HIV-1 functional activity.

Author contribution

Concept and planning were done by FA, UKA, and MKA; data preparing and computational analysis were done by FA and UKA; the manuscript was written by FA and UKA; the manuscript was reviewed and rewritten by FA, UKA, and MKA.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The researchers strictly declare that there is no potential conflict of

interests regarding the publication of this research article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.micpath.2019.103791.

References

[1] D.B. Cluck, R.F. Underwood, A therapeutic perspective of living with human immunodeficiency virus/AIDS in 2017, Nurs. Clin. N. Am. 53 (2018) 97–110, https://doi.org/10.1016/j.cnur.2017.10.009.

[2] G. Li, E. De Clercq, HIV genome-wide protein associations: a review of 30 Years of research, Microbiol. Mol. Biol. Rev. 80 (2016) 679–731, https://doi.org/10.1128/MMBR.00065-15.

[3] Y. Pommier, A.A. Johnson, C. Marchand, Integrase inhibitors to treat HIV/Aids, Nat. Rev. Drug Discov. 4 (2005) 236–248, https://doi.org/10.1038/nrd1660.

[4] I. Stratov, R. DeRose, D.F.J. P, S.J. Kent, Vaccines and vaccine strategies against HIV, Curr. Drug Targets 5 (2004) 71–88, https://doi.org/10.2174/1389450043490686.

[5] J. Shu, X. Fan, J. Ping, X. Jin, P. Hao, Designing peptide-based HIV vaccine for Chinese, BioMed Res. Int. 2014 (2014), https://doi.org/10.1155/2014/272950.

[6] R. Mahendran, S. Jeyabaskar, G. Sitharaman, R.D. Michael, A.V. Paul, Computer-aided vaccine designing approach against fish pathogens Edwardsiella tarda and Flavobacterium columnare using bioinformatics softwares, Drug Des. Dev. Ther. 2016 (10) (2016) 1703–1714, https://doi.org/10.2147/DDDT.S95691.

[7] P. Chiarella, E. Massi, M. De Robertis, V.M. F, E. Signori, Recent advances in epitope design for immunotherapy of cancer, recent pat. Anticancer, Drug Discov 4 (2009) 227–240, https://doi.org/10.2174/157489209789206922.

[8] C.L. Slingluff Jr., The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination? Cancer J. 17 (2011) 343–350, https://doi.org/10.1097/PPO.0b013e318233e5b2.

[9] R.K. Pandey, R. Ojha, V.S. Aathmanathan, M. Krishnan, V.K. Prajapati, Immunoinformatics approaches to design a novel multi-epitope subunit vaccine against HIV infection, Vaccine 36 (2018) 2262–2272, https://doi.org/10.1016/j.vaccine.2018.03.042.

[10] M. V Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction, BMC Bioinf. 8 (2007) 424, https://doi.org/10.1186/1471-2105-8-424.

[11] M. Andreatta, M. Nielsen, Gapped sequence alignment using artificial neural networks: application to the MHC class I system, Bioinformatics 32 (2015) 511–517, https://doi.org/10.1093/bioinformatics/btv639.

[12] M. Nielsen, O. Lund, NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction, BMC Bioinf. 10 (2009) 296, https://doi.org/10.1186/1471-2105-10-296.

[13] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinf. 8 (2007) 4, https://doi.org/10.1186/1471-2105-8-4.

[14] J.J.A. Calis, M. Maybeno, J.A. Greenbaum, D. Weiskopf, A.D. De Silva, A. Sette, C. Keşmir, B. Peters, Properties of MHC class I presented peptides that enhance immunogenicity, PLoS Comput. Biol. 9 (2013), https://doi.org/10.1371/journal.pcbi.1003266.

[15] H.-H. Bui, J. Sidney, W. Li, N. Fusseder, A. Sette, Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines, BMC Bioinf. 8 (2007) 361, https://doi.org/10.1186/1471-2105-8-361.

[16] S. Saha, G.P.S. Raghava, AlgPred: prediction of allergenic proteins and mapping of IgE epitopes, Nucleic Acids Res. 34 (2006) W202–W209, https://doi.org/10.1093/nar/gkl343.

[17] I. Dimitrov, I. Bangov, D.R. Flower, I. Doytchinova, AllerTOP v.2—a server for in silico prediction of allergens, J. Mol. Model. 20 (2014) 2278, https://doi.org/10.1007/s00894-014-2278-5.

[18] I. Dimitrov, L. Naneva, I. Doytchinova, I. Bangov, AllergenFP: allergenicity prediction by descriptor fingerprints, Bioinformatics 30 (2013) 846–851, https://doi.org/10.1093/bioinformatics/btt619.

[19] H.-H. Bui, J. Sidney, K. Dinh, S. Southwood, M.J. Newman, A. Sette, Predicting population coverage of T-cell epitope-based diagnostics and vaccines, BMC Bioinf. 7 (2006) 153, https://doi.org/10.1186/1471-2105-7-153.

[20] G. Morris, R. Huey, W. Lindstrom, M. Sanner, R. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791, https://doi.org/10.1002/jcc.21256.

[21] T. Oleg, A.J. O, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461, https://doi.org/10.1002/jcc.21334.

[22] V.L. Murthy, L.J. Stern, The class II MHC protein HLA-DR1 in complex with an endogenous peptide: implications for the structural basis of the specificity of peptide binding, Structure 5 (1997) 1385–1396, https://doi.org/10.1016/S0969-2126(97)00288-8.

[23] E.F. Rosloniec, R.A. Ivey, K.B. Whittington, A.H. Kang, H.-W. Park, Crystallographic structure of a rheumatoid arthritis MHC susceptibility allele, HLA-DR1 (DRB1*0101), complexed with the immunodominant determinant of human type II collagen, J. Immunol. 177 (2006) 3884–LP – 3892, https://doi.org/10.4049/jimmunol.177.6.3884.

[24] G. Kaur, S. Gras, J.I. Mobbs, J.P. Vivian, A. Cortes, T. Barber, S.B. Kuttikkatte, L.T. Jensen, K.E. Attfield, C.A. Dendrou, M. Carrington, G. McVean, A.W. Purcell, J. Rossjohn, L. Fugger, Structural and regulatory diversity shape HLA-C protein expression levels, Nat. Commun. 8 (2017) 15924, https://doi.org/10.1038/ncomms15924.

[25] S. Zhang, J. Liu, H. Cheng, S. Tan, J. Qi, J. Yan, G.F. Gao, Structural basis of cross-allele presentation by HLA-A*0301 and HLA-A*1101 revealed by two HIV-derived peptide complexes, Mol. Immunol. 49 (2011) 395–401, https://doi.org/10.1016/j.molimm.2011.08.015.

[26] J.C. Boyington, S.A. Motyka, P. Schuck, A.G. Brooks, P.D. Sun, Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand, Nature 405 (2000) 537–543, https://doi.org/10.1038/35014520.

[27] M. Hahn, M.J. Nicholson, J. Pyrdol, K.W. Wucherpfennig, Unconventional topology of self peptide–major histocompatibility complex binding by a human autoimmune T cell receptor, Nat. Immunol. 6 (2005) 490–496, https://doi.org/10.1038/ni1187.

[28] P. Ghosh, M. Amaya, E. Mellins, D.C. Wiley, The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3, Nature 378 (1995) 457–462, https://doi.org/10.1038/378457a0.

[29] C. Gerstner, A. Dubnovitsky, C. Sandin, G. Kozhukh, H. Uchtenhagen, E.A. James, J. Rönnelid, A.J. Ytterberg, J. Pieper, E. Reed, K. Tandre, M. Rieck, R.A. Zubarev, L. Rönnblom, T. Sandalova, J.H. Buckner, A. Achour, V. Malmström, Functional and structural characterization of a novel HLA-DRB1*04:01-restricted α-enolase T cell epitope in rheumatoid arthritis, Front. Immunol. 7 (2016) 494, https://doi.org/10.3389/fimmu.2016.00494.

[30] A. Lamiable, P. Thévenet, J. Rey, M. Vavrusa, P. Derreumaux, P. Tufféry, PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex, Nucleic Acids Res. 44 (2016) W449–W454, https://doi.org/10.1093/nar/gkw329.

[31] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: an open chemical toolbox, J. Cheminf. 3 (2011) 33, https://doi.org/10.1186/1758-2946-3-33.

[32] Y. EL-Manzalawy, D. Dobbs, V. Honavar, Predicting linear B-cell epitopes using string kernels, J. Mol. Recognit. 21 (2008) 243–255, https://doi.org/10.1002/jmr.893.

[33] S.H. van der Burg, M.S. Bijker, M.J.P. Welters, R. Offringa, C.J.M. Melief, Improved peptide vaccine strategies, creating synthetic artificial infections to maximize immune efficacy, Adv. Drug Deliv. Rev. 58 (2006) 916–930, https://doi.org/10.1016/j.addr.2005.11.003.

[34] T. Mohan, C. Sharma, A.A. Bhat, D.N. Rao, Modulation of HIV peptide antigen specific cellular immune response by synthetic α- and β-defensin peptides, Vaccine 31 (2013) 1707–1716, https://doi.org/10.1016/j.vaccine.2013.01.041.

[35] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R. Appel, A. Bairoch, Protein identification and analysis tool on the ExPASy server, Proteomics Protoc. Handb, 2005, pp. 571–607, , https://doi.org/10.1385/1-59259-890-0:571.

[36] C. Geourjon, G. Deleage, SOPMA: significant improvements in protein secondary structure prediction by prediction from multiple alignments, Bioinformatics 11 (1995) 681–684, https://doi.org/10.1093/bioinformatics/11.6.681.

[37] J. Peng, J. Xu, RaptorX: exploiting structure information for protein alignment by statistical inference, Proteins 79 (2011) 161–171, https://doi.org/10.1002/prot.23175.

[38] M. Ali, R.K. Pandey, N. Khatoon, A. Narula, A. Mishra, V.K. Prajapati, Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection, Sci. Rep. 7 (2017) 9232, https://doi.org/10.1038/s41598-017-09199-w.

[39] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (2013) 384–388, https://doi.org/10.1093/nar/gkt458.

[40] R.A. Laskowski, J. Jabłońska, L. Pravda, R.S. Vařeková, J.M. Thornton, PDBsum: structural summaries of PDB entries, Protein Sci. 27 (2018) 129–134, https://doi.org/10.1002/pro.3289.

[41] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407–W410, https://doi.org/10.1093/nar/gkm290.

[42] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519, https://doi.org/10.1002/pro.5560020916.

[43] N. Khatoon, R.K. Pandey, V.K. Prajapati, Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach, Sci. Rep. 7 (2017) 8285, https://doi.org/10.1038/s41598-017-08842-w.

[44] D.B. Craig, A.A. Dombkowski, Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins, BMC Bioinf. 14 (2013) 346, https://doi.org/10.1186/1471-2105-14-346.

[45] A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel, D. Jahn, JCat: a novel tool to adapt codon usage of a target gene to its potential expression host, Nucleic Acids Res. 33 (2005) W526–W531, https://doi.org/10.1093/nar/gki376.

[46] R.K. Pandey, B.V. Kumbhar, S. Srivastava, R. Malik, S. Sundar, A. Kunwar, V.K. Prajapati, 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 (2017) 141–158, https://doi.org/10.1080/07391102.2015.1135298.

[47] S.R. Comeau, D.W. Gatchell, S. Vajda, C.J. Camacho, ClusPro: a fully automated algorithm for protein-protein docking, Nucleic Acids Res. 32 (2004) W96–W99, https://doi.org/10.1093/nar/gkh354.

[48] H. Zahroh, A. Ma'rup, U.S.F. Tambunan, A.A. Parikesit, Immunoinformatics approach in designing epitope-based vaccine against meningitis-inducing bacteria

(Streptococcus pneumoniae, Neisseria meningitidis, and Haemophilus influenzae type b), Drug Target Insights 10 (2016) 19–29, https://doi.org/10.4137/DTI.S38458.
[49] E. Depla, A. Van der Aa, B.D. Livingston, C. Crimi, K. Allosery, V. De Brabandere, J. Krakover, S. Murthy, M. Huang, S. Power, L. Babe, C. Dahlberg, D. McKinney, A. Sette, S. Southwood, R. Philip, M.J. Newman, L. Meheus, Rational design of a multiepitope vaccine encoding T-lymphocyte epitopes for treatment of chronic Hepatitis B virus infections, J. Virol. 82 (2008) 435–450, https://doi.org/10.1128/jvi.01505-07.
[50] M. Nosrati, H. Mohabatkar, M. Behbahani, A novel multi-epitope vaccine for cross protection against Hepatitis C virus (HCV): an immunoinformatics approach, Res. Mol. Med. 5 (2017) 17–26, https://doi.org/10.29252/rmm.5.1.17.
[51] A. Ikram, T. Zaheer, F.M. Awan, A. Obaid, A. Naz, R. Hanif, R.Z. Paracha, A. Ali, A.K. Naveed, H.A. Janjua, Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches, Sci. Rep. 8 (2018) 1–14, https://doi.org/10.1038/s41598-018-34254-5.
[52] S.I. Bazhan, D.V. Antonets, L.I. Karpenko, S.F. Oreshkova, O.N. Kaplina, E.V. Starostina, S.G. Dudko, S.A. Fedotova, A.A. Ilyichev, In silico designed ebola virus T-cell multi-epitope DNA vaccine constructions are immunogenic in mice, Vaccines 7 (2019), https://doi.org/10.3390/vaccines7020034.
[53] A. Narula, R.K. Pandey, N. Khatoon, A. Mishra, V.K. Prajapati, Excavating chikungunya genome to design B and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection, Infect. Genet. Evol. 61 (2018) 4–15, https://doi.org/10.1016/j.meegid.2018.03.007.
[54] M. Hasan, P.P. Ghosh, K.F. Azim, S. Mukta, R.A. Abir, J. Nahar, M.M. Hasan Khan, Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus, Microb. Pathog. 130 (2019) 19–37, https://doi.org/10.1016/j.micpath.2019.02.023.
[55] R. Kumar Pandey, R. Ojha, A. Mishra, V. Kumar Prajapati, Designing B- and T-cell multi-epitope based subunit vaccine using immunoinformatics approach to control Zika virus infection, J. Cell. Biochem. 119 (2018) 7631–7642, https://doi.org/10.1002/jcb.27110.
[56] J. Ying, X. Dong, Y. Chen, Preliminary evaluation of a candidate multi-epitope-vaccine against the classical swine fever virus, Tsinghua Sci. Technol. 13 (2008) 433–438, https://doi.org/10.1016/S1007-0214(08)70070-1.
[57] R. Ojha, A. Pareek, R.K. Pandey, D. Prusty, V.K. Prajapati, Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus Zoonotic infection, ACS Omega 4 (2019) 13069–13079, https://doi.org/10.1021/acsomega.9b00944.
[58] K.F. Azim, M. Hasan, M.N. Hossain, S.R. Somana, S.F. Hoque, M.N.I. Bappy, A.T. Chowdhury, T. Lasker, Immunoinformatics approaches for designing a novel multi epitope peptide vaccine against human norovirus (Norwalk virus), Infect. Genet. Evol. 74 (2019) 103936, https://doi.org/10.1016/j.meegid.2019.103936.
[59] O.S. Koutsoni, J.G. Routsias, I.D. Kyriazis, M. Barhoumi, I. Guizani, A. Tsakris, E. Dotsika, In silico analysis and in vitro evaluation of immunogenic and immunomodulatory properties of promiscuous peptides derived from Leishmania infantum eukaryotic initiation factor, Bioorg. Med. Chem. 25 (2017) 5904–5916, https://doi.org/10.1016/j.bmc.2017.07.013.
[60] A.H. Gutiérrez, C. Loving, L. Moise, F.E. Terry, S.L. Brockmeier, H.R. Hughes, W.D. Martin, A.S. De Groot, In vivo validation of predicted and conserved T cell epitopes in a swine influenza model, PLoS One 11 (2016), https://doi.org/10.1371/journal.pone.0159237 e0159237–e0159237.
[61] A.W. Purcell, J. McCluskey, J. Rossjohn, More than one reason to rethink the use of peptides in vaccine design, Nat. Rev. Drug Discov. 6 (2007) 404, https://doi.org/10.1038/nrd2224.
[62] D. Gourav, L. Neha, J. Sahil, B. Manoj, Metadherin peptides containing CD4+ and CD8+ T cell epitopes as a therapeutic vaccine candidate against cancer, Microbiol. Immunol. 60 (2016) 646–652, https://doi.org/10.1111/1348-0421.12436.
[63] U.K. Adhikari, M.M. Rahman, Overlapping CD8 + and CD4 + T-cell epitopes identification for the progression of epitope-based peptide vaccine from nucleocapsid and glycoprotein of emerging Rift Valley fever virus using immunoinformatics approach, Infect. Genet. Evol. 56 (2017) 75–91, https://doi.org/10.1016/j.meegid.2017.10.022.
[64] M. Negahdaripour, N. Golkar, N. Hajighahramani, S. Kianpour, N. Nezafat, Y. Ghasemi, Harnessing self-assembled peptide nanoparticles in epitope vaccine design, Biotechnol, Adv 35 (2017) 575–596, https://doi.org/10.1016/j.biotechadv.2017.05.002.
[65] G. Dhiman, N. Lohia, S. Jain, M. Baranwal, Metadherin peptides containing CD4+ and CD8+ T cell epitopes as a therapeutic vaccine candidate against cancer, Microbiol. Immunol. 60 (2016) 646–652, https://doi.org/10.1111/1348-0421.12436.
[66] S. Jain, M. Baranwal, Computational analysis in designing T cell epitopes enriched peptides of Ebola glycoprotein exhibiting strong binding interaction with HLA molecules, J. Theor. Biol. 465 (2019) 34–44, https://doi.org/10.1016/j.jtbi.2019.01.016.
[67] A.S. De Groot, L. Moise, J.A. McMurry, W. Martin, Epitope-based immunome-derived vaccines: a strategy for improved design and safety, in: A. Falus (Ed.), Clin. Appl. Immunomics, Springer US, New York, NY, 2009, pp. 39–69, , https://doi.org/10.1007/978-0-387-79208-8_3.
[68] T. Farhadi, R.S. Ovchinnikov, M.M. Ranjbar, In silico designing of some agonists of toll-like receptor 5 as a novel vaccine adjuvant candidates, Netw. Model. Anal. Heal. Informatics Bioinforma. 5 (2016) 31, https://doi.org/10.1007/s13721-016-0138-1.
[69] X. Liu, L.M. Wetzler, P. Massari, The PorB porin from commensal Neisseria lactamica induces Th1 and Th2 immune responses to ovalbumin in mice and is a potential immune adjuvant, Vaccine 26 (2008) 786–796, https://doi.org/10.1016/j.vaccine.2007.11.080.
[70] M. Quinones-Mateu, M.M. Lederman, Z. Feng, B. Chakraborty, J. Weber, H. Rangel, M. Marotta, M. Mirza, B. Jiang, P. Kiser, K. Medvik, S. Sieg, A. Weinberg, Human epithelial beta-defensins 2 and 3 inhibit HIV-1 replication, AIDS 17 (2003) F39–F48, https://doi.org/10.1097/01.aids.0000096878.73209.4f.
[71] R.K. Pandey, T.K. Bhatt, V.K. Prajapati, Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating Anopheles salivary protein, Sci. Rep. 8 (2018) 1125, https://doi.org/10.1038/s41598-018-19456-1.
[72] G. Pantaleo, C. Graziosi, A.S. Fauci, The role of lymphoid organs in the immunopathogenesis of HIV infection, AIDS 7 (1993) S19–S23, https://doi.org/10.1097/00002030-199301001-00003.
[73] P. Oyarzún, B. Kobe, Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production, Hum. Vaccines Immunother. 12 (2016) 763–767, https://doi.org/10.1080/21645515.2015.1094595.
[74] M.A. Khan, M. Hossain, M. Islam, S.M. Rakib-Uz-Zaman, N. Morshed, Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study, Scand. J. Immunol. 82 (2015), https://doi.org/10.1111/sji.12302.
[75] L. Klavinskis, J.L. Whitton, M.B.A. Oldstone, Monoclonal engineered vaccine which expresses an immunodominant T cell epitope induces CTL that confer protection from lethal virus infection, J. Virol. 63 (1989) 4311–4316.
[76] B. Walker, A. McMichael, The T-cell response to HIV, Cold Spring Harb. Perspect. Med. 2 (2012), https://doi.org/10.1101/cshperspect.a007054.
[77] K.C. Garcia, L. Teyton, I.A. Wilson, Structural basis OF T cell recognition, Annu. Rev. Immunol. 17 (1999) 369–397, https://doi.org/10.1146/annurev.immunol.17.1.369.
[78] D.S. Rosa, S.P. Ribeiro, E. Cunha-Neto, CD4+ T cell epitope discovery and rational vaccine design, Arch. Immunol. Ther. Exp. (Warsz). 58 (2010) 121–130, https://doi.org/10.1007/s00005-010-0067-0.
[79] S. Kalams, S. Buchbinder, E. Rosenberg, J. Billingsley, D.S. Colbert, N. Jones, A.K. Shea, A.K. Trocha, B.D. Walker, Association between virus-specific cytotoxic T-lymphocyte and helper responses in human immunodeficiency virus type 1 infection, J. Virol. 73 (1999) 6715–6720.
[80] E.W. Hewitt, The MHC class I antigen presentation pathway: strategies for viral immune evasion, Immunology 110 (2003) 163–169, https://doi.org/10.1046/j.1365-2567.2003.01738.x.
[81] S. Moir, A.S. Fauci, B cells in HIV infection and disease, Nat. Rev. Immunol. 9 (2009) 235–245, https://doi.org/10.1038/nri2524.
[82] U.K. Adhikari, M. Tayebi, M.M. Rahman, Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging Oropouche virus, J. Immunol. Res. 2018 (2018), https://doi.org/10.1155/2018/6718083.
[83] P.M. Sharp, B.H. Hahn, Origins of HIV and the AIDS pandemic, Cold Spring Harb. Perspect. Med. 1 (2011) a006841, , https://doi.org/10.1101/cshperspect.a006841.
[84] A. Wlodawer, Stereochemistry and validation of macromolecular structures, Methods Mol. Biol. 1607 (2017) 595–610, https://doi.org/10.1007/978-1-4939-7000-1_24.
[85] M. Akke, S. Forsén, Protein stability and electrostatic interactions between solvent exposed charged side chains, Proteins Struct. Funct. Bioinforma. 8 (1990) 23–29, https://doi.org/10.1002/prot.340080106.
[86] S.F. Betz, Disulfide bonds and the stability of globular proteins, Protein Sci. 2 (1993) 1551–1558, https://doi.org/10.1002/pro.5560021002.
[87] S.C. Makrides, Strategies for achieving high-level expression of genes in Escherichia coli, Microbiol. Rev. 60 (1996) 512–538 https://www.ncbi.nlm.nih.gov/pubmed/8840785.