Heliyon 9 (2023) e22121

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Prediction of an immunogenic peptide ensemble and multi-subunit vaccine for Visceral leishmaniasis using bioinformatics approaches

Manu Kupani [a], Rajeev Kumar Pandey [b], Sharad Vashisht [c], Satyendra Singh [d], Vijay Kumar Prajapati [d], Sanjana Mehrotra [a,*]

[a] *Department of Human Genetics, Guru Nanak Dev University, Amritsar, 143005, Punjab, India*
[b] *Research & Development, Thermo Fisher Scientific, Bangalore, 560066, Karnataka, India*
[c] *Regional Centre for Biotechnology, NCR Biotech Science Cluster, 3rd Milestone, Faridabad-Gurugram Expressway, Faridabad, 121001, Harayana, India*
[d] *Department of Biochemistry, University of Delhi South Campus, Benito Juarez Road, Dhaula Kuan, New Delhi, 110021, India*

ARTICLE INFO

Keywords:
Visceral leishmaniasis
Vaccine
Mapping
Immunogenicity
Antigen
HLA
Epitope

ABSTRACT

Visceral Leishmaniasis (VL) is a neglected tropical disease of public health importance in the Indian subcontinent. Despite consistent elimination initiatives, the disease has not yet been eliminated and there is an increased risk of resurgence from active VL reservoirs including asymptomatic, post kala azar dermatitis leishmaniasis (PKDL) and HIV-VL co-infected individuals. To achieve complete elimination and sustain it in the long term, a prophylactic vaccine, which can elicit long lasting immunity, is desirable. In this study, we employed immunoinformatic tools to design a multi-subunit epitope vaccine for the Indian population by targeting antigenic secretory proteins screened from the *Leishmania donovani* proteome. Out of 8014 proteins, 277 secretory proteins were screened for their cellular location and proteomic evidence. Through NCBI BlastP, unique fragments of the proteins were cropped, and their antigenicity was evaluated. B-cell, HTL and CTL epitopes as well as IFN-ɣ, IL-17, and IL-10 inducers were predicted, manually mapped to the fragments and common regions were tabulated forming a peptide ensemble. The ensemble was evaluated for Class I MHC immunogenicity and toxicity. Further, immunogenic peptides were randomly selected and used to design vaccine constructs. Eight vaccine constructs were generated by linking random peptides with GS linkers. Synthetic TLR-4 agonist, RS09 was used as an adjuvant and linked with the constructs using EAAK linkers. The predicted population coverage of the constructs was ~99.8 % in the Indian as well as South Asian populations. The most antigenic, nontoxic, non-allergic construct was chosen for the prediction of secondary and tertiary structures. The 3D structures were refined and analyzed using Ramachandran plot and Z-scores. The construct was docked with TLR-4 receptor. Molecular dynamic simulation was performed to check for the stability of the docked complex. Comparative *in silico* immune simulation studies showed that the predicted construct elicited humoral and cell-mediated immunity in human host comparable to that elicited by Leish-F3, which is a promising vaccine candidate for human VL.

* Corresponding author.
E-mail addresses: sanjana.human@gndu.ac.in, sanjana.bhu@gmail.com (S. Mehrotra).

https://doi.org/10.1016/j.heliyon.2023.e22121
Received 12 July 2023; Received in revised form 2 November 2023; Accepted 5 November 2023
Available online 14 November 2023

1. Introduction

Visceral leishmaniasis (VL) is a severe clinical form of leishmaniasis and is fatal in most cases if untreated. It is a protozoal, vector-borne disease caused by two closely related species of *Leishmania*, *L. donovani* and *L. infantum*. An estimated 50,000 to 700,000 new human VL cases are reported with more than 50 % global burden from South Asia [1,2]. In India, the disease is endemic in Bihar, Jharkhand and West Bengal where it affects people from low socio-economic strata and rural communities [3]. Joint VL elimination initiative, launched in 2005, by the World Health Organization (WHO) and governments of India, Bangladesh and Nepal reduced the disease incidence, but complete elimination threshold (one case per 10,000 population) in Indian subcontinent is yet to be achieved [3]. The Possibility of transmission still exists due to asymptomatic infections, PKDL and human immunodeficiency virus (HIV)-VL coinfections. These can act as potential reservoirs of *Leishmania* and could be a possible cause of future epidemics. Further, the development of parasite resistance to common chemotherapeutic drugs: antimonials, miltefosine and ambisome has been reported in India [4]. In such a scenario, prevention remains the key to sustainable elimination and vaccines can complement leishmaniasis elimination campaigns. Even though several vaccine development strategies have been tested in animal models including live attenuated vaccines, DNA vaccines and recombinant vaccines, there are no prophylactic or therapeutic vaccines available against human VL (hVL). LEISH-F1 (formerly called Leish-111f), LEISH-F3, ChAd63-KHd and genetically modified live attenuated parasites ($LdCen^{-/-}$) created through CRISPR-Cas9 genome editing tools are some promising candidates under different stages of preclinical and clinical trials [5].

Since the traditional methods to design a vaccine are costly and time consuming, focus has now shifted to reverse vaccinology (RV) based computational approaches which rely on screening antigenic peptides from the genome and proteomes of pathogens. The *in silico* tools can reduce the list of potential epitope candidates for experimental testing and hence save time and cost. The predicted epitopes from different target proteins can then be combined to construct multi-epitope vaccines (MEVs), capable of eliciting potent immune responses in the host and are safer alternatives to classical vaccines [6]. Many computational studies have been performed to design MEVs for VL [7–10], however, most studies have focused on individual epitopes from parasite proteins rather than unique immunogenic regions which may possess overlapping epitopes thus enhancing epitope promiscuity. Epitope promiscuity enables the use of a minimal set of peptides to construct a vaccine that can target multiple HLA alleles and increases population coverage [11]. Since *Leishmania* is known to evade defense mechanism by inhibiting antigen cross priming [12], inclusion of peptides with promiscuous epitopes, capable of binding multiple HLA alleles, becomes important.

The present study employs an immunoinformatics approach to predict an epitope ensemble from *L. donovani,* which was further used to create a vaccine construct targeting the Indian subcontinent population. Then entire L. *donovani* proteome was filtered using *in silico* tools to find the most antigenic and immunogenic regions of the proteins, which are non-homologous to humans and contain T& B cell epitopes and cytokine inducers, unlike most studies that have attempted to predict either epitope [10,13,14]. Finally, peptides with promiscuous epitopes were selected to design a nontoxic, non-allergic vaccine construct that can elicit efficient immune response *in silico* against the South Asian population (specifically the Indian population).

2. Methodology

2.1. Sequence retrieval and classification of proteins

The complete sequences of all proteins of *L. donovani* BPK282A1 strain were accessed from the NCBI database (https://www.ncbi.nlm.nih.gov/) and further evaluated for the presence of *N*-terminal signal sequence, which is a hallmark of classically secreted proteins. SignalP-5.0 server (https://services.healthtech.dtu.dk/service.php?SignalP-5.0) was used to predict the secreted proteins. The server uses deep convolutional and recurrent neural network approaches to predict the presence of signal peptides and cleavage sites in archaea, bacteria and eukarya proteins [15]. To identify the subcellular location of secreted proteins, DeepLoc 1.0 server (https://services. healthtech.dtu.dk/service.php?DeepLoc-1.0) was used and predictions were made by choosing "profiles". This server is also based on deep neural network approach and predicts subcellular localization of eukaryotic protein in 10 different categories with good accuracy (78 % for 10 categories; 92 % for membrane-bound or soluble) [16]. Of all 10 localizations, proteins belonging to the cell membrane and extracellular locations were selected for subsequent analysis. The selected proteins were verified for their location based on the presence of either GPI anchor or transmembrane helix (none or no more than 1) or both through Pred-GPI (http://gpcr2.biocomp.unibo.it/predgpi/index.htm) and TOPCONS (https://topcons.%20net/pred/) respectively. TOPCONS is a consensus-based membrane protein topology prediction tool with an average performance of 80 % [17]. Pred-GPI, on the other hand, predicts the GPI anchor positions and omega cleavage sites using Support Vector Machine (SVM) and Hidden Markov Model (HMM) algorithms, with a low false positive prediction rate [18]. Completion status of the protein sequences was evaluated through Uniprot database (https://www.uniprot.org/) and proteomic evidence of proteins with completed sequences was manually searched in TriTrypDb database (https://tritrypdb.%20org/tritrypdb/app/).

2.2. Selection of antigenic protein fragments

Since an antigenic protein should be unique and foreign to the host, the homology of proteins was tested using NCBI BlastP tool (https://blast.ncbi.%20nlm.nih.gov/Blastp) against non-redundant human protein sequences. All proteins were analyzed and unique, non-homologous fragments $\geq$15 amino acids from each protein sequence were selected. Antigenicity of these protein fragments was

evaluated through VaxiJen 2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.%20html) under the category of "Parasite" at a threshold of 0.5. VaxiJen is an alignment independent antigenicity prediction server that classifies antigens based on physico-chemical properties. The classification made is based on the auto cross-covariance (ACC) transformation of protein sequences into uniform vectors of amino acid properties [19].

2.3. Prediction of the HTL, CTL and B cell epitopes

Before evaluation of immunogenicity, presence of MHC I binders, MHC II binders and B cell epitope in the selected fragments were assessed. Firstly, most common HLA alleles belonging to the South Asian region were obtained from the Allele Frequency Net Database (AFND) (https://academic.oup.com/nar/article/48/D1/D783/5624967). The choice of region selection was based on the target population for the present study. AFND is a freely available database containing frequency data of several immune system related gene polymorphisms [20]. MHC I binder prediction was made through NetMHCpan4.1 server (https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1). All predicted 9mers with percentile rank less than 2 restricted against the 30 Class I HLA alleles were selected. For HTL epitope prediction, all 15mers restricted against 11 HLA-DRB1 alleles with a percentile rank less than 10 were obtained through NetMHCIIpan 4.0 server (https://services.healthtech.dtu.dk/service.php?NetMHCIIpan-4.0). Both these evaluations were conducted before October 2020. For linear B cell epitope prediction, all antigenic sequences were subjected to IEDB Antibody Epitope Prediction tool (http://tools.iedb.org/bcell/). Bepipred Linear Epitope Prediction 2.0 was chosen as the prediction method and all peptides $\geq$5 amino acids at the default threshold (0.5) were considered B cell epitopes.

2.4. Prediction of the peptide ensemble

To evaluate immunogenic regions in the fragments, epitopes capable of inducing IFN-γ, IL-17 and IL-10 responses were searched. For this, the sequences were scanned in IFNepitope (https://webs.iiitd.edu.in/raghava/ifnepitope/index.php), IL-17escan (http://metagenomics.iiserb.ac.in/IL17eScan/index.html), and IL-10pred (http://crdd.osdd.net/raghava/IL-10pred/) servers to find respective inducers at default thresholds. The servers allow users to choose from different methods of prediction. While a hybrid approach was selected for IFN epitope, Dipeptide Composition (DPC) and SVM based methods, which also were the preset defaults, were selected for IL-17escan and IL-10pred servers, respectively. IFN epitope and IL-10 pred predictions are based on three approaches, SVM based, motif based and hybrid (SVM + Motif based) [21,22]. IL-17escan, on the other hand, is trained on composition-based models and machine learning methods [23]. All the epitopes obtained so far were mapped to antigenic fragments by using different color schemes for each. Since IL-10 is documented to play a role in disease progression [24], IL-10 non-inducing regions were focussed. Thus, overlapping regions of all epitopes and IL-10 non-inducers which were $\geq$5aa were highlighted and tabulated as peptide ensemble.

2.5. Refining of the peptide ensemble

The peptides obtained could be of different lengths, even lower than the minimum length required for an HTL epitope which is 15aa. Therefore, the shorter peptides were lengthened to reach at least 15aa by addition of few neighbouring amino acids. Finally, modified peptides were once again checked through IL-10pred and Class I Immunogenicity tool (http://tools.iedb.org/immunogenicity/) of IEDB to make sure all of them were IL-10 non-inducers and had MHC I immunogenicity, respectively. All the peptides lacking immunogenic MHC I binders were removed, and a final immunogenic peptide ensemble was obtained. Those were then evaluated for toxicity through ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) under default parameters. ToxinPred is an *in-silico* tool to predict toxicity of peptides and proteins based on dipeptide composition and hybrid models (dipeptide composition + motif composition), with 94.5 % accuracy of the former model [25].

2.6. Prediction of vaccine construct

The construction of multi epitope subunit vaccines requires suitable linkers for joining immunogenic peptides along with an adjuvant to enhance the immunogenicity. In this study, TLR4 agonist RS09, which is a lipopolysaccharide (LPS)-mimotope, was selected [26]. To conjugate an adjuvant with peptides, EAAK linker was used and for peptide-peptide conjugation Glycine-Serine (GS) linkers were used. The *N*- and C-terminal were also defined by GS linkers. The placement of the agonist and the order of peptides significantly affects the overall performance of the fusion constructs. Therefore, the peptides were placed randomly, and maximum possible conformations were tried to achieve the best model.

2.7. Analysis of primary structure

2.7.1. Homology, physicochemical properties and population coverage

To check if the constructs formed were non-homologous to humans and homologous to *Leishmania* species, homology was tested against non-redundant databases of *Homo sapiens* and *Leishmania* sp. using the NCBI BlastP tool (https://blast.ncbi.nlm.nih.gov/Blastp). Physicochemical properties of the constructs were checked through Expasy ProtParam tool (https://web.expasy.org/protparam/). Finally, population coverage of each construct was assessed for the Indian and South Asian populations through IEDB population coverage tool (http://tools.iedb.org/population/).

2.7.2. Toxicity, allergenicity and antigenicity

The toxicity and allergenicity of the constructs were estimated through ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) and AlgPred2.0 (https://webs.iiitd.edu.in/raghava/algpred2/index.html) respectively. AlgPred2.0, an updated version of AlgPred server, is trained on large dataset of allergens, non-allergens and IgE epitope database, and displays highest accuracy in a hybrid model (Machine Learning + BLAST + MERCI) [27]. Files were submitted in the recommended format and results were obtained with default parameter settings. Antigenicity was assessed through VaxiJen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) and Secret-AAR (http://microbiomics.ibt.unam.mx/tools/aar/index.php) tools. Secret-AAR tool allows assessment of the antigenic density of a protein and hence, evaluation of their antigenic potential [28]. Following construct submission, abundance of antigenic region (AAR) was calculated.

2.8. Prediction of secondary structure

Secondary structure of the protein constructs was predicted using Raptor-X Property online server (http://raptorx.uchicago.edu/StructurePropertyPred/predict/). The tool serves best especially for the proteins without close homologs in the Protein Data Bank (PDB) or with very sparse sequence profile. The server predicts based on machine learning model called Deep Convolutional Neural Fields and can predict solvent accessibility, secondary structure and disordered regions. The analysis was done to predict solvent accessible regions and 3-state secondary structure (SS3) elements: helices, globular regions, beta-turn regions, all of which are necessary features for protein stability [29].

2.9. Prediction of tertiary structure and its validation

To predict tertiary structure Raptor X Contact (http://raptorx.uchicago.edu/ContactMap/) was used. The server predicts distance-based protein folding which is based on deep convolutional neural network model [30]. The predicted models were refined using the GalaxyRefine (https://galaxy.seoklab.org/cgi-bin/submit. cgi?type = REFINE) web server [31]. The refined tertiary structures predicted were validated, based on Ramachandran plot analysis, through PROCHECK server (https://saves.mbi.ucla.edu/) and Swiss Structure Assessment tool (https://swissmodel.expasy.org/assess). Another validation web server, PRO-SA (https://prosa. services. came.sbg.ac.at/prosa.php) was also used to check the overall model quality in terms of Z-score and residue scores. Z-score describes the overall model quality by checking whether the score of the submitted protein falls within the range of scores found for native proteins of similar size. Residue scores depict the plotting energies as a function of amino acid at specific position such that positive scores correspond to flawed parts of the submitted protein [32].

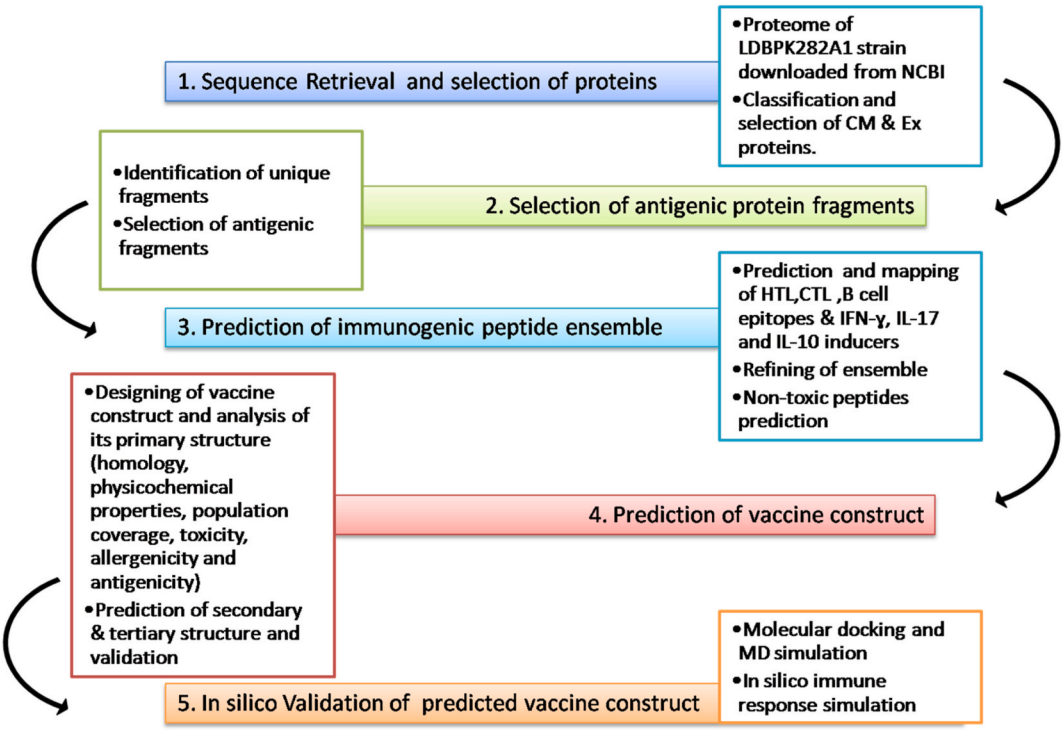


Fig. 1. Schematic representation of the workflow.

2.10. In silico validation of predicted vaccine construct

2.10.1. Molecular docking and molecular dynamic (MD) simulation

Activation of the antigen presenting cells (APCs) via Toll Like Receptors (TLRs) primes the innate immune system for efficient antigen processing and presentation. To this end, TLR ligands are widely exploited to enhance the efficacy of vaccine constructs. To analyze the interaction of the designed construct with TLR4, molecular docking was performed using Cluspro2.0 (https://cluspro.bu.edu/login.php). The tertiary structure of TLR4 (PDB id: 3FXI) was downloaded from Protein Data Bank (PDB) and submitted as "Receptor" file. The finalized vaccine construct was submitted as "Ligand" file. Cluspro2.0 performs rigid body docking based on Fast Fourier Transform (FFT) correlation approach, root mean square deviation (RMSD) based clustering of 1000 lowest energy models followed by the energy minimization refining of selected models [33,34]. The resulting models were ranked according to decreasing cluster size. The selected models were visualized using Pymol software (http://www.pymol.Org). Molecular docking was followed by molecular dynamic (MD) simulation to check for flexibility and stability of the complex. The simulation was carried out by GROMACS version 5.0 software and force field GROMOS96a. The stability of the complex was calculated in terms of temperature, pressure, RMSD and root mean square fluctuation (RMSF) of the complex.

2.10.2. In silico immune response simulation

To simulate the immune response generated by the vaccine construct in human host, C-ImmSimm (https://kraken.iac.rm.cnr.it/C-IMMSIM/) server was used. The predictions are based on the Position Specific Scoring Matrix (PSSM) in terms of T &B cell epitopes generated, and their interaction with the immune system [35,36]. For the simulations, default parameters and single injection was used with population specific HLA alleles. The primary amino acid sequence of Leish F3 and the predicted vaccine constructs were submitted to the server. The primary sequence of Leish-F3 was obtained from Ref. [37] (Supplementary File 3). A schematic representation of the workflow is given in Fig. 1.

3. Results

3.1. Classification of proteins

Reference proteome sequence of *Leishmania donovani* BPK282A1 was obtained from NCBI on June 2, 2020. A total of 8014 protein sequences were retrieved in FASTA file format and submitted to SignalP5.0 for classification. The secretome of a pathogen consists of all the secreted proteins. These proteins are vital for carrying out diverse biological processes including cell migration, cell adhesion, proliferation, differentiation and virulence , and facilitate interaction of the pathogen with the host cells. Due to the fact that the parasite secretome exhibits direct interaction with the host cells, these proteins in principle should make good candidates as immunogenic targets [38]. Using this rationale, in the present study, secretory proteins were chosen for further evaluation. 277 classically secreted proteins were obtained from the server, which were further subjected to DeepLoc 1.0 for classification based on their cellular location. Of 10 predicted localizations, 67 cell membrane proteins and 69 extracellularly located proteins were identified, which were further validated for the presence or absence of transmembrane helices (TMH) and GPI-anchors, the two differentiating characteristics of membrane associated and cleavable (soluble) proteins. Presence of either 1 TMH or GPI- anchor or both was used as a criterion for classifying membrane proteins. Proteins with >1 TMH in addition to a GPI anchor were excluded from the prioritized list due to potential difficulties perceived with their *in vitro* synthesis and purification [39]. Absence of both was a prerequisite for a protein to be classified as a soluble protein. 47 cell membrane proteins with one TMH and 52 extracellular proteins with no TMH were obtained from TOPCONS. Pred-GPI found 7 cell membrane proteins with GPI anchor and 62 extracellular proteins without GPI anchor. A total of 47 cell membrane proteins (with either one TMH or GPI anchor or both) and 48 soluble proteins (with neither TMH nor GPI anchor) were obtained.

Complete sequence information for *L. donovani* proteins was accessed from two of the most regularly updated databases, Uniprot and TriTrypDb. Complete sequence reads were obtained for 46 cell membrane and 37 extracellular proteins from Uniprot. However, two extracellular proteins identifiers, XP_003864652.1 and XP_024329258.1 shared same sequence and information, and hence were clubbed and used as XP_003864652.1, resulting in 36 soluble proteins. Of these, proteomic evidence was found for only 23 cell membrane and 12 extracellular proteins through TriTrypDb, listed in Supplementary Tables S1 and S2 and hence, were taken up for further analysis.

3.2. Selection of the antigenic protein fragments

After classification and validation of proteins for the mentioned class-associated features, homology with human proteins, was evaluated. NCBI BlastP server was used to find unique, ≥15 amino acids long *L. donovani* sequences, lacking similarity to identified hits at default threshold values against the non-redundant protein database of *Homo sapiens*. The length of unique regions was based on the HTL epitope length considered in this study, which is 15aa. The corresponding non-homologous sequences of proteins were cropped and saved. Proteins, for which no significant similarity was found, were included as such without any modification to the native sequence. Except for XP_003858308.1, all 22 cell membrane proteins presented >15aa non-homologous segments that were cropped and saved as 41 fragments. Similarly, 13 fragments were obtained from all 11 extracellular proteins, except XP_003858320.1. Further, VaxiJen server was used to check the antigenicity of the fragments obtained. Of all fragments, 27 cell membrane and 12 extracellular antigenic fragments were obtained with probability of antigenicity greater than 0.5.

3.3. Prediction of the immunogenic peptide ensemble

3.3.1. Prediction of HTL, CTL and B cell epitopes

To search for fragments containing epitopes restricted by both MHC classes, most frequent MHC alleles of South Asian population were extracted from AFND. A total of 30 HLA Class I alleles (10 each from HLA-A, HLA-B, HLA-C subset) and 11 HLA-DRB1 alleles, based on literature as well as AFND, were analyzed for epitope restriction (Table 1). The fragments were evaluated for presence of MHC Class I & II epitopes through NetMHCpan 4.1 and NetMHCIIpan 4.0 server respectively, against listed MHC alleles. All fragments had MHC epitopes except XP_003864707.1_1 extracellular protein fragment which lacked HTL epitope and hence was not considered for further analysis. The objective next in line was to identify fragments containing B cell epitope. Since the present study dealt with protein fragments, prediction of linear epitopes instead of conformational epitopes was considered owing to the fact that the latter might lie in the homologous regions which have been eliminated. Therefore, IEDB antibody epitope prediction tool was used for linear B cell epitope prediction, selecting all epitopes ≥5aa at the default threshold. All fragments contained B cell epitopes, except for XP_003865497.1_1 extracellular fragment, which was dropped for subsequent analysis. All the B-cell and T-cell epitopes of respective cell membrane and extracellular fragments were tabulated and saved for further evaluations.

3.3.2. Prediction of immunogenic peptides and epitope mapping

To check whether selected fragments can elicit immune response downstream of binding to MHC Class II, IFN-γ, IL-17 and IL-10 inducers were scanned in the fragments through IFNepitope, IL-17escan and IL-10pred servers respectively. IFN-γ inducers were found in all fragments, IL-10 inducers were found in all except XP_003862129.1_1 cell membrane fragment and IL-17 inducers were found in only one fragment (XP_003863771.1_1 cell membrane fragment). All inducers and B &T cell epitopes were mapped onto antigenic peptide fragments and regions ≥5 amino acids, having all the epitopes, inducers and IL-10 non-inducer were highlighted. Because IL-10 is known to facilitate disease progression and enhanced IFN-γ is known to promote anti-parasitic responses, the present study focused on regions that induced inflammatory IFN-γ response but not anti-inflammatory IL-10 response. All highlighted peptide sequences were tabulated to achieve a peptide ensemble, which consisted of 71 cell membrane and 22 extracellular peptides (Supplementary File S3).

3.3.3. Refining of the peptide ensemble

Since the ensemble has the MHC II binding regions capable of inducing cytokines downstream, they could be called as Class II immunogenic regions. The peptides in the ensemble ranged from 5aa to 52aa in length. The peptides smaller than 15aa in length were modified by addition of a few neighbouring amino acids to reach a length of at least 15aa such that they contain epitopes and inducers. While adding the neighbouring regions, it was also checked that MHC Class I epitopes in that region remain immunogenic. This evaluation was conducted using Class I Immunogenicity tool of IEDB. After modification, several peptides were found either to be closely present in the main protein sequence or overlapping; such peptides were clustered to form a single peptide, resulting in a total of 66 cell membrane peptides and 20 extracellular peptides. Finally, the modified ensemble was once again evaluated through IL-10pred server and Class I Immunogenicity tool to reconfirm their IL-10 non-inducing capacity as well as Class I immunogenicity. All peptides were IL-10 non-inducers, however, the ones lacking immunogenic epitopes were removed from the ensemble, thus, restricting the final number to a total of 43 cell membrane and 17 extracellular peptides in the ensemble. Finally, toxicity of each of these was evaluated through ToxinPred. Only Ex15 peptide derived from protein XP_003865232. 1 peptide was found to be toxic and hence was removed for further analysis. At last, a total of 43 cell membrane and 16 extracellular peptides, derived from 17 cell membrane and 6 extracellular proteins were obtained. The peptides along with their toxicity status are listed in Supplementary File S3. Fig. 2 shows a schematic diagram summarizing the screening of protein fragments for the non-toxic immunogenic protein ensemble.

3.4. Prediction of vaccine constructs

3.4.1. Core properties and structure of the vaccine constructs

Eight constructs were developed (Supplementary File S3), each comprising of different set of randomly arranged peptides linked with GSGGSG linkers. The TLR4 agonist RS09 was used as an adjuvant. Addition of adjuvant increases the vaccine efficacy and protection against variable pathogens by broadening the immune response [40]. The agonist was placed at *N*- terminal following GSG sequence and EAAK linker was used to link adjuvant with peptides. A schematic representation of vaccine construct design is given in Fig. 3A.

Table 1
List of HLA alleles selected for epitope restriction.

MHC Class I			MHC Class II
HLA-A	HLA-B	HLA-C	HLA-DRB1
01:01, 02:01, 02:11, 02:24, 03:01, 11:01, 24:02, 31:08, 33:01, 68:01	14:05, 18:07, 35:03, 40:06, 44:04, 44:06, 51:01, 51:10, 52:01, 57:01	03:02, 04:01, 05:09, 06:03, 07:01, 06:02, 07:02, 12:02, 14:02, 15:02	0301, 0101, 0701, 0403, 1001, 1101, 1301, 1404, 1501, 1502, 1504

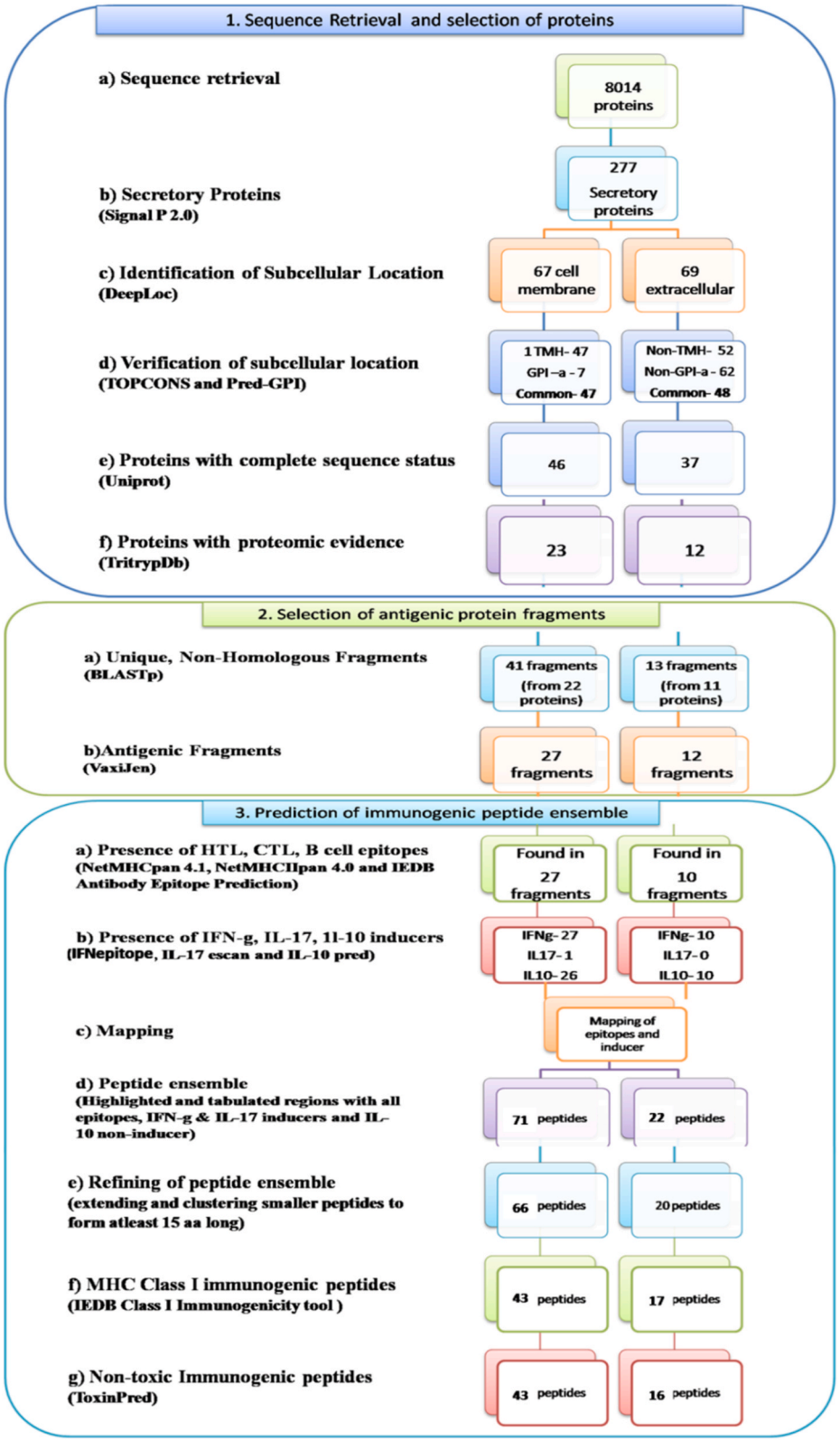


Fig. 2. Schematic representation of extracting non-toxic immunogenic peptide ensemble from whole proteome sequence of LDBPK282A.1. TMH- Trans membrane helix; GPI-a- Glycophostidyl anchored; non-GPI-a- non-glycophostidyl anchored.

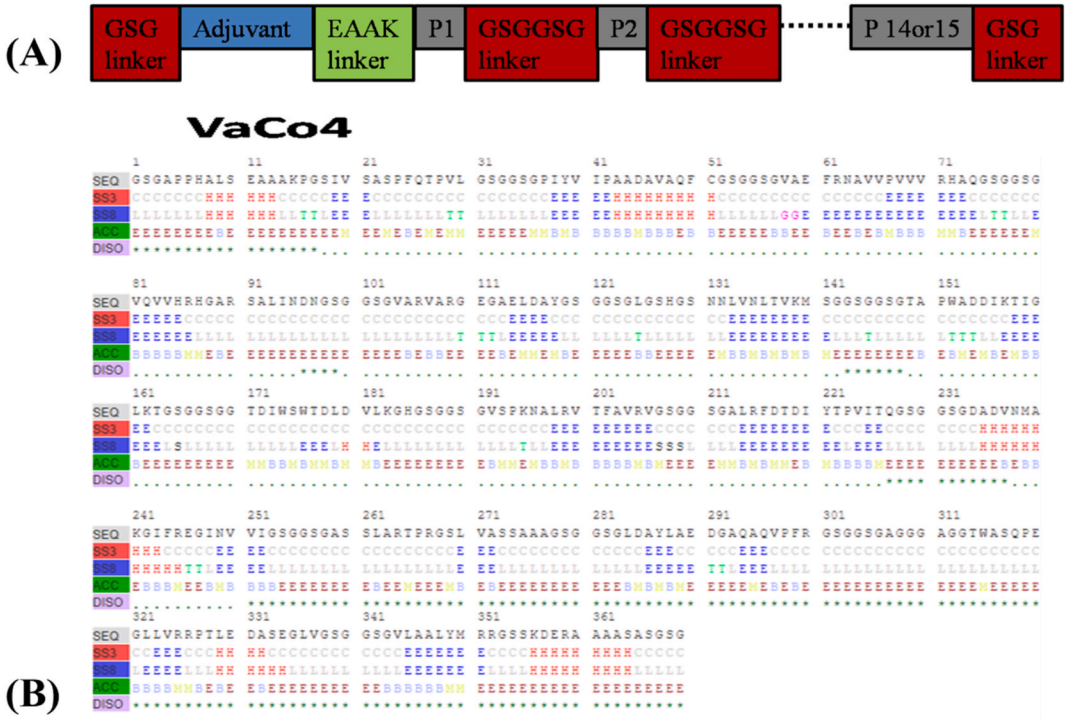


Fig. 3. (A) Representation of vaccine construct design; P1-Peptide1, P2- Peptide2, P14or 15-Peptide14 or15. (B) Secondary structure prediction of vaccine construct VaCo4. For SS3 predictions, H, E, and C represent alpha-helix, beta-sheet and coil, respectively. For solvent accessibility predictions, B, M and E represent Buried, Medium and Exposed respectively.

3.4.2. Analysis of vaccine constructs

3.4.2.1. Homology, physicochemical properties and population coverage. Vaccine sequence construction was followed by homology analysis using PSI-BLAST against non-redundant databases of *Homo sapiens* and *Leishmania*. On submitting the query sequence, no significant similarity was obtained with the human proteome, however, there was significant similarity with different leishmania species (including *L. infantum, L. major, L. braziliensis*) in addition to *L. donovani*. Next, physicochemical properties of all the constructs were listed using ProtParam server. The molecular weight of all the constructs was more than 30 kDa, indicating a higher probability of their immunogenicity. However, varied theoretical PI was obtained for VaCo3 and VaCo5 and therefore these were not considered for further evaluation. An instability index of less than 40 indicates a stable protein, which was not observed for VaCo6, therefore it was excluded. A higher aliphatic index (AI) indicates thermostability [41]; hence, Vaco2 and VaCo8 with low AI were also removed. GRAVY scores of all the constructs indicated their hydrophilic nature. Based on the results, only three vaccine constructs VaCo1, VaCo4, and VaCo7 were selected for further analysis.

3.4.2.2. Toxicity, allergenicity and antigenicity. Toxicity of the constructs was scanned using ToxiPred server and all three constructs were found to be non-toxic. AlgPred2 was used for allergenicity evaluation and VaCo7 was found to be an allergen, hence removed from further analysis. VaxiJen and Secret-AAR tools were used to estimate antigenicity of the constructs. Both VaCo1 and VaCo4 were found to be antigenic, however, VaCo4 was found to be more antigenic in both prediction tools, having higher probability and number of antigenic regions compared to VaCo1. Therefore, VaCo4 was taken up for further analysis. Analysis of all vaccine constructs has been tabulated in Table 2.

3.4.3. Prediction of secondary structure

To predict the secondary structure of the vaccine constructs, Raptor-X Property server was used. The VaCo4 construct consisted of **10 %** helix, **21 %** beta sheets, **68 %** coils. The server also predicted the solvent accessible and buried sequence percentages which were 57 %, and 26 % for VaCo4 respectively. Interestingly, the server predicted 6 of total 7 amino acids in agonist sequence of VaCo4 exposed to the solvent, indicating a higher probability of agonist interactions (Fig. 3B).

3.4.4. Prediction of tertiary structure and its validation

Prediction of tertiary structure was done through RaptorX server (Fig. 4A) which predicted five different models. They were refined using GalaxyRefine server (Fig. 4B). For each model, the first refined model predicted was saved. It is because according to the server, in model 1 structural changes are applied only to side-chain clusters, compared to the rest of the models where additional changes are

Table 2
Analysis of all eight designed vaccine constructs based on their properties.

Vaccine Construct	BLAST		Physicochemical properties						Population Coverage		Toxicity	Allergenicity	Antigenicity	
	Human	Leishmania	No. of amino acids	Molecular Weight	Theoretical PI	Instability Index	Aliphatic Index	GRAVY	India	South Asia			Antigenicity (VaxiJen)	Antigenic Regions (AAR)
VaCo1	No SS	SS	371	35109.96	6.78	36.36	60.65	−0.128	99.74	99.8	Nt	NAlg	0.9047	12 (31)
VaCo2	No SS	SS	372	35499.38	7.34	31.53	56.52	−0.183	99.74	99.8	/	/	/	/
VaCo3	No SS	SS	362	34408.2	9.58	34.48	57.49	−0.319	99.74	99.8	/	/	/	/
VaCo4	**No SS**	**SS**	**369**	**35202.57**	**7.2**	**28.61**	**73.04**	**−0.052**	**99.74**	**99.8**	**Nt**	**NAlg**	**0.955**	**14 (26.43)**
VaCo5	No SS	SS	376	36152.24	5.18	36.63	61.04	−0.129	99.74	99.8	/	/	/	/
VaCo6	No SS	SS	372	36031.79	8.03	40.65	49.62	−0.367	99.74	99.8	/	/	/	/
VaCo7	No SS	SS	361	34275.36	8.81	37.19	65.68	−0.125	99.74	99.8	Nt	Alg	/	/
VaCo8	No SS	SS	374	36038.66	6.43	29.69	56.71	−0.389	99.74	99.8	/	/	/	/

No SS–No Significant Similarity; SS- Significant Similarity; Nt- Non-toxic; NAlg- Non- Allergen; Bold-Properties of the vaccine construct chosen for the study.

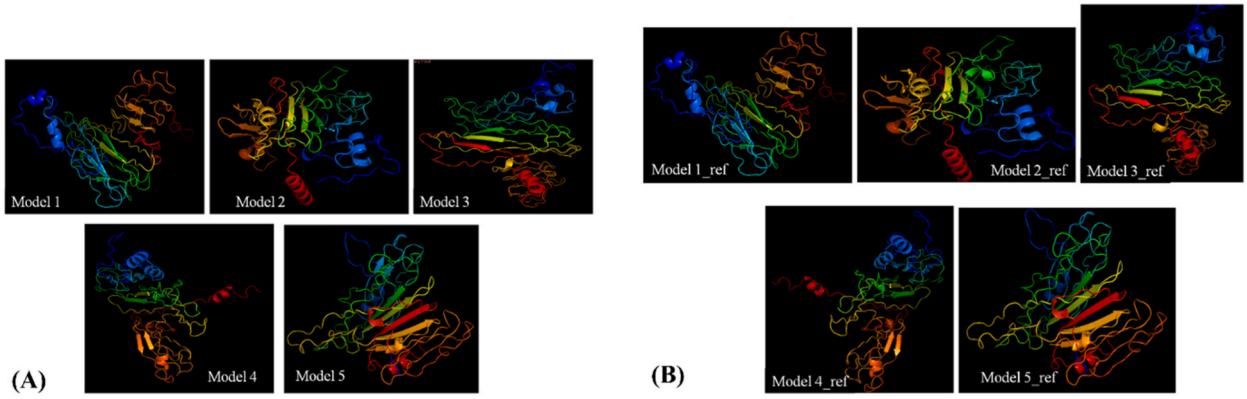
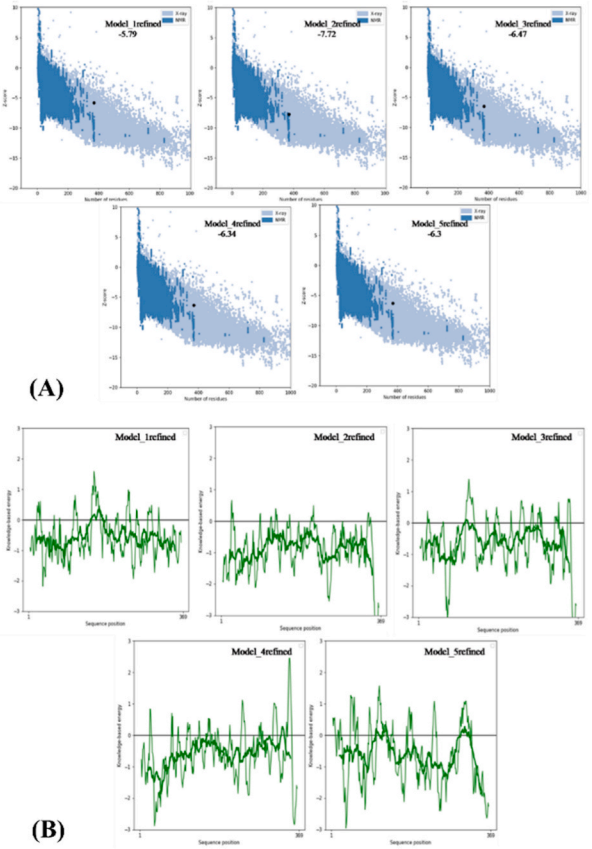


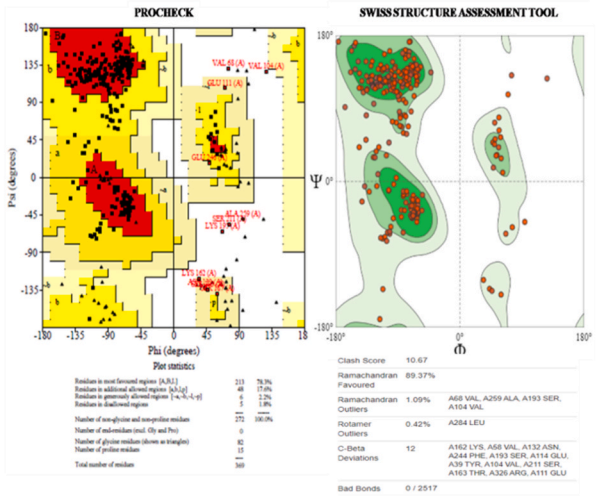
Fig. 4. Prediction of tertiary structures of VaCo4 and their refinement. (A) Tertiary structures generated for VaCo4 using RaptorX. (B) Each tertiary model obtained in previous step was refined using GalaxyRefine for VaCo4.

made to secondary structures and elements [31]. For validation, Ramachandran's score was calculated through ProCheck server and Swiss Structure Assessment tool. Model_2ref and Model_5ref were the best predicted models according to Procheck (Fig. 5C). However, according to Swiss Structure Assessment tool, the best models were Model_1ref and Model_4ref, although rest of the models also had comparable scores indicating all models were of good quality (Fig. 5C). According to ProSA web server, all models had Z-scores within permissible limits. However, Model_2ref obtained a Z-score of −7.72 which was in range of experimentally validated proteins through both sources (X-ray and NMR) (Fig. 5 A&C). Also, the local model quality graph also suggested this to be the best model (Fig. 5B). Therefore Model_2ref was selected for further analysis. Comparison of Ramachandran plot of Model_2ref obtained from Swiss



(C)

Model	Ramachandran Plot Analysis		Z-Score (ProSA)
	PROCHECK	Swiss Structure Assessment	
Model_1ref	77.6	92.92	-5.79
Model_2ref	78.3	89.37	-7.72
Model_3ref	77.9	89.65	-6.47
Model_4ref	77.6	91.83	-6.34
Model_5ref	80.5	91.55	-6.3



(D)

Fig. 5. Comparison of all refined models. Comparison of (A) Z-scores and (B) local model quality of all refined models using PRO-SA web server, respectively.(C) Comparison of Ramachandran Plot and Z-scores of all refined models. (D) Ramachandran Plot analysis of the selected refined model, which was VaCo4_Model_2ref, using PROCHECK and Swiss structure Assessment tool, respectively.

Structure Assessment tool and PROCHECK is shown in Fig. 5D.

3.5. In silico validation of predicted vaccine construct

3.5.1. Molecular Docking and Molecular Dynamic Simulation

Predictions were followed by molecular docking of the models with TLR4 to demonstrate how well the vaccine constructs interacted with TLR4. Murray and coworkers [42] reported that triggering TLR4 signaling upregulates macrophage antileishmanial activity, strengthened resistance and promotes effect of chemotherapy in the livers of *L. donovani* infected C57BL/6 mice. Hence, to this end, the tertiary structures of VaCo4 were docked against the tertiary structure of TLR4 using ClusPro2.0 server. On submitting the tertiary structure of the VaCo4 construct and TLR4, 29 docked complexes were obtained. The model with largest cluster size was ranked first by the server and hence was selected for further simulation. The complex structure had 42 members in the cluster, with the lowest binding energy of −718.3 and center energy of −641, suggesting good binding affinity. A clear interaction between the

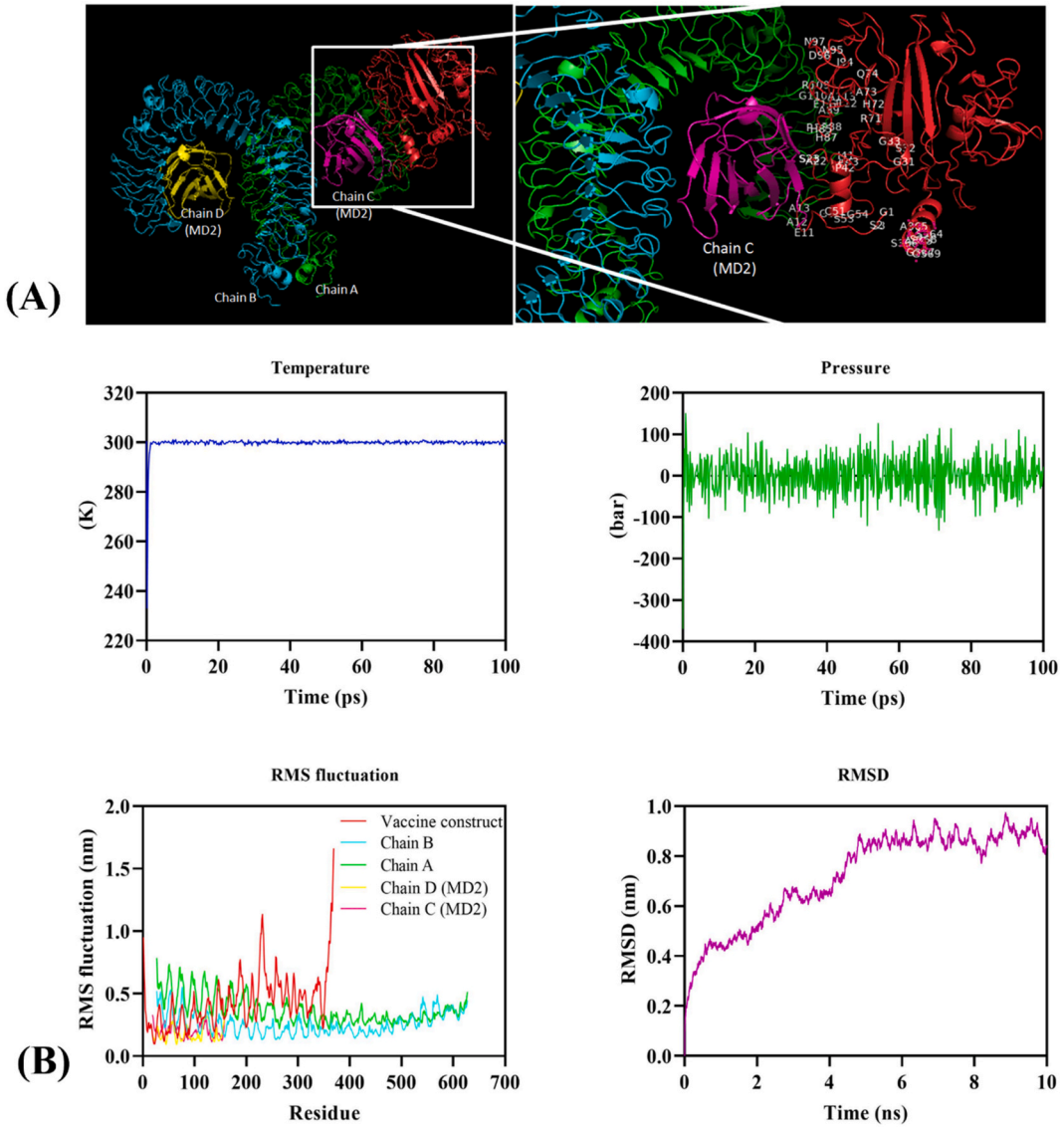


Fig. 6. Molecular Docking and Molecular Dynamic Simulation. (A) Molecular Docking of proposed vaccine construct VaCo4_Model_2ref (red) and TLR (green, cyan, yellow and magenta). Inset image depicts interaction of TLR chains A&C (MD2) (green and magenta) and *N*-terminal region of proposed vaccine construct. Few random residues from position 1–113 and C-terminal have been labeled in vaccine construct to ease visualization. (B) MD Simulation of VaCo4-TLR4 complex indicated through temperature, pressure, RMSF and RMSD plots respectively. RMSD plot shows protein stability. RMSF plot indicates great variation indicating flexible protein. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

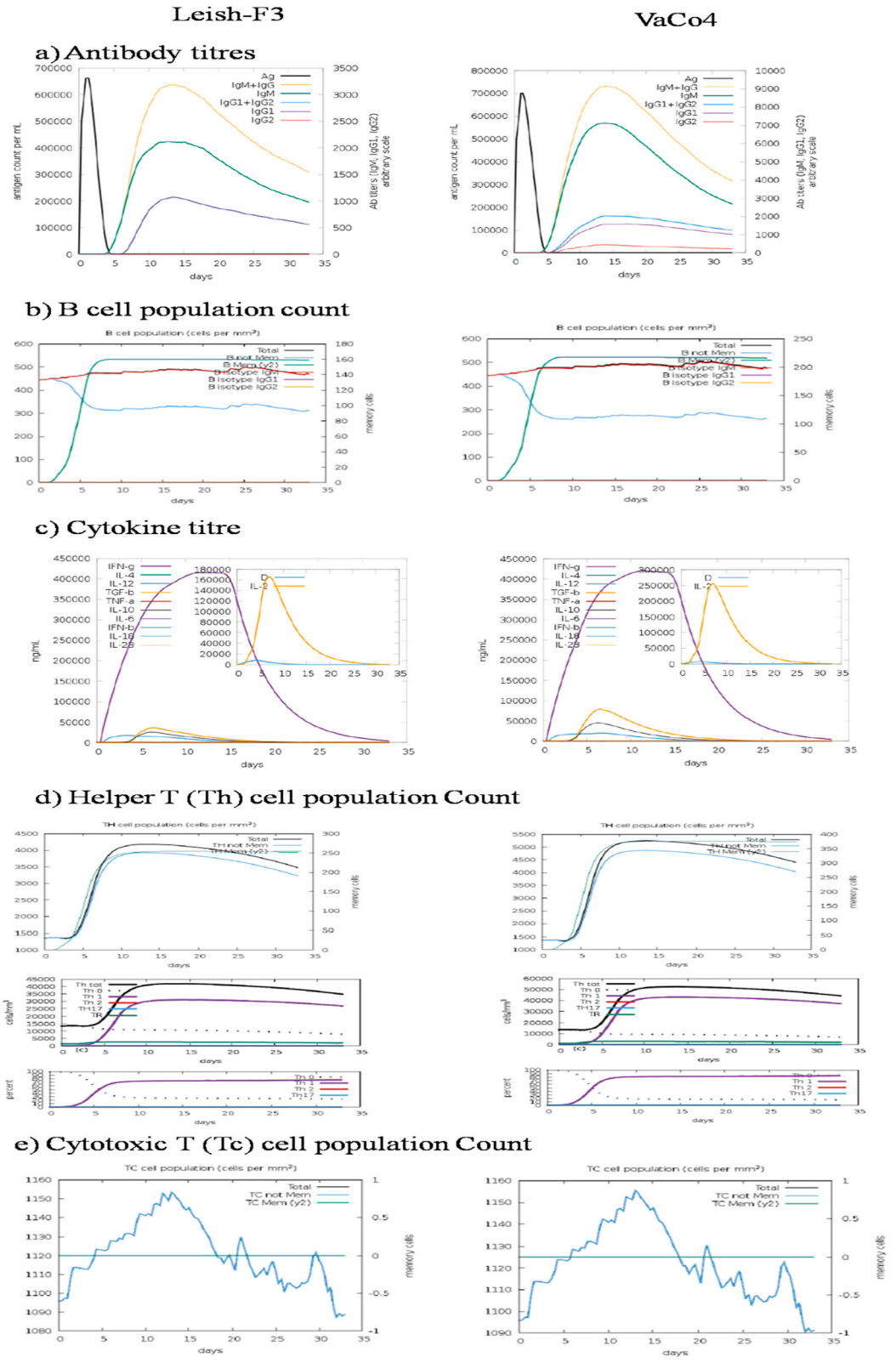


Fig. 7. Comparison of immune response elicited in response to predicted vaccine constructs Leish F3 (positive control) and VaCo4. Plots in (a) depicts comparison of antibody titres, (b) depicts comparison of B cell count; (c) depicts comparison of cytokine titre, (d) depicts comparison of

helper T Cell population and (e) depicts comparison of cytotoxic T cell population elicited in response to control and predicted vaccine constructs, respectively.

N-terminal region of VaCo4 was visualized with TLR4 chains A and C (MD2) (Fig. 6A).

The next step was to look at the stability and flexibility profile of the complex, for which molecular dynamic simulation was performed for 10 ns (nanoseconds) on Groningen Machine for chemical simulations GROMACS version 5.0 The constancy, flexibility, and folding was determined through temperature, pressure, RMSD, and RMSF plots generated in the software. The RMSD indicated protein stability after 5 ns for VaCo4 whereas RMSF graph suggested a flexible and thermostable complex. The fluctuations of the amino acids ranged from 0.25 to 0.75 nm for VaCo4 (Fig. 6B), and got stable after 350 residues. Moreover, at constant temperature and pressure, the complex was stable.

3.5.2. In silico immune response simulation

C-ImmSimm server was used to check the immune response generated by the vaccine constructs. Comparative analysis was run with Leish-F3 and single injection was placed with no adjuvant. Immune response was generated against homozygous HLA alleles (HLA-A*1101, HLA-B*4006 and HLA-DRB1*0701). Activation of humoral immunity in response to the predicted construct was evident by the rise of antibody titers and B cell count (Fig. 7 a&b). It was found that in response to the positive control Leish- F3 the IgM + IgG, IgM and IgG1 titers were raised whereas for the predicted construct, in addition to the mentioned classes, IgG1+ IgG2 and IgG2 were also found to be raised. For VaCo4, IgM, IgM + IgG titers (7000 and 9000 on arbitrary scale respectively) were maximally recorded. Further, higher numbers of B cells were found to be activated in response to the predicted construct as compared to Leish-F3, with major rise in memory B cells up to 200–250 cells/mm^3 as compared to 160 cells/mm^3 in case of the positive control. The up-regulated memory B cell numbers lasted for 35 days which also was the entire simulation period (Fig. 7 a&b).

In addition to humoral immune responses, activation of cell-mediated immunity was also evaluated. It was found that roughly equal levels of IFN-ɣ were produced against both vaccines (between $4.0x10^5$-$4.5x10^5$ ng/ml), however, IL-2 and anti-inflammatory TGF-β were produced in higher titers in response to the predicted constructs as compared to the positive control (IL-2: $2.5x10^5$-$3.0x10^5$ ng/ml as compared to $1.6x10^5$ ng/ml and TGF-β: $0.5x10^5$- $1.0x10^5$ ng/ml as compared to 0–$0.5x10^5$ ng/ml respectively). Besides, IL-10 and IL-12 were produced sub-optimally below $0.5x10^5$ ng/ml in both tested vaccines (Fig. 7c). Similarly, for helper T cells, relatively higher population was activated in response to the predicted constructs (5000–5500 cells/mm^3) as compared to Leish-F3 (4000–4500 cells/mm^3), of which 80 % population comprised of Th1 cells (Fig. 7d). Nevertheless, no major difference was observed with respect to the cytotoxic T cell population in both the tested constructs (1150–1160 cells/mm^3). The peak values mentioned above were all observed at 10 days post injection, however none of them dropped to nil even by the end of the simulation period (Fig. 7e).

The results observed indicated generation of a robust immune response against the *in silico* designed vaccine constructs.

4. Discussion

The conventional methods of vaccine development are time-consuming and can take up to a decade [43]. Reverse Vaccinology (RV) is a widely used approach to identify potential vaccine candidates (PVCs) by screening the genome or proteome of a pathogen through computational analyses [44]. One of the earliest reports on the application of immunoinformatics to develop an anti-leishmanial vaccine made use of a T cell predictive algorithm (AMPHI) with parasite protein gp63 as the candidate [45]. Since then, there has been a consistent improvement in the accuracy and thereby, reliability of these algorithms, enabling prediction of proteome-wide peptide epitopes [46,47]. The present study aimed to find the most immunogenic regions across the *L. donovani* proteome using immunoinformatic tools and to design vaccine constructs using a combination of those epitopes.

The L. *donovani* BPK282.A1 proteome was classified and secretory proteins were selected for candidate analysis. The *Leishmania* parasite, which undergoes morphological transitions in the two hosts, releases a set of secretory proteins in a temporal or life cycle stage-specific manner; this facilitates adaptation to the changing surroundings and thereby the survival of the parasite. Exploring such proteins and understanding the metabolic pathways they are involved in could help identify better candidates for the development of prophylactic vaccines [48]. Additionally, several *Leishmania* proteins which are located at the plasma membrane undergo processing and are shed subsequently. A few of these are also involved in facilitating the interaction of the parasite with the host and modulation of host's immune responses [38]. Several published reports on vaccine constructs have focused on previously characterized secretory proteins of *Leishmania* [7,13,37,49], but in the present study, the whole *L. donovani* proteome was classified based on their sub-cellular location; the proteins found to be localized in the extracellular region or cell membrane were selected as initial pool for analysis. The selected subcellular locations were further verified by the presence or absence of a GPI- anchor or TM-helix. Further, in contrast to Singh *et al* [9] who presented an epitope ensemble by screening epitopes from 4 antigenic *L. donovani* proteins, the present study investigated whether the reported protein sequence is complete and if any protein expression data is available for the selected proteins, using Uniprot and TriTrypDb databases respectively. This approach was in line with a similar study by Khan *et al* [13] which used mass spectrometry (MS)-derived data on comparative abundance of proteins in promastigotes and amastigotes for epitope prediction.

Non-homology or foreignness of an antigen increases its immunogenicity and prevents the generation of autoimmune responses [50–52]. However, instead of outrightly rejecting the homologous proteins, as demonstrated in several past reports [53–57], the current approach was to carefully read along protein's full length and crop unique regions, which presented no homology to non-redundant human protein sequences using NCBI BLASTp. The rationale for the application of this approach lies in the fact that any foreign protein is processed into shorter peptides before being presented for binding to a B- or T-cell receptor, and therefore any

resulting peptide showing non-homology to human proteins may in principle be antigenic. These short antigenic peptides are called epitopes [58,59]. In addition to the proteins with "no significant similarity", non-homologous fragments ≥15 aa in length were also selected from each protein. All such peptides were termed as fragments and evaluated for antigenicity using VaxiJen 2.0.

NetMHCpan-4.1 and NetMHCIIpan-4.0 were used for predicting MHC I & II binders (T cell epitopes) respectively. These were preferred over other prediction servers such as NetMHC, NetMHCpan, SYFPEITHI, as NetMHCpan-4.1 and NetMHCIIpan-4.0 are trained by integrating binding affinity (BA) and eluted ligand (EL) data derived from mass spectrometry (MS) experiments, allowing the capability of annotating sequences to single MHC restrictions, and thereby offering more accurate and reliable predictions [60,61]. B cell epitopes were predicted using Antibody Prediction tool of IEDB server. Although several RV based *in silico* studies prefer conformational B cell epitopes [62–65], the current study sought linear epitopes only. As mentioned before, one obvious reason was the expected loss of conformational epitopes in the cropped sequences [66] and the fact that the predictive accuracy of conformational epitopes for shorter peptides is not great. The predicted epitopes were manually mapped to the fragments.

Analysis of immunogenicity, which refers to the capability of a peptide to trigger downstream immune responses when presented to T cells and B cells, was done by prediction and mapping of IFN-ɣ, IL-17 and IL-10 cytokine inducing regions. Unlike most studies, which used online servers to predict the immunogenicity potential of selected epitopes or vaccine constructs [13,49,67–69], the present study sought for immunogenic regions in the fragments, which were manually mapped to fragment sequences. This approach avoided the loss of fragments due to sequential mapping and ensured inclusion of the maximum possible number of fragments which were immunogenic for Class I & Class II and qualified for inducers. Non-immunogenic fragments were removed, and toxicity evaluation was performed *in silico* to identify candidate peptides for the next steps. The pool of peptide fragments thus obtained was termed as immunogenic ensemble, comprising of peptides which were antigenic, immunogenic (positive for the presence of MHC Class I, Class II and B cell epitopes and could also induce T cell responses), non-toxic and non-homologous (unique) to the human proteome, indicating high probability of their success as vaccine candidates.

Finally, peptides from the immunogenic ensemble were used to design MEV constructs against VL using a RV approach. These MEVs have certain advantages over the monovalent vaccines in terms of potency to cumulatively raise innate, humoral, and cellular immune responses [71]. Khatoon et al. [7], Vakili *et al* [8], Khatoon *et al* [70] and Singh et al. [65], have previously reported the theoretical potential of *in silico* designed vaccines for VL. While the studies by Khatoon *et al* [7] and Singh *et al* [65] largely utilized available genomic databases of *L. donovani* to select vaccine targets, Vakili *et al* [8] extracted epitopes from four antigenic *Leishmania* proteins and constructed a MEV. The chimeric construct was administered in mice and a robust Th1-type immune response was observed as shown by increased levels of IL-2, IFN-γ, TNF-α, and IgG2a, demonstrating the immunogenic potential of computationally identified candidates [8].

In this study, the predicted immunogenic fragments were linked through the most common flexible linker GSGGSG to allow appropriate flexibility [49]. To enhance the immunogenicity of the vaccine, TLR-4 agonist (RS09) was linked to the fragment combination using EAAK linker. Eight different constructs were formed and subjected to homology analyses. Uniqueness of an antigen not only increases its immunogenicity but also prevents molecular mimicry which is a significant mechanism causing autoimmunity. While homology analyses of the shortlisted peptides against the human proteome revealed no significant similarity, rendering the autoimmunity inducing potential of these fragments highly unlikely, the fragments exhibited significant similarity to the proteome from closely related species of the Leishmania parasite (*L. infantum*, *L. major*, *L. braziliensis*) in addition to *L. donovani*. However, in the background of extensive genetic diversity of *L. donovani*, experimental validation on the efficacy of the constructs remains to be demonstrated. Based on their physicochemical properties like molecular weight, GRAVY index, theoretical pI, instability index and aliphatic index, the construct VaCo4 was selected for further analysis. Computational analysis predicted the vaccine constructs to be thermostable, antigenic, non-toxic, and non-allergic. Furthermore, a population coverage of ~99.8 % was found in the Indian as well as South Asian region which was significantly better as compared to other reported studies. As next steps, tertiary structures of the constructs were predicted and refined; the models were validated based on the Ramachandran score, Z-score and the local model quality using a step wise approach to identify the models with the most favourable tertiary structure. Predictions were followed by molecular docking of the models with TLR4, which showed good binding energy values. MD simulation predicted good protein stability and a flexible and thermostable complex. Finally, a comparative *in silico* immune simulation was carried out and the constructs showed promising results in terms of their capacity of eliciting humoral and cell mediated immunity as well as memory cell response in the human host as compared to that elicited by Leish-F3 (control), a promising vaccine candidate for human VL.

Despite the inclusion of multiple important parameters for design, testing and careful considerations around additional aspects such as toxicity and allergenicity, experimental validation data will be required to prove the efficacy of the proposed constructs. Beyond the design of the recombinant constructs, additional components, e.g., TLR ligands, adjuvants and the overall formulation may also be of critical importance to reach the desired efficacy range and the optimal safety profile. At the level of construct design, the present approach has accounted for induction of autoimmunity due to molecular mimicry. However, multiple instances of autoimmune syndrome induced by adjuvants (ASIA) are reported [72], which should be seriously considered during *in vivo* validation. Even in the absence of an established causal relationship, adjuvants may contribute to autoimmunity with variably long latency periods and depending on factors like co-infection and the distinct immunological background of a vaccinated individual. There may be rarer or less frequently reported conditions of epitope spreading, or of polyclonal B-cell activation in response to non-protein antigens like bacterial lipopolysaccharide, leading to induction of autoimmunity [72,73]. Owing to multifactorial and at times idiopathic causation, it remains a major challenge to predict and account for such relatively rarer outcomes while computationally designing vaccine constructs. Notwithstanding, based on the outcomes received during the *in vivo* testing phase, one or several components in the formulation could be tweaked to obtain optimal safety and efficacy profile.

5. Conclusion

The study proposes a multi-subunit vaccine based on a combination of select immunogenic sequences from the *Leishmania* proteome. Unlike most of the previously reported studies targeting already characterized antigenic proteins, we have screened the entire *Leishmania* proteome to build a peptide ensemble for epitope selection, which is a major strength of the study. Another highlight of the study is the approach taken up to predict immunogenic ensemble. To the best of our knowledge, it is novel in finding unique non-homologous segments followed by prediction of epitopes and inducers, and their manual mapping to the protein fragments. The predicted peptides were combined to construct a potential vaccine candidate for hVL, which is non-toxic, non-allergic, antigenic as well as immunogenic, and has a stable and legitimate tertiary structure. Molecular docking and molecular dynamic simulation studies demonstrate good binding affinity with TLR-4 and formation of a stable complex for the vaccine constructs. Thus, it can be concluded that being a multi-peptide fusion construct with good predicted stability, existing standard protocols for protein expression and purification can be adopted for manufacturing the proposed vaccine construct and testing experimentally in *vitro* and *in vivo* settings.

Consent for publication

Not applicable.

Ethical approval

The present study does not involve human and animal samples.

Data availability statement

The data associated with our study has not been deposited into a publicly available repository. It is because all data generated or analyzed during the study are included in the submitted manuscript and supplementary data. All data generated or analyzed during the study are included in the submitted manuscript and supplementary data.

CRediT authorship contribution statement

Manu Kupani: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. **Rajeev Kumar Pandey:** Methodology, Writing – review & editing, Validation. **Sharad Vashisht:** Methodology. **Satyendra Singh:** Software. **Vijay Kumar Prajapati:** Software. **Sanjana Mehrotra:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Rajeev Kumar Pandey reports a relationship with Thermo Fisher Scientific that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the valuable suggestions of Henrik Nielson, Associate Professor, Department of Health Technology, Technical University of Denmark. We are also thankful for the financial assistance given by Department of Science and Technology, Science & Engineering Research Board (DST-SERB), Grant No. SB/YS/LS-140/2014 and Department of Science and Technology, Promotion of University Research and Scientific Excellence (DST-PURSE), Grant No. SR/PURSEPhase 2/18 (G). Also, MK would like to acknowledge CSIR for senior research fellowship (CSIR-SRF(Direct)) provided to her.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e22121.

References

[1] S. Burza, S.L. Croft, M. Boelaert, Leishmaniasis, Lancet 392 (2018) 951–970, https://doi.org/10.1016/S0140-6736(18)31204-2.
[2] M. Boelaert, F. Meheus, A. Sanchez, S.P. Singh, V. Vanlerberghe, A. Picado, S. Sundar, The poorest of the poor: a poverty appraisal of households affected by visceral leishmaniasis in Bihar, India, Trop. Med. Int. Health 14 (6) (2009) 639–644, https://doi.org/10.1111/j.1365-3156.2009.02279.x.

[3] S. Sundar, A. Singh, Recent developments and future prospects in the treatment of visceral leishmaniasis, Therapeutic advances in infectious disease 3 (3–4) (2016) 98–109, https://doi.org/10.1177/2049936116646063.

[4] S. Sundar, O.P. Singh, J. Chakravarty, Visceral leishmaniasis elimination targets in India, strategies for preventing resurgence, Expert Rev. Anti-infect. Ther. 16 (11) (2018) 805–812, https://doi.org/10.1080/14787210.2018.1532790.

[5] G. Volpedo, T. Pacheco-Fernandez, P. Bhattacharya, T. Oljuskin, R. Dey, S. Gannavaram, H.L. Nakhasi, Determinants of innate immunity in visceral leishmaniasis and their implication in vaccine development, Front. Immunol. 12 (2021), 748325, https://doi.org/10.3389/fimmu.2021.748325.

[6] Y. Zhang, S. Liang, S. Zhang, S. Zhang, Y. Yu, H. Yao, G. Liu, Development and evaluation of a multi-epitope subunit vaccine against group B Streptococcus infection, Emerg. Microb. Infect. 11 (1) (2022) 2371–2382, https://doi.org/10.1080/22221751.2022.2122585.

[7] 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 (1) (2017) 1–12, https://doi.org/10.1038/s41598-017-08842-w.

[8] B. Vakili, M. Eslami, G.R. Hatam, B. Zare, N. Erfani, N. Nezafat, Y. Ghasemi, Immunoinformatics-aided design of a potential multi-epitope peptide vaccine against Leishmania infantum, Int. J. Biol. Macromol. 120 (2018) 1127–1139, https://doi.org/10.1016/j.ijbiomac.2018.08.125.

[9] G. Singh, M. Pritam, M. Banerjee, A.K. Singh, S.P. Singh, Genome based screening of epitope ensemble vaccine candidates against dreadful visceral leishmaniasis using immunoinformatics approach, Microb. Pathog. 136 (2019), 103704, https://doi.org/10.1016/j.micpath.2019.103704.

[10] D.P. Lage, P.A. Ribeiro, D.S. Dias, D.V. Mendonça, F.F. Ramos, L.M. Carvalho, M. Christodoulides, A candidate vaccine for human visceral leishmaniasis based on a specific T cell epitope-containing chimeric protein protects mice against Leishmania infantum infection, npj Vaccines 5 (1) (2020) 75, https://doi.org/10.1038/s41541-020-00224-0.

[11] W. Fleri, S. Paul, S.K. Dhanda, S. Mahajan, X. Xu, B. Peters, A. Sette, The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design, Front. Immunol. 8 (2017) 278, https://doi.org/10.3389/fimmu.2017.00278.

[12] D. Matheoud, N. Moradin, A. Bellemare-Pelletier, M.T. Shio, W.J. Hong, M. Olivier, A. Descoteaux, Leishmania evades host immunity by inhibiting antigen cross-presentation through direct cleavage of the SNARE VAMP8, Cell Host Microbe 14 (1) (2013) 15–25, https://doi.org/10.1016/j.chom.2013.06.003.

[13] M.A.A. Khan, J.Q. Ami, K. Faisal, R. Chowdhury, P. Ghosh, F. Hossain, D. Mondal, An immunoinformatic approach driven by experimental proteomics: in silico design of a subunit candidate vaccine targeting secretory proteins of Leishmania donovani amastigotes, Parasites Vectors 13 (2020) 1–21, https://doi.org/10.1186/s13071-020-04064-8.

[14] D.P. Lage, P.A. Ribeiro, D.S. Dias, D.V. Mendonça, F.F. Ramos, L.M. Carvalho, E.A. Coelho, Liposomal formulation of ChimeraT, a multiple T-cell epitope-containing recombinant protein, is a candidate vaccine for human visceral leishmaniasis, Vaccines 8 (2) (2020) 289, https://doi.org/10.3390/vaccines8020289.

[15] J.J. Almagro Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat. Biotechnol. 37 (4) (2019) 420–423, https://doi.org/10.1038/s41587-019-0036-z.

[16] J.J. Almagro Armenteros, C.K. Sønderby, S.K. Sønderby, H. Nielsen, O. Winther, DeepLoc: prediction of protein subcellular localization using deep learning, Bioinformatics 33 (21) (2017) 3387–3395, https://doi.org/10.1093/bioinformatics/btx431.

[17] K.D. Tsirigos, C. Peters, N. Shu, L. Käll, A. Elofsson, The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides, Nucleic Acids Res. 43 (W1) (2015) W401–W407, https://doi.org/10.1093/nar/gkv485.

[18] C. Rita, M. Pier, P. Andrea, PredGPI: a GPI-anchor predictor, BMC (Biomed. Chromatogr.) (2008), https://doi.org/10.1186/1471-2105-9-392.

[19] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinf. 8 (1) (2007) 1–7, https://doi.org/10.1186/1471-2105-8-4.

[20] F.F. Gonzalez-Galarza, A. McCabe, E.J.M.D. Santos, J. Jones, L. Takeshita, N.D. Ortega-Rivera, A.R. Jones, Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools, Nucleic Acids Res. 48 (D1) (2020) D783–D788, https://doi.org/10.1093/nar/gkz1029(2020).

[21] S.K. Dhanda, P. Vir, G.P. Raghava, Designing of interferon-gamma inducing MHC class-II binders, Biol. Direct 8 (1) (2013) 1–15, https://doi.org/10.1186/1745-6150-8-30.

[22] G. Nagpal, S.S. Usmani, S.K. Dhanda, H. Kaur, S. Singh, M. Sharma, G.P. Raghava, Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential, Sci. Rep. 7 (1) (2017), 42851, https://doi.org/10.1038/srep42851.

[23] S. Gupta, P. Mittal, M.K. Madhu, V.K. Sharma, IL17eScan: a tool for the identification of peptides inducing IL-17 response, Front. Immunol. 8 (2017) 1430, https://doi.org/10.3389/fimmu.2017.01430.

[24] S. Nylén, D. Sacks, Interleukin-10 and the pathogenesis of human visceral leishmaniasis, Trends Immunol. 28 (9) (2007) 378–384, https://doi.org/10.1016/j.it.2007.07.004.

[25] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, Open Source Drug Discovery Consortium, G.P. Raghava, In silico approach for predicting toxicity of peptides and proteins, PLoS One 8 (9) (2013), e73957, https://doi.org/10.1371/journal.pone.0073957.

[26] A. Shanmugam, S. Rajoria, A.L. George, A. Mittelman, R. Suriano, R.K. Tiwari, Synthetic Toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants, PLoS One 7 (2) (2012), e30839, https://doi.org/10.1016/j.it.2007.07.004.

[27] N. Sharma, S. Patiyal, A. Dhall, A. Pande, C. Arora, G.P. Raghava, AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes, Briefings Bioinf. 22 (4) (2021) bbaa294, https://doi.org/10.1093/bib/bbaa294.

[28] F. Cornejo-Granados, J.M. Hurtado-Ramírez, R. Hernández-Pando, A. Ochoa-Leyva, Secret-AAR: a web server to assess the antigenic density of proteins and homology search against bacterial and parasite secretome proteins, Genomics 111 (6) (2019) 1514–1516, https://doi.org/10.1016/j.ygeno.2018.10.007(2019).

[29] S. Wang, W. Li, S. Liu, J. Xu, RaptorX-Property: a web server for protein structure property prediction, Nucleic Acids Res. 44 (W1) (2016) W430–W435, https://doi.org/10.1093/nar/gkw306.

[30] S. Wang, S. Sun, Z. Li, R. Zhang, J. Xu, Accurate de novo prediction of protein contact map by ultra-deep learning model, PLoS Comput. Biol. 13 (1) (2017), e1005324, https://doi.org/10.1371/journal.pcbi.1005324.

[31] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (W1) (2013) W384–W388, https://doi.org/10.1093/nar/gkt458.

[32] 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 (suppl_2) (2007) W407–W410, https://doi.org/10.1093/nar/gkm290.

[33] D. Kozakov, D.R. Hall, B. Xia, K.A. Porter, D. Padhorny, C. Yueh, S. Vajda, The ClusPro web server for protein–protein docking, Nat. Protoc. 12 (2) (2017) 255–278, https://doi.org/10.1038/nprot.2016.169.

[34] I.T. Desta, K.A. Porter, B. Xia, D. Kozakov, S. Vajda, Performance and its limits in rigid body protein-protein docking, Structure 28 (9) (2020) 1071–1081, https://doi.org/10.1016/j.str.2020.06.006.

[35] F. Castiglione, D. Deb, A.P. Srivastava, P. Liò, A. Liso, From infection to immunity: understanding the response to SARS-CoV2 through in-silico modeling, Front. Immunol. 12 (2021), 646972, https://doi.org/10.3389/fimmu.2021.6469.

[36] N. Rapin, O. Lund, M. Bernaschi, F. Castiglione, Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system, PLoS One 5 (4) (2010), e9862, https://doi.org/10.1371/journal.pone.0009862.

[37] R. Ojha, R.K. Pandey, V.K. Prajapati, Vaccinomics strategy to concoct a promising subunit vaccine for visceral leishmaniasis targeting sandfly and leishmania antigens, Int. J. Biol. Macromol. 156 (2020) 548–557, https://doi.org/10.1016/j.ijbiomac.2020.04.097.

[38] S. Ranganathan, G. Garg, Secretome: clues into pathogen infection and clinical applications, Genome Med. 1 (11) (2009) 1–7, https://doi.org/10.1186/gm113.

[39] A. Olaya-Abril, I. Jimenez-Munguia, L. Gomez-Gascon, I. Obando, M.J. Rodriguez-Ortega, Identification of potential new protein vaccine candidates through pan-surfomic analysis of pneumococcal clinical isolates from adults, PLoS One 8 (7) (2013), e70365, https://doi.org/10.1371/journal.pone.0070365.

[40] M. Luchner, S. Reinke, A. Milicic, TLR agonists as vaccine adjuvants targeting cancer and infectious diseases, Pharmaceutics 13 (2) (2021) 142, https://doi.org/10.3390/pharmaceutics13020142.

[41] A. Ikai, Thermostability and aliphatic index of globular proteins, J. Biochem. 88 (6) (1980) 1895–1898.

[42] H.W. Murray, Y. Zhang, Y. Zhang, V.S. Raman, S.G. Reed, X. Ma, Regulatory actions of Toll-like receptor 2 (TLR2) and TLR4 in Leishmania donovani infection in the liver, Infect. Immun. 81 (7) (2013) 2318–2326, https://doi.org/10.1128/IAI.01468-12.
[43] M.A. Soltan, N. Elbassiouny, H. Gamal, E.B. Elkaeed, R.A. Eid, M.A. Eldeen, A.A. Al-Karmalawy, In silico prediction of a multitope vaccine against Moraxella catarrhalis: reverse vaccinology and immunoinformatics, Vaccines 9 (6) (2021) 669, https://doi.org/10.3390/vaccines9060669.
[44] M. Dalsass, A. Brozzi, D. Medini, R. Rappuoli, Comparison of open-source reverse vaccinology programs for bacterial vaccine antigen discovery, Front. Immunol. 10 (2019) 113, https://doi.org/10.3389/fimmu.2019.00113.
[45] D.M. Yang, M.V. Rogers, F.Y. Liew, Identification and characterization of host-protective T-cell epitopes of a major surface glycoprotein (gp63) from Leishmania major, Immunology 72 (1) (1991) 3.
[46] D.M. Resende, A.M. Rezende, N.J. Oliveira, I.C. Batista, R. Corrêa-Oliveira, A.B. Reis, J.C. Ruiz, An assessment on epitope prediction methods for protozoa genomes, BMC Bioinf. 13 (1) (2012) 1–11, https://doi.org/10.1186/1471-2105-13-309.
[47] R.C. Brito, F.G. Guimarães, J.P. Velloso, R. Corrêa-Oliveira, J.C. Ruiz, A.B. Reis, D.M. Resende, Immunoinformatics features linked to Leishmania vaccine development: data integration of experimental and in silico studies, Int. J. Mol. Sci. 18 (2) (2017) 371, https://doi.org/10.3390/ijms18020371.
[48] G. Garg, K. Singh, V. Ali, Proteomic approaches unravel the intricacy of secreted proteins of Leishmania: an updated review, Biochim. Biophys. Acta Protein Proteonomics 1866 (8) (2018) 913–923, https://doi.org/10.1016/j.bbapap.2018.05.011.
[49] A. Atapour, F. Ghalamfarsa, S. Naderi, G. Hatam, Designing of a novel fusion protein vaccine candidate against human visceral leishmaniasis (vl) using immunoinformatics and structural approaches, Int. J. Pept. Res. Therapeut. 27 (3) (2021) 1885–1898, https://doi.org/10.1007/s10989-021-10218-8.
[50] C. Aguttu, B.A. Okech, A. Mukisa, G.W. Lubega, A potential malaria vaccine candidate identified using an insilico approach, American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS) 48 (1) (2018) 1–13.
[51] S. Ahmad, S.S. Azam, A novel approach of virulome based reverse vaccinology for exploring and validating peptide-based vaccine candidates against the most troublesome nosocomial pathogen: acinetobacter baumannii, J. Mol. Graph. Model. 83 (2018) 1–11, https://doi.org/10.1016/j.jmgm.2018.04.020.
[52] A.M. Arega, K.P. Pattanaik, S. Nayak, R.K. Mahapatra, Computational discovery and ex-vivo validation study of novel antigenic vaccine candidates against tuberculosis, Acta Trop. 217 (2021), 105870, https://doi.org/10.1016/j.actatropica.2021.105870.
[53] A.N. Sarangi, R. Aggarwal, Q. Rahman, N. Trivedi, Subtractive genomics approach for in silico identification and characterization of novel drug targets in Neisseria Meningitides Serogroup B, J. Comput. Sci. Syst. Biol. 2 (5) (2009) 255–258, https://doi.org/10.4172/jcsb.1000038.
[54] L. John, G.J. John, T. Kholia, A reverse vaccinology approach for the identification of potential vaccine candidates from Leishmania spp, Appl. Biochem. Biotechnol. 167 (2012) 1340–1350, https://doi.org/10.1007/s12010-012-9649-0(2012).
[55] S.K. Panda, R.K. Mahapatra, In-silico screening, identification and validation of a novel vaccine candidate in the fight against Plasmodium falciparum, Parasitol. Res. 116 (2017) 1293–1305, https://doi.org/10.1007/s00436-017-5408-z.
[56] V. Solanki, M. Tiwari, V. Tiwari, Prioritization of potential vaccine targets using comparative proteomics and designing of the chimeric multi-epitope vaccine against Pseudomonas aeruginosa, Sci. Rep. 9 (1) (2019) 1–19, https://doi.org/10.1038/s41598-019-41496-4.
[57] M. Mushtaq, S. Khan, M. Hassan, A.I. Al-Harbi, A.R. Hameed, K. Khan, S. Ahmad, Computational design of a chimeric vaccine against plesiomonas shigelloides using pan-genome and reverse vaccinology, Vaccines 10 (11) (2022) 1886, https://doi.org/10.3390/vaccines10111886.
[58] M.H. Van Regenmortel, What is a B-cell epitope?, in: *Epitope Mapping Protocols* Human Press, 2009, pp. 3–20.
[59] J.S. Blum, P.A. Wearsch, P. Cresswell, Pathways of antigen processing, Annu. Rev. Immunol. 31 (2013) 443–473, https://doi.org/10.1146/annurev-immunol-032712-095910.
[60] B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data, Nucleic Acids Res. 48 (W1) (2020) W449–W454, https://doi.org/10.1093/nar/gkaa379.
[61] B. Reynisson, C. Barra, S. Kaabinejadian, W.H. Hildebrand, B. Peters, M. Nielsen, Improved prediction of MHC II antigen presentation through integration and motif deconvolution of mass spectrometry MHC eluted ligand data, J. Proteome Res. 19 (6) (2020) 2304–2315, https://doi.org/10.1021/acs.jproteome.9b00874.
[62] S. Delfani, A.A.I. Fooladi, A.M. Mobarez, M. Emaneini, J. Amani, H. Sedighian, In silico analysis for identifying potential vaccine candidates against Staphylococcus aureus, Clinical and experimental vaccine research 4 (1) (2015) 99–106, https://doi.org/10.7774/cevr.2015.4.1.99.
[63] O.S. Chukwudozie, C.M. Gray, T.A. Fagbayi, R.C. Chukwuanukwu, V.O. Oyebanji, T.T. Bankole, E.M. Daniel, Immuno-informatics design of a multimeric epitope peptide based vaccine targeting SARS-CoV-2 spike glycoprotein, PLoS One 16 (3) (2021), e0248061, https://doi.org/10.1371/journal.pone.0248061.
[64] V. Kumar, S. Kancharla, P. Kolli, M. Jena, Reverse vaccinology approach towards the in-silico multiepitope vaccine development against SARS-CoV-2, F1000Research 10 (2021), https://doi.org/10.12688/f1000research.36371.1.
[65] G. Singh, M. Pritam, M. Banerjee, A.K. Singh, S.P. Singh, Designing of precise vaccine construct against visceral leishmaniasis through predicted epitope ensemble: a contemporary approach, Comput. Biol. Chem. 86 (2020), 107259, https://doi.org/10.1016/j.compbiolchem.2020.107259.
[66] Y. He, Z. Xiang, Databases and in silico tools for vaccine design, in: In Silico Models for Drug Discovery, 2013, pp. 115–127, https://doi.org/10.1007/978-1-62703-342-8_8.
[67] R. Dong, Z. Chu, F. Yu, Y. Zha, Contriving multi-epitope subunit of vaccine for COVID-19: immunoinformatics approaches, Front. Immunol. 11 (2020) 1784, https://doi.org/10.3389/fimmu.2020.01784.
[68] T. Kar, U. Narsaria, S. Basak, D. Deb, F. Castiglione, D.M. Mueller, A.P. Srivastava, A candidate multi-epitope vaccine against SARS-CoV-2, Sci. Rep. 10 (1) (2020), 10895, https://doi.org/10.1038/s41598-020-67749-1.
[69] P. Bhatt, M. Sharma, S. Sharma, Prediction and identification of T cell epitopes of COVID-19 with balanced cytokine response for the development of peptide based vaccines, Silico Pharmacology 9 (2021) 1–17, https://doi.org/10.1007/s40203-021-00098-7.
[70] N. Khatoon, R.K. Pandey, R. Ojha, V.S. Aathmanathan, M. Krishnan, V.K. Prajapati, Exploratory algorithm to devise multi-epitope subunit vaccine by investigating Leishmania donovani membrane proteins, J. Biomol. Struct. Dyn. 37 (9) (2019) 2381–2393, https://doi.org/10.1080/07391102.2018.1484815.
[71] I.J. Amanna, M.K. Slifka, Contributions of humoral and cellular immunity to vaccine-induced protection in humans, Virology 411 (2) (2011) 206–215, https://doi.org/10.1016/j.virol.2010.12.016.
[72] L.E. Guimarães, B. Baker, C. Perricone, Y. Shoenfeld, Vaccines, adjuvants and autoimmunity, Pharmacol. Res. 100 (2015) 190–209, https://doi.org/10.1016/j.phrs.2015.08.003.
[73] C.L. Montes, E.V. Acosta-Rodríguez, M.C. Merino, D.A. Bermejo, A. Gruppi, Polyclonal B cell activation in infections: infectious agents' devilry or defense mechanism of the host? J. Leukoc. Biol. 82 (5) (2007) 1027–1032, https://doi.org/10.1189/jlb.0407214.