



OPEN

Identification and prioritisation of potential vaccine candidates using subtractive proteomics and designing of a multi-epitope vaccine against *Wuchereria bancrofti*

Murali Aarthy¹, G. Navaneetha Pandiyan¹, R. Paramasivan¹, Ashwani Kumar^{2,3} & Bhavna Gupta¹✉

This study employed subtractive proteomics and immunoinformatics to analyze the *Wuchereria bancrofti* proteome and identify potential therapeutic targets, with a focus on designing a vaccine against the parasite species. A comprehensive bioinformatics analysis of the parasite's proteome identified 51 probable therapeutic targets, among which "Kunitz/bovine pancreatic trypsin inhibitor domain-containing protein" was identified as the most promising vaccine candidate. The candidate protein was used to design a multi-epitope vaccine, incorporating B-cell and T-cell epitopes identified through various tools. The vaccine construct underwent extensive analysis of its antigenic, physical, and chemical features, including the determination of secondary and tertiary structures. Docking and molecular dynamics simulations were performed with HLA alleles, Toll-like receptor 4 (TLR4), and TLR3 to assess its potential to elicit the human immune response. Immune simulation analysis confirmed the predicted vaccine's strong binding affinity with immunoglobulins, indicating its potential efficacy in generating an immune response. However, experimental validation and testing of this multi-epitope vaccine construct would be needed to assess its potential against *W. bancrofti* and even for a broader range of lymphatic filarial infections given the similarities between *W. bancrofti* and *Brugia*.

Lymphatic filariasis is a neglected tropical disease that poses a significant public health burden worldwide. It is primarily transmitted through the bites of infected mosquitoes carrying parasites of the *Wuchereria bancrofti*, *Brugia malayi*, and *B. timori* species. Among these, *W. bancrofti* accounts for more than 90% of lymphatic filariasis cases globally^{1,2}. The impact of lymphatic filariasis on affected individuals and communities is severe. The disease causes chronic and debilitating conditions such as lymphedema, elephantiasis, and hydrocele³. These conditions not only lead to physical disability but also result in social stigma, economic hardship, and reduced productivity. The burden of the disease falls disproportionately on marginalized populations in low-resource settings, further exacerbating health inequities.

To address this global health challenge, the World Health Organization (WHO) launched the Global Program to Eliminate Lymphatic Filariasis (GPELF) in 2000. The program aimed to eliminate lymphatic filariasis as a public health problem by 2020⁴ and this timeline for elimination has been extended to 2030 now⁵. GPELF adopted a comprehensive approach that combined mass drug administration (MDA) with other preventive measures such as vector control, morbidity management, and disability prevention. MDA involves the administration of anti-filarial drugs, including Albendazole, Ivermectin, and Diethylcarbamazine (DEC)⁶, to entire at-risk populations in endemic areas. These drugs target the microfilariae, the larval stage of the parasites circulating in the

¹ICMR-Vector Control Research Centre (VCRC), Field Station, Madurai, Tamil Nadu 625002, India. ²ICMR-Vector Control Research Centre (VCRC), Puducherry, India. ³Present address: Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Tandalam, Chennai, Tamil Nadu 602105, India. ✉email: bhavnaguptaster@gmail.com

bloodstream, and aim to reduce their numbers, prevent transmission, and alleviate symptoms. The success of MDA campaigns in reducing disease burden has been significant, and several countries have made remarkable progress toward elimination goals^{7,8}. However, challenges remain on the path toward complete elimination of lymphatic filariasis. One of the primary challenges is the persistence of adult worms in infected individuals, as the currently available drugs primarily target the microfilariae^{9,10}. These adult worms can continue to produce microfilariae, perpetuating the transmission cycle of the disease. Additionally, there are concerns about the emergence of drug resistance in the parasite populations^{11–16}. There have been some reports of poor responses to DEC^{12,13,16} treatment as well as to Ivermectin^{11,17}. Moreover, a mutation associated with resistance to Albendazole has been detected in *W. bancrofti* populations¹⁵. Such challenges could impact the effectiveness of MDA interventions. Innovative strategies are required to overcome these challenges and achieve sustained elimination. One promising approach is the development of vaccines against lymphatic filariasis. Vaccines have the potential to provide long-lasting protection, target multiple stages of the parasite's life cycle, and complement existing control measures^{18–20}. Extensive research efforts have been dedicated to identifying suitable vaccine candidates using various approaches, including subtractive genomics and reverse vaccinology^{21–24}.

Subtractive genomics involves comparing the genomes of the parasite with the host organisms to identify unique proteins or metabolic pathways that could be potential targets for interventions. This approach helps in prioritizing proteins essential for the survival or reproduction of the parasite²⁵, that could be explored to develop potential vaccines or new drug therapies. Reverse vaccinology complements subtractive genomics to predict immunogenic epitopes, design and evaluate multi-epitope vaccines^{26,27} using various immune-informatics tools. This computational approach aids in the selection of vaccine candidates with the highest potential for success, minimizing the need for extensive experimental testing²⁸.

Furthermore, significant advancements in high-throughput sequencing technologies and proteomics have facilitated the identification and characterization of potential vaccine targets. By analyzing the proteome of the parasite, the proteins involved in host-parasite interactions, immune evasion mechanisms, and pathogenesis can be identified²⁹. This knowledge guides the development of effective vaccines that can stimulate robust and specific immune responses against parasites^{30,31}. Several potential vaccine targets have been identified and characterized in *B. malayi* using these innovative approaches. Some of the promising vaccine candidates for combating lymphatic filariasis include abundant larval transcript 2 (ALT-2)³², thioredoxin peroxidase (TPX)³², vespid venom allergen (VAH)³³, heat shock protein (HSP)³⁴, glutathione S-transferases (GST)³⁵. However, these antigens have not yielded satisfactory results individually. Combining these candidates into multivalent vaccines has shown potential in addressing lymphatic filariasis. For instance, recombinant multivalent fusion protein vaccine (rBm-HAXT) the combination of ALT-2, HSP, thioredoxin peroxidase 2, and tetraspanin large extracellular loop is among the most promising candidates. It has demonstrated 88–94% protection against challenge infections in rodents^{36,37} and approximately 57% protection in rhesus macaques³⁸. Another multivalent vaccine candidate is the combination of *B. malayi* VAH and ALT-2³⁹. Further expanding and diversifying the database of potential vaccine targets is essential for advancing vaccine development efforts in the fight against lymphatic filariasis.

In this study, subtractive proteomics and reverse vaccinology techniques were employed to identify potential vaccine candidates for lymphatic filariasis by screening the putative proteome of *W. bancrofti*. Through prioritization of the proteins with anti-parasitic properties, the study provides valuable insights into promising targets that could be explored further for their vaccine development potential. The protein named “Kunitz type inhibitor domain-containing protein” (VDM15541), emerged as the most promising vaccine candidate among the 51 therapeutic targets identified. This protein was further analyzed to design a multi-epitope vaccine construct. Extensive analysis of the antigenic, physical, and chemical features of the vaccine construct, including its secondary and tertiary structure, confirmed its potential immunogenicity. Multi-epitope vaccines contain multiple epitopes or antigenic regions from a pathogen's proteins. Unlike traditional single-antigen vaccines, multi-epitope vaccines provide several advantages. Multiple epitopes enhance the immune response and reduces the chances of immune escape as well as development of resistance in the pathogens^{40,41}. Multi-epitope vaccines hold great promise for the development of effective and versatile vaccine solutions.

Materials and methods

Data retrieval

The entire set of 13,058 proteins linked to *W. bancrofti* Genome Project PRJEB536 with the Taxonomy ID: 6293 (BioSample: SAMEA1686771) was obtained from the NCBI database in January 2022. The human proteome dataset with the Genome Assembly ID: GrCh38.p14 was retrieved from NCBI.

Subtractive proteomic analysis to identify and prioritize the potential vaccine candidates

Subtractive proteomics helps to identify potential proteins with specific features from a large dataset using several subtractive steps as shown in Fig. 1. *W. bancrofti* proteome was BLAST analyzed against the human proteome to identify proteins with similarity. Proteins showing any similarity to human proteins were excluded from further analysis. CD-HIT analysis (<https://www.bioinformatics.org/cd-hit/>) was performed using a 75% similarity threshold to eliminate paralogous proteins and reduce redundancy. The remaining non-homologous and non-paralogous proteins were selected for further analysis. Using BLAST analysis, the selected proteins were compared to the Database of Essential Genes (DEG)⁴². Essential proteins were identified based on stringent criteria such as an e-value of 10^{-10} , bit score 100, and percent identity of > 30. Although this strict approach may have excluded some potentially valuable information, the selected genes ensure high-quality candidates for further analysis. The subcellular localization of the identified proteins was predicted using the WoLF PSORT⁴³ bioinformatics tool. Proteins predicted to be membrane bound or extracellular in nature were considered potential vaccine targets. The presence of a signal peptide, indicative of potential protein secretion, was analyzed using the SignalP 5.0

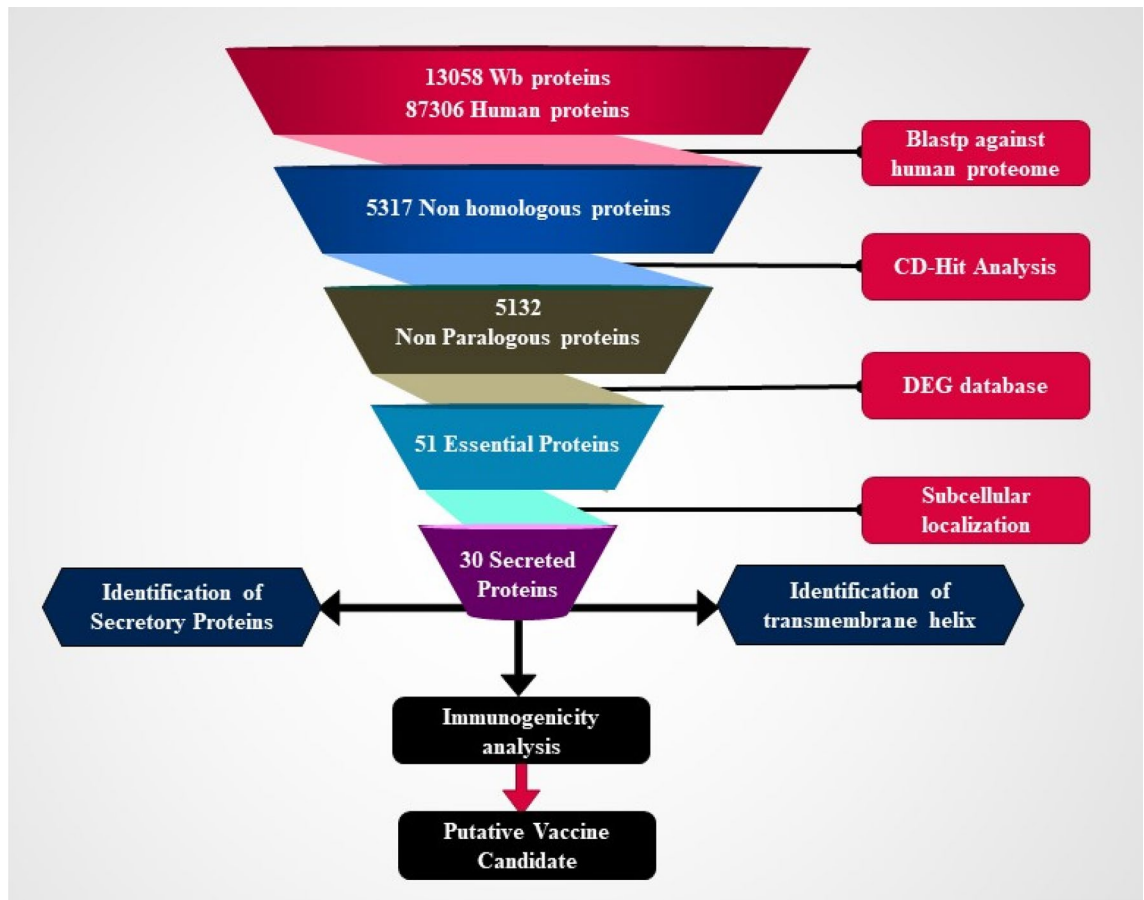


Figure 1. Workflow of subtractive proteomics to identify and potential vaccine candidates against *Wuchereria bancrofti*.

tool⁴⁴. The SecretomeP tool⁴⁵ was used to predict proteins potentially exported through non-classical pathways. Proteins predicted to be secretory through both classical and non-classical pathways were considered secreted proteins. Furthermore, the transmembrane domains in each protein were identified using the TMHMM server (<https://services.healthtech.dtu.dk/services/TMHMM-2.0/>)⁴⁶. Proteins with more than one transmembrane helix were considered membrane-bound and excluded from further analysis. The VaxiJen v2.0⁴⁷ and AntigenPro⁴⁸ tools were utilized to assess the antigenicity and immunogenicity of the selected proteins, respectively. Proteins with high antigenicity and immunogenicity scores were prioritized as potential vaccine targets. Various bioinformatics analyses, such as sequence alignment, and domain analysis, were performed to further characterize the selected vaccine target.

Reverse vaccinology to develop a putative vaccine against *W. bancrofti*

Further analysis of the potential vaccine candidate was performed using immunoinformatics analysis (Fig. 2). MHC-I binding epitopes were predicted using the NetCTL server⁴⁹ with a threshold for epitope identification of 0.75, weight on TAP transport efficiency 0.05 and weight on C terminal cleavage of 0.15. Immune Epitope Database Analysis Resource (IEDB AR)⁵⁰ was used to predict the binding of shortlisted epitopes with MHC-I. MHC-II binding epitopes were predicted through IEDB 2023.05. B-cell epitopes were predicted using ABCPred⁵¹, BCePred prediction server⁵², and Bepipred 2.0⁵³. Potent B-cell and MHC-I epitopes were combined to form a multi-epitope vaccine construct. Adjuvant β defensin was added at the N-terminal end and linked to MHC-I epitopes using the EAAAK linker. MHC-I epitopes were joined using the EAAAK linker, and B-cell epitopes were joined using the GPGPG linker.

The physical and chemical properties of the vaccine construct, including molecular weight, aliphatic index, instability index, GRAVY score, and solubility, were determined using ProtParam tool (<https://web.expasy.org/protparam/protpar-ref.html>).

Evaluation of immune response of the vaccine construct

Antigenicity and immunogenicity of the vaccine construct were assessed using AntigenPro and VaxiJen v2.0, respectively. The allergic response of the vaccine construct was determined using AlgPred server⁵⁴ and AllerCatPro server 2.0⁵⁵.

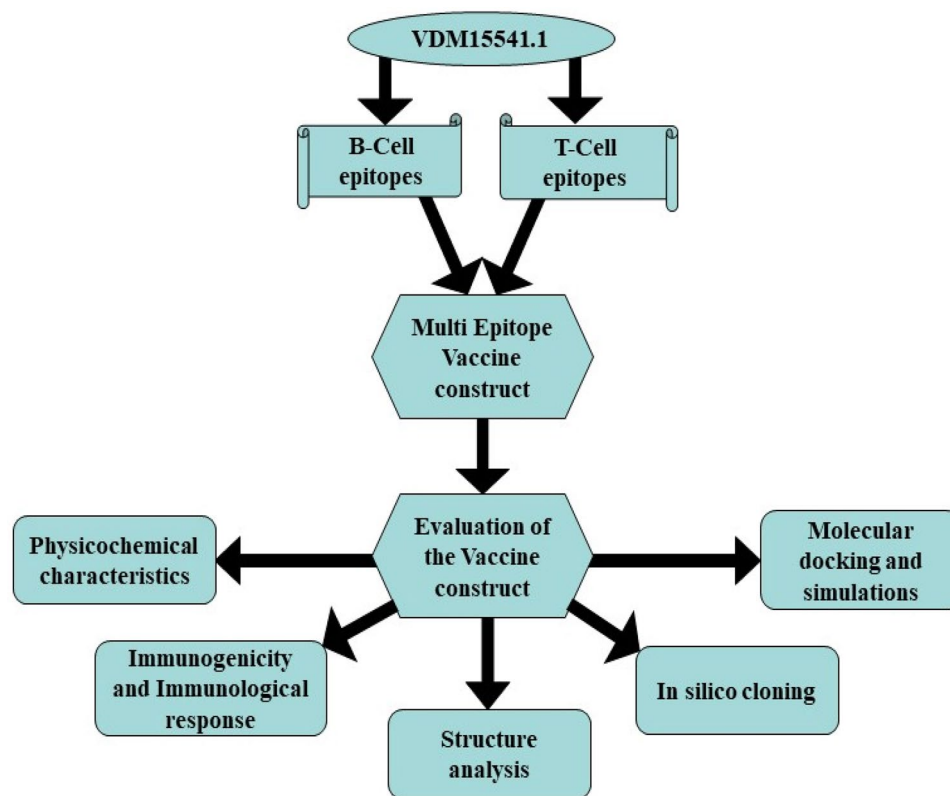


Figure 2. Workflow of reverse vaccinology analysis to design and evaluate multi-epitope vaccine against *Wuchereria bancrofti*.

Prediction of post-translational modifications (PTMs)

The prediction of post-translational modifications including glycosylation, phosphorylation, and acetylation of the vaccine construct was carried out using NetNGlyc 1.0, NetPhos 3.0, and NetAct 1.0 servers, which are available at <http://www.cbs.dtu.dk/services>. Additionally, N-terminal glycines myristoyl and GPI-modifications were assessed using MyrPS/NMT (<https://mendel.imp.ac.at/myristate/>)⁵⁶ and big-PI/GPI servers (https://mendel.imp.ac.at/gpi/gpi_server.html)⁵⁷, respectively.

Vaccine model construction

To design and construct the model of the vaccine, the sequence VDM15541 was retrieved from the UniProtKB database⁵⁸. The sequence was checked for suitable template through BLAST using default parameters. The protein with the query coverage of 24% and percent identity of 95.56% was obtained. The length of the residues for template protein was only 45 when the length of query protein was 165 amino acids in length. Since, the query coverage was very low, we tried ab initio threading of the protein through i-TASSER-MTD⁵⁹ and obtained five models. The five models were analyzed and validated through PDBSum generated for the Ramachandran plot analysis⁶⁰. The best model was further used for the interaction studies with the human TLR3 and TLR4 protein^{61,62} and HLA alleles.

Protein–protein docking

The best model of the vaccine construct was used for the docking with TLR3 encompassing the PDB ID: 1ZIW and TLR4 with the PDB ID: 3FXI in HDock (<http://hdock.phys.hust.edu.cn/>) along with the HLA alleles. The HDock server is a highly integrated suite of molecular modeling techniques. This includes homology search, template-based modeling, structure prediction, docking etc., This docking platform is different from other docking platforms which uses the amino acids as input and the strategy of docking in which the information about the binding site is retrieved from the experimental studies. Moreover, the small angle X-ray scattering is incorporated during the docking and post docking processes. HDock supports the docking function through the intrinsic scoring function^{63,64}. Further, the binding affinity of the protein and alleles was calculated through the PPI binding tool (<https://protcal-suite.cbe.bci.tu-dortmund.de/PPIAffinity/>). This tool uses binding affinity predictors from support vector machines to screen datasets of protein–protein and protein–peptide complexes⁶⁵.

Molecular dynamics simulation

Molecular dynamics simulation was carried out with the help of MD Simulation tool incorporated in SiBI-OLEAD which uses GROMACS simulation application (<https://sibiolead.com/MDSIM>). This web server uses

five steps which includes preprocessing, energy minimization, equilibration, production dynamics and trajectory analysis. The protein complex of vaccine and TLR and HLA alleles were used for the simulation process. The protein complex was placed in a Cubic box filled with an SPC water model. Further, the protein complex was neutralized with the NaCl ions with the concentration of 0.15 molar. The steepest descent integrator with the 5000 steps was applied for the energy minimization parameters. The default equilibration type NVT/NPT with the temperature of 300 K, pressure of 1 bar with 100 ps was performed before the production dynamics in order to equilibrate the system. Finally, the Leapfrog integrator with the time period of 10 ns was carried out for the production dynamics with the 5000 frames⁶⁶.

In-silico cloning and codon optimisation

The JCat web server⁶⁷ was used for back-translation and codon optimization of the vaccine construct. Codons were modified based on the codon usage preferences of the target host, *E. coli* K12, to enhance expression rates. The modified sequence was analyzed for GC content and codon adaptation index using the JCat webserver.

Simulations for evaluating the immune response of the vaccine construct in host cells

The C-ImmSim server⁶⁸ was used to simulate the immune response to the predicted vaccine construct. Three doses of the vaccine construct were administered at three-month intervals. The simulation tracked the levels of IgM, IgG1, and IgG2 antibodies over time to assess the immune response. The frequency and activation of CD8+ and CD4+ T-cells, as well as the formation of memory cells, were also monitored. Cytokine levels during the immune reaction were analyzed using the Simpson index D graph to evaluate potential risks and safety of the vaccine construct.

Results and discussion

In this study, we aimed to identify potential vaccine targets against *W. bancrofti* by screening its entire proteome. Total 13,058 proteins were retrieved from the PRJEB536 genome sequence project. To ensure these targets would not induce adverse immunogenic effects by cross-reacting with human proteins, BLAST analysis was conducted against the human proteome, to exclude any similar proteins. Subsequently, the non-homologous proteins were analyzed using CD-HIT analysis (75% similarity threshold) to remove redundancy, resulting in 5132 proteins. We further subjected these proteins to BLAST analysis against the DEG database to identify essential proteins crucial for pathogen survival. A subset of 51 essential proteins (Table 1) was identified, which were selected based on stringent criteria. These 51 proteins represent promising candidates for further analysis in vaccine development against *W. bancrofti*, offering potential therapeutic targets. Further analysis of these 51 proteins was conducted to assess their immunogenicity, efficacy, and safety, providing valuable insights into their potential as vaccine candidates against *W. bancrofti*.

Prioritization of putative vaccine targets

To prioritize the putative vaccine targets several factors were taken into consideration. Firstly, the subcellular localization of the 51 identified proteins was determined using the WoLF PSORT bioinformatics tool. It was observed that 24 proteins were associated with the plasma membrane, six were extracellular, five were in the organelles, and 13 were in the cytosol (Table 1). Theoretically, the membrane-associated proteins are more exposed to the human immune system and are thus considered potential vaccine targets. Cytoplasmic membrane proteins are also important for the physiology of bacteria, as they are involved in many important metabolic functions and thus are potential drug targets. Out of the 51 proteins, 30 proteins (associated with plasma-membrane and extracellular) were found as the potential vaccine targets based on this analysis. However, as these predictions are based on bioinformatics algorithms, it is crucial to corroborate the findings with multiple analyses. For example, the secretory potential of the proteins was assessed using several methods. First, SignalP tool was used to identify the signal peptide, which facilitates transport to the extracellular region. Among the 51 proteins, only three were found to have a signal peptide (Table 1), suggesting their potential for secretion. Secondly, SecretomeP tool was employed using a mammalian secreted proteins database to identify the proteins which are likely to be exported through non-classical pathways that do not require a signal peptide⁶⁹. This method detected four proteins (Table 1). Thirdly, the presence of transmembrane hydrophobic α -helices was determined using TMHMM server. The secretory proteins can have diverse locations, including the periplasm, outer membrane, or periplasmic side of the inner membrane. Membrane-bound proteins typically contain transmembrane hydrophobic α -helices which helps in their embedding in the membrane. However, membrane-bound proteins are difficult to purify^{70,71}. The proteins with more than one transmembrane helix were considered membrane-bound and were excluded from further analysis. Among the analyzed 51 proteins, two proteins, namely VDM10251 and VDM15541, met the criteria of being secretory either having signal peptide or through the non-secretory pathway and having less than one transmembrane helix (Table 1). However, it is important to note that VDM10251 is a collagen protein typically found in the extracellular matrix, and it does not exhibit secretory properties. While protein VDM15541 is an orthologue of mlt-11 in *C. elegans* which is a secretory protein and is associated with molting⁷². VDM15541 also showed high antigenicity (0.825) and immunogenicity (0.943) as assessed by AntigenPro and Vaxijen tools, respectively.

Through this comprehensive analysis combining subcellular localization, secretion prediction, membrane localization, and antigenicity assessment, VDM15541, a 165 amino acid long protein was identified as the most promising vaccine candidate for further evaluation as a potential vaccine target against *W. bancrofti*.

Sl no.	Gene ID	Sub-cellular location	Signal IP	Secretome	TMHMM	Vaxijen Score
1	VDM06647.1	Plasma membrane			0	0.331
2	VDM06730.1	Plasma membrane			1	0.5839
3	VDM06817.1	Plasma membrane			6	0.5399
4	VDM06914.1	Plasma membrane			10	0.6464
5	VDM06935.1	Plasma membrane		Yes	12	0.4645
6	VDM07450.1	Plasma membrane			11	0.4436
7	VDM07534.1	Extracellular		Yes	3	0.6416
8	VDM09497.1	Plasma membrane			1	0.6711
9	VDM09927.1	Cytosol			0	0.5954
10	VDM09984.1	Plasma membrane			3	0.6092
11	VDM10251.1	Plasma membrane		Yes	1	0.3749
12	VDM10465.1	Cytosol			0	0.5102
13	VDM10950.1	Extracellular		Yes	4	0.4195
14	VDM10964.1	Cytosol			0	0.7157
15	VDM11313.1	Plasma membrane			7	0.574
16	VDM11729.1	Plasma membrane			4	0.6573
17	VDM12133.1	Plasma membrane			3	0.5382
18	VDM12167.1	Cytosol			0	0.2702
19	VDM12641.1	Cytosol			0	0.3612
20	VDM12677.1	Plasma membrane			2	0.5437
21	VDM12793.1	Plasma membrane			0	0.6694
22	VDM13458.1	Cytosol			0	0.7542
23	VDM13776.1	Plasma membrane			2	0.5851
24	VDM14680.1	Cytoskeleton			5	0.5977
25	VDM14801.1	Plasma membrane	Yes	Yes	3	0.6152
26	VDM15121.1	Extracellular			0	0.5909
27	VDM15129.1	Cytosol			0	0.4969
28	VDM15521.1	Nuclear			0	0.6647
29	VDM15541.1	Plasma membrane	Yes	Yes	1	0.8248
30	VDM15698.1	Cytosol			0	0.5138
31	VDM16641.1	Cytoskeleton			0	0.5628
32	VDM16965.1	Extracellular			0	0.1302
33	VDM17544.1	Extracellular			0	0.7654
34	VDM17973.1	Cytosol			1	0.6007
35	VDM19390.1	Cytosol			0	0.3326
36	VDM19792.1	Plasma membrane			2	0.6621
37	VDM20216.1	Cytoskeleton			3	0.5835
38	VDM20372.1	Plasma membrane			4	0.6906
39	VDM20417.1	Plasma membrane		Yes	4	0.4638
40	VDM20455.1	Mitochondria			0	0.401
41	VDM20976.1	Plasma membrane	Yes	Yes	12	0.6795
42	VDM21381.1	Cytosol			0	0.487
43	VDM21514.1	Extracellular			0	0.4478
44	VDM21712.1	Plasma membrane			10	0.6351
45	VDM21861.1	Cytosol			0	0.4379
46	VDM21901.1	Plasma membrane			3	0.557
47	VDM21936.1	Plasma membrane			9	0.5519
48	VDM22966.1	Endoplasmic reticulum			0	0.4424
49	VDM23249.1	Cytosol			0	0.515
50	VDM23299.1	Mitochondria			0	0.4192
51	VDM23483.1	Mitochondria			0	0.3879

Table 1. List of 51 probably therapeutic targets identified from the analysis of *Wuchereria bancrofti* proteome.

Importance of the selected vaccine target

Due to incomplete annotation of the *W. bancrofti*, further analysis of the identified protein (VDM15541) was conducted by comparing it to known proteins from *Brugia* and other nematodes. While there has not been an in-depth analysis of inter- and intra-species variations in the genome of lymphatic filariasis (LF) parasites, data comparing selected genes between *W. bancrofti* and *B. malayi* have demonstrated significant homology, with over 95% similarity⁷³. The first 100 hits for VDM15541 from BLAST analysis are given in Table S1. The protein showed similarity with several nematodes including *Brugia*, *Caenorhabditis*, *Dirofilaria*, *Onchocerca*, *Loa loa* with percent identity ranging from 40 to 89% and query coverage from 26 to 95%. VDM15541 showed 80.13% identity with 92% query coverage with *B. malayi* Kunitz/Bovine pancreatic trypsin inhibitor domain-containing protein (BMA-MLT-11). BMA-MLT-11 is a secretory protein and has three known isoforms (Bm6109a, Bm6109b and Bm6109c). These proteins have not been characterized extensively. However, its orthologue *mlt-11* in *C. elegans* is known to be associated with molting⁷². Knockdown of *mlt-11* by RNAi induced Larval (L1) arrest and defects during larval molting^{74–76}. *W. bancrofti* protein covered only a part of the BMA-MLT-11 (isoform Bm6109a) which is ~3000 amino-acid long. However, 100% similar protein was identified in all the four BioSamples sequenced under three different BioProjects (PRJNA526, PRJNA37759 and two samples under PRJNA275548) (Figure S1). The InterPro analysis of VDM15541 provided valuable insights into the protein's domain and other features (Figure S2). Notably, a pancreatic trypsin inhibitor Kunitz domain from 21 to 105 amino-acid was identified. Kunitz inhibitors have been recognized as promising vaccine targets, as they play a crucial role in modulating host immunity^{77–79}. It is worth noting that Kunitz inhibitors have already shown promise as vaccine candidates in other helminths, such as *Fasciola hepatica* and *Schistosoma mansoni*^{78,80–83}. Notably, Kunitz protease inhibitor gene *Smp_147730* in *S. mansoni* is 146 amino-acid long and contains only one Kunitz/bovine protease inhibitor domain⁸¹ similar to VDM15541.

Kunitz-type inhibitors play a crucial role in the context of host blood feeding and pathogen transmission, particularly in vector-borne diseases⁷⁷. For example, these inhibitors are often produced by blood-feeding arthropods, such as mosquitoes, ticks, and sandflies, as well as by the pathogens they transmit. When a blood-feeding arthropod bites a host, it injects saliva containing various bioactive molecules, including Kunitz-type inhibitors. These inhibitors are essential for the arthropod's successful blood meal acquisition by modulating host hemostasis and immune responses⁸⁴. By inhibiting host proteases involved in blood clotting and inflammation, Kunitz-type inhibitors prevent the formation of blood clots and suppress host immune reactions at the bite site. This allows the arthropod to feed on blood more efficiently and remain undetected by the host's immune system.

In the context of pathogen transmission, Kunitz-type inhibitors have additional implications. They can interfere with the host's immune response to the pathogen, aiding in its establishment and survival within the host⁸². Interestingly, several helminths themselves produce Kunitz-type inhibitors as a means to manipulate the host's immune system. These pathogen-derived inhibitors can modulate host immune responses, hinder immune cell activation, and suppress immune defenses, ultimately promoting pathogen survival and dissemination within the host⁸².

Given the critical role of Kunitz-type inhibitors in host blood feeding and pathogen transmission, they have emerged as potential targets for intervention strategies. By targeting these inhibitors, it may be possible to disrupt the delicate balance between the vector, pathogen, and host interactions. However, it is important to note that further research is needed to fully understand the complexity of these interactions and the potential of targeting Kunitz-type inhibitors as a control strategy.

Evaluation of selected protein as a vaccine target

Identification of T-cell and B-cell epitopes

The identification of T-cell and B-cell epitopes is crucial for designing an effective vaccine. T-cells are classified as CD4+ (HTLs) or CD8+ (CTLs) based on the receptors present on their membranes. These cells interact with epitopes presented by MHC-I and MHC-II molecules. CTLs bind to MHC-I binding epitopes, while HTLs interact with MHC-II molecules. Strong interactions between antigenic epitopes and MHC molecules lead to a robust immune response. Therefore, predicting high-affinity epitopes is important.

NetCTL identified 26 MHC ligands with a threshold score of 0.75 for epitope identification that gives 65% sensitivity and 97% specificity. Out of these 26 predicted epitopes, six showed binding to MHC_I alleles (Table S2). However, none of the MHC-II binding epitopes with a percentile rank below 0.2 were identified. Thus only MHC-I epitopes were further assessed for their antigenicity, conservancy, and toxicity. All the epitopes were found 100% conserved and non-toxic to the host cells. Out of six epitopes, four had antigenicity > 0.5 and thus those epitopes were selected for reverse vaccinology analysis (Table S2).

B-cell epitopes were predicted using servers ABCPred, BCPred, and BepiPred. ABCpred, employs an Artificial Neural Network (ANN) based method that considers amino acid sequences and their physicochemical properties to predict B cell epitopes. It calculates a propensity score for each amino acid in the sequence and subsequently assigns a binary classification (epitope or non-epitope). Threshold value of 0.5 and length of linear B-cell epitopes of 16 were used for epitope prediction on ABCpred server. Similarly, predicted residue score of 0.2 was used for BepiPred-2.0 which identifies the B cell epitopes based on a random forest algorithm trained on epitopes annotated from antibody-antigen protein structures. The ABCPred server predicted 12 B-cell epitopes, BepiPred identified seven epitopes and BCPred predicted six epitopes (Table S3), with the default threshold values for hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptides chains. On comparing all the epitopes six of them were identified as overlapping in two or more methods and were used for further analysis (Table 2).

S. no.	MHC-I	B-cell epitopes
1	TTIVVAIEK	NDVNVCKRQPFGRGRCPSVGGKGA
2	ANDENEPVL	CVSY
3	RSQFVLRYY	FGHCANDENEPVLYR
4	ALLAQVAFV	LNKWDTNDTTIQSTTIQNDNEMEKSENSS
5		ITDNDNDSDS
6		STTLSSSFPIIDNNNDNDT

Table 2. MHC-I and B-cell epitopes selected for designing multi-epitope vaccine.

Designing and characterization of a multi-epitope vaccine construct

The shortlisted potent B-cell and MHC-I epitopes were linked together to form a putative vaccine construct. An adjuvant, β -defensin, comprising 45 amino acids, was introduced at the N-terminal end of the vaccine construct. This adjuvant was connected to MHC-I alleles using the EAAAK linker. To enhance the overall effectiveness of the vaccine, MHC-I epitopes were linked together utilizing the EAAAK linker, while MHC-I alleles were connected to each other using the GPGPG linker. The final construct contained 182 amino acids with a molecular weight of 19 kDa.

An important aspect to consider when evaluating a vaccine candidate is its immunogenicity, which reflects its ability to stimulate an immune response, whether cell-mediated or humoral, against the targeted pathogen. According to computational findings from our study, the candidate vaccine construct demonstrated an immunogenicity score of 0.77 and antigenic probability score of 0.9425, as predicted by the VaxiJen and ANTIGENPro server, respectively. An average antigenic probability score of ≥ 0.8 is typically preferred for a vaccine candidate⁸⁵. Therefore, our obtained score suggests that the newly designed subunit vaccine candidate is likely to exhibit antigenic properties. Moreover, the solubility probability determined by SolPro⁸⁶ was 0.694. The schematic representation of the construct is shown in Fig. 3. The vaccine construct fulfilled all the criteria of a potential vaccine construct, warranting future evaluation.

Prediction of physicochemical property

The physicochemical property of the vaccine was analyzed through the ProtParam online server⁸⁷. This analysis provides various information about the protein like molecular weight, theoretical pI of the protein, hydrophatic index, aliphatic index etc., The molecular weight of the predicted vaccine is 19 kDa and the theoretical pI was 8.72 which states that the protein is stable. Further, the average of the hydrophatic index was -0.609 suggesting the hydrophilic nature of the protein that represents strong interactions with the water molecules. The aliphatic index of the vaccine was 62.75 indicating the stability of the protein in various temperatures. The instability score was 36.76 stating the stability which initiates an immunogenic reaction.

The vaccine construct was analyzed for various PTMs. The analysis conducted using the NetCGlyc-1.0 server indicated the absence of C-glycosylation sites in the vaccine construct. Phosphorylation modification prediction by the NetPhos-3.0 server identified 11 phosphorylation sites (Serine: 5, Threonine: 4, Tyrosine: 2) within the construct. Phosphorylated peptides or epitopes show high affinity towards cytotoxic T lymphocytes and thus play an important role fostering specific and robust immune responses⁸⁸. No N-acetylation site was detected by the NetAcet-1.0 server. Furthermore, there were no lipid PTMs at the N-terminal glycine myristoyl, and GPI-modification, as predicted by the MyrPS/NMT and the big-PI/GPI animals servers, respectively.

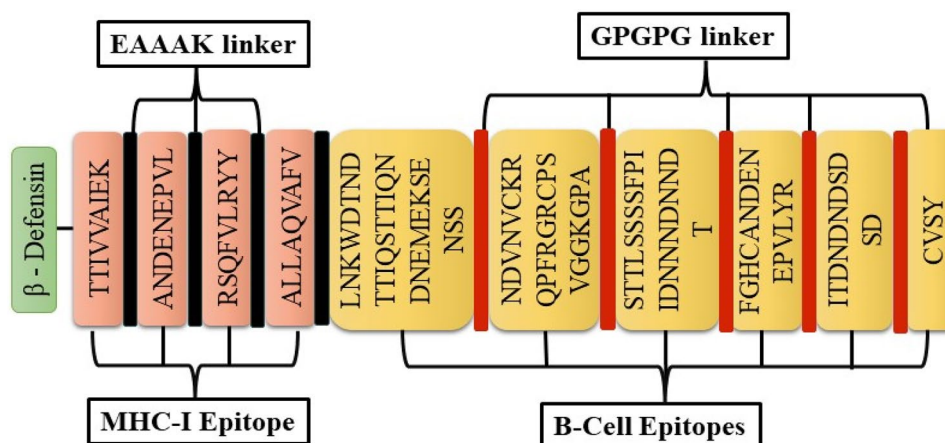


Figure 3. Schematic representation of the multi-epitope vaccine construct.

energy and can be taken further for molecular modeling studies. Further, the validation of the multi epitope fusion vaccines modelled has been presented in Supplementary Fig. S5 stating that the model has acquired stability and can be used further for the molecular modelling studies.

Interaction of the constructed vaccine with TLRs and HLA alleles

The modelled vaccine was further analyzed for the interaction with the TLR4, TLR3 human proteins and HLA alleles. Nine HLA alleles that showed binding with the MHC-I epitopes in this study were selected for docking. The interaction studies were performed using HDock. The HDock docking protocol showed a high negative score implying the strong binding energy due to good interaction with human proteins (TLR3 & TLR4 and HLA alleles). The docking scores along with the confidence score of the complexes and the binding affinity of the complexes are given in Table 3.

Molecular dynamics simulation of the Apo-vaccine and complex

The stability of the modeled structure along with the protein–protein complexes were subjected to a 10 ns simulation in the explicit water using SiBioLead which is an online biocomputing platform for computational drug discovery which uses GROMACS (Groningen Machine for Chemical Simulations). The modeled structure attained stability after 4 to 6 ns and level of stability fluctuated till 10 ns time period (Fig. 5A). The comparison of complexes and the modeled proteins showed that the complex attained better stability than the modeled structure (Fig. 5A). The gyration graph supported the RMSD of the complex and the vaccine proteins (Fig. 5B). Further, Fig. 5C represented the Root mean square fluctuation of the modeled vaccine target obtained, whereas Fig. 5D represented the RMSF of the bounded complex. There is a break between the residues ranging from 600 to 700 showing the bounded state between the TLRs and the vaccine (Fig. 5D). TLR3 showed minimal fluctuations compared with the TLR4. When compared with the apoprotein, the fluctuations in TLRs-vaccine are minimal, possibly due to the bounded state.

Conformational analysis of the HLA allele and the vaccine construct

The HLA alleles bound with the modeled vaccine were retrieved from HDock. A total of nine complexes were selected for the docking strategy (Table 3). Based on the docking and confidence score, four complexes were analyzed further for the molecular dynamics simulation. The protein complex with the respective PDB ID's 7M8S, 5VUF, 6AT5 and 6UZP showed highest docking scores ranging from -256.23 to -253.37 (Table 3). Whereas the confidence scores varied from 0.89 to 0.8877 (Table 3). While the numbers appear similar, the decimal points were considered in distinguishing the scores. The HLA alleles with the highest confidence scores, also exhibited higher docking scores when compared to the other docked complexes. Therefore, we selected the four HLA alleles that corresponded to the highest confidence scores, and higher docking scores, for subsequent molecular dynamics simulations. The RMSD deviated initially to attain equilibrium at around 8 ns (Fig. 6A). The complex of vaccine and 5VUF, showed stable RMSD after 7 ns around 2.55 nm (Fig. 6B). The same impression was observed in the complex vaccine–6AT5 (Fig. 6C) and vaccine-6UZP (Fig. 6D). Further, the radius of gyration is presented in the supplementary Figure S3 showing the correlation between the RMSD and Radius of gyration. The Hydrogen bond of the complexes presented in the supplementary Fig. S4 indicated the strength of the binding between the vaccine and the HLA alleles throughout the simulation period. Moreover, the RMSD, Radius of Gyration and Hydrogen bond observations were correlated with each other. These analyses strongly suggested a robust interaction between the vaccine and the alleles during the simulation period with the root mean square deviation getting stabilized after the initial equilibrium state. The hydrogen bond interaction showed no loss of hydrogen bonds indicating stable binding. All these features indicated that the HLA alleles when bound with the vaccine exhibited strong stability proving it to be the potential candidate.

S. no.	Protein	PDB ID	Docking score	Confidence score	Binding affinity (kcal/mol)
1	TLR3	1ZIW	-318.95	0.9670	-13.7
2	TLR4	3FX1	-351.41	0.9825	-14.2
3	HLA A 01:01	4NQX	-237.59	0.8522	-12.1
4	HLA B 57:01	5VUF	-255.20	0.8913	-13.1
5	HLA A 02:06	3OXR	-245.72	0.8715	-12.7
6	HLA A 02:03	3OX8	-245.65	0.8714	-12.6
7	HLA B 07:02	6AT5	-254.49	0.8899	-12.9
8	HLA A 02:01	7M8S	-256.23	0.8933	-13.1
9	HLA A 68:01	6PBH	-248.79	0.8782	-12.5
10	HLA B 15:01	6UZP	-253.37	0.8877	-12.7
11	HLA A 11:01	5WJL	-240.03	0.8782	-11.6

Table 3. Docking scores obtained during the interaction analysis of vaccine with TLRs and HLA alleles.

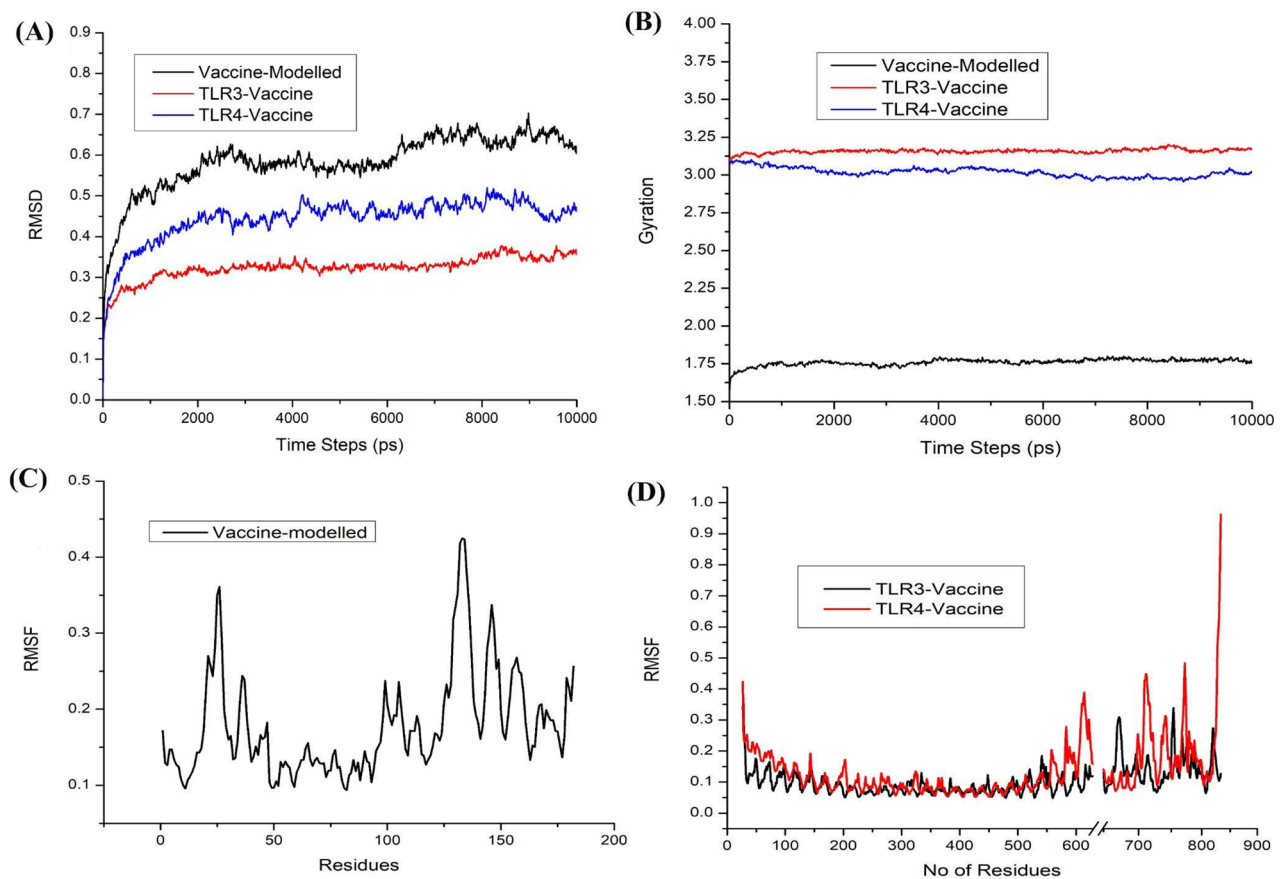


Figure 5. Molecular dynamics simulations of Apo-vaccine and the complex. (A) RMSD and (B) radius of gyration of the modeled vaccine and docked complex of TLR3-vaccine and TLR4-vaccine. (C) RMSF of the modelled vaccine and (D) RMSF of the docked complex TLR3-vaccine and TLR4-vaccine.

In silico cloning of vaccine construct

Codon adaptation is a method employed to enhance the expression rates of bacterial genomes by optimizing codons. In the present study, this strategy was implemented to increase the synthesis rate of a vaccine construct in the *E. coli* K12 system, considering the divergent codon usage between humans and the target host. To achieve this, the JCat webservice was used to back-translate and modify the codons of the vaccine construct sequence. Analysis of the modified sequence indicated 52% GC content, which fell within the acceptable range of 30–70%. Additionally, the codon adaptation index was found to be 0.98, suggesting a high potential for the expression of the vaccine construct.

Immune simulation of predicted vaccine construct

The vaccine construct was also analyzed using the C-ImmSim server to simulate the natural immune response in the human body. Upon encountering antigens, the immune reaction is initiated, leading to the production of IgM antibodies, with a smaller proportion of IgG antibodies (Fig. 7A). The production of antibodies is typically low initially and depends on the type of antigenic pathogen. We injected three doses of the vaccine at intervals of three months. By administering the vaccine dosage (antigens) for the first time, the levels of IgM antibodies started to rise (Fig. 7A). Subsequent antigen exposures with different time intervals resulted in a robust immune response characterized by higher levels of IgM and IgG antibodies (Fig. 7A). The levels of IgM, IgG1, and IgG2 antibodies significantly increased, indicating a stronger binding affinity of immunoglobulins for the vaccine construct (antigens) and the establishment of immunological memory. This enhanced immune response leads to improved pathogen elimination upon additional antigenic doses.

The B-cell, CD8+ and CD4+ T-cells exhibited a strong reaction in different cells, accompanied by the formation of memory cells (Fig. 7B–D). The frequency of Helper T cells remained higher throughout the exposure (Fig. 7C). The levels of cytokines during the immune reaction were assessed to determine any potential risks that might lead to complications in the human system, using the Simpson index D graph (Fig. 7E). The low D value indicated that the vaccine construct is safe to use.

For allergenicity assessment of the vaccine, we employed the AlgPred server, which integrates six distinct approaches: (i) IgE mapping, (ii) MEME/Mast motif analysis, (iii, iv) SVM modules based on both amino acid and dipeptide composition, (v) BLAST search against ARPs, and (vi) a hybrid approach combining all parameters. Among these approaches, MEME/Mast motif analysis, BLAST search against ARPs, and IgE mapping consistently indicated that the vaccine construct is non-allergenic. However, SVM modules, using both amino

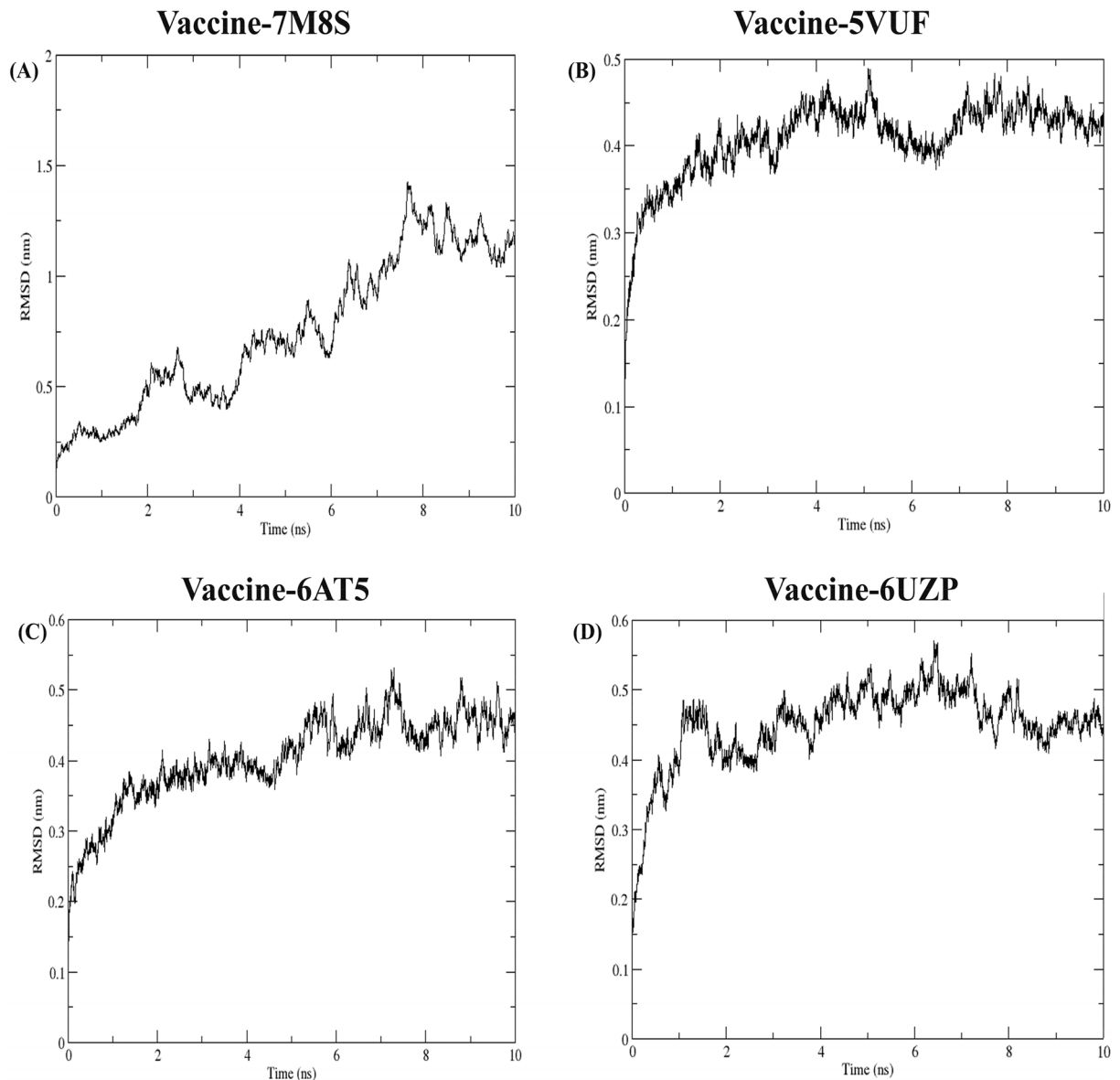


Figure 6. RMSD of the vaccine and the four HLA alleles docked through the HDock web server. (A) Vaccine-7M8S complex, (B) Vaccine-5VUF complex, (C) Vaccine-6AT5, (D) Vaccine-6UZP complex.

acid and dipeptide compositions, as well as the hybrid approach, suggested allergenicity of the vaccine construct. To ensure a comprehensive evaluation, we further conducted allergenicity analysis using the AllerTop server⁸⁹. The results obtained from this tool corroborated the non-allergenic nature of the vaccine. In light of the combined results from both servers, we concluded that the vaccine candidate is likely non-allergenic, thereby minimizing the risk of allergic reactions during vaccination protocols.

Analysis of VAH, ALT2, and the fusion of two proteins

To assess the usability of the analysis pipeline for vaccine predictions we tested it with known vaccine targets: VAH (AAD16985), ALT-2 (AAC35355), and the fusion of these two proteins (VAH-ALT-2), created by linking the two proteins with a GGGGS(3) linker. These proteins were selected as multivalent vaccine formulation of *B. malayi* VAH and ALT-2 provided higher protection compared to monovalent formulations of each protein³⁹.

Both T- and B-cell epitopes of each protein were identified. Interestingly, the T/B cell epitopes of the fused protein were nearly identical to those of VAH and ALT-2 (Supplementary Table S4). Tertiary structure models were developed for all three proteins, and the rationality of their structures was determined using Ramachandran plots (Supplementary Fig. S5). For ALT-2, where no similar sequences were found in the protein database, i-TASSER-MTD was employed for modeling. In the case of VAH and the fusion protein, similarity was observed with the *B. malayi* venom allergen-like protein-1 (ID: 6ANY) from the Protein Data Bank, leading to the use of MODELLER version 10.4 for modeling⁹⁰. The modeled proteins were then analyzed for interaction with TLR4, TLR3 human proteins, and HLA alleles used for the new vaccine construct. The HDock docking protocol revealed that the fusion protein exhibited the highest negative score in most complexes (Table 4), indicating

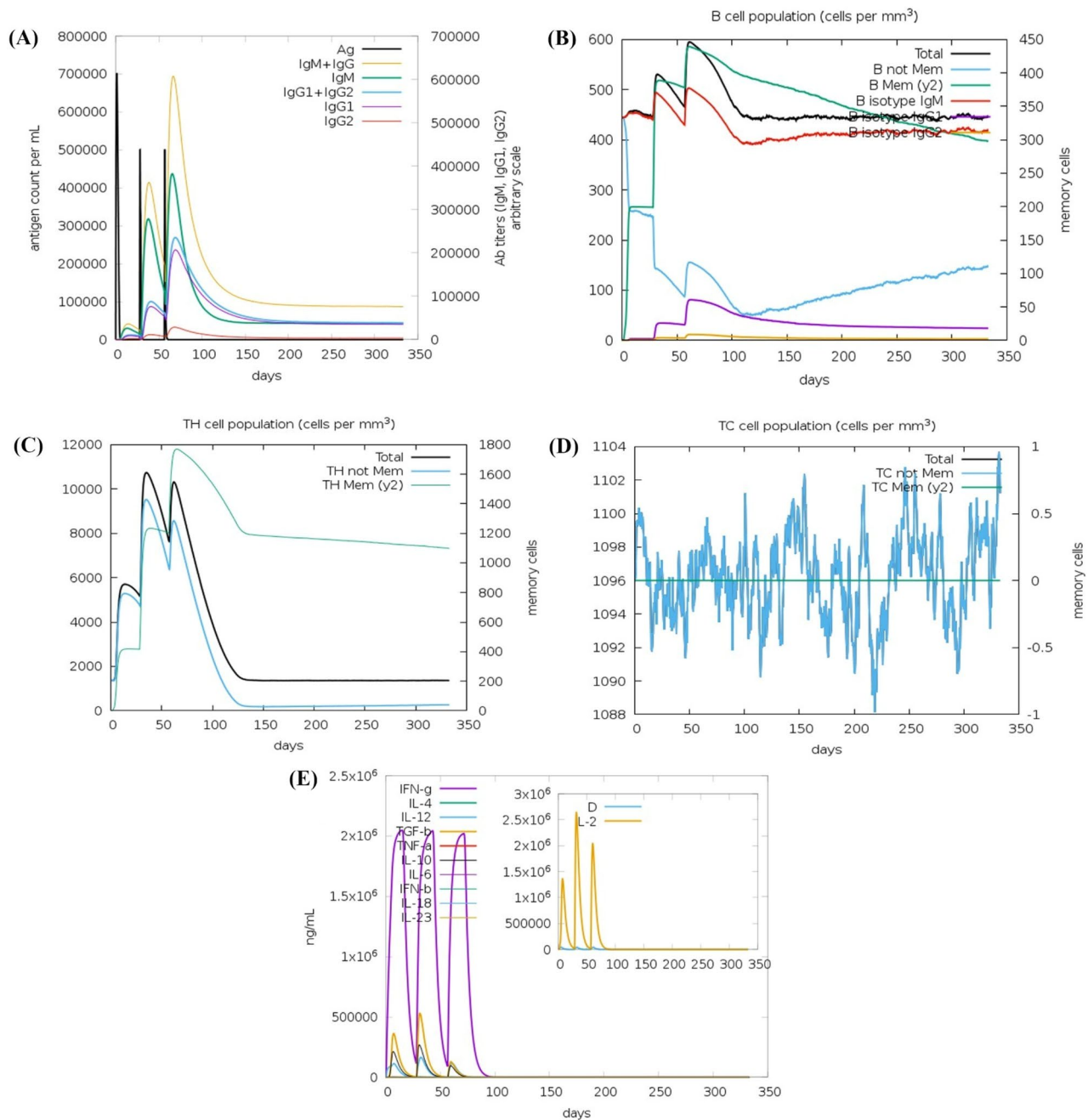


Figure 7. C-ImmSim plots showing the immune response of the multi-epitope vaccine construct. (A) The immunoglobulins and their subclasses in response to the injection of the vaccine construct (antigen) at three different time intervals; Antigen inoculations are shown as black vertical lines and immunoglobulin subclasses are indicated as colored peaks. (B) B-cell populations following three injections. (C) The population of T-helper cells. (D) The evolution of CD4 T-helper cell classes during the course of vaccination. (E) Levels of cytokine and interleukin after three injections; inset plot shows danger signal and interleukin growth factor IL-2.

strong binding energy and favorable interaction with human proteins. The analysis indicated that fusion protein had the combination of the epitopes from both the genes and, importantly, improved binding affinity to human proteins compared to the individual protein. This comprehensive analysis demonstrated the effectiveness of the in-silico methods employed in this study in predicting and validating the proposed vaccine construct.

Furthermore, all the three proteins were examined for assessing their potential to elicit immune response using C-ImmSim server. The VAH and the VAH-ALT-2 exhibited a comparable immune response pattern, while ALT-2 showed distinct characteristics. Specifically, ALT-2 displayed a decreased level of IgM compared to the other two proteins (Fig. 8A). The secondary and tertiary responses (IgG + IgM, and IgG1 + IgG2 antibodies) were more robust than the primary response for ALT-2 (Fig. 8A). Despite the enhanced secondary and tertiary responses in ALT-2, concerns arose as Interleukin-2 (IL-2) levels were observed to decrease (Fig. 8B). IL-2 is

S. no.	Protein	PDB ID	Docking score		
			AAC35355.1	AAD16985	Fusion vaccine
1	TLR3	1ZIW	-320.48	-318.54	-342.80
2	TLR4	3FX1	-315.14	-293.76	-326.63
3	HLA A 01:01	4NQX	-310.29	-291.10	-341.44
4	HLA B 57:01	5VUF	-298.33	-269.47	-316.32
5	HLA A 02:06	3OXR	-301.10	-317.54	-316.32
6	HLA A 02:03	3OX8	-282.33	-305.58	-294.93
7	HLA B 07:02	6AT5	-269.88	-258.83	-297.47
8	HLA A 02:01	7M8S	-280.93	-278.36	-297.88
9	HLA A 68:01	6PBH	-274.22	-312.78	-286.76
10	HLA B 15:01	6UZP	-303.66	-272.48	-316.70
11	HLA A 11:01	5WJL	-295.73	-315.71	-318.68

Table 4. Docking scores obtained during the interaction analysis of multi epitope fusion vaccine with TLRs and HLA alleles.

a critical signaling molecule in the immune system, important for the activation and proliferation of CD4+ T helper cells⁹¹. These cells are involved in coordinating immune responses, providing assistance to other immune cells, and facilitating the differentiation of B cells into antibody-producing plasma cells. Additionally, increased Transforming Growth Factor-beta (TGF- β) levels were noted in ALT-2 (Fig. 8B). TGF- β plays a dual role exhibiting both immunosuppressive and immune-regulatory effects in vaccines and the immune response⁹². It is recognized to inhibit the proliferation of T cells, including CD4+ and CD8+ T cells. The observed rise in TNF- β suggests a complex interplay of factors in the immune response to ALT-2. Furthermore, ALT-2 exhibited a lesser degree of cell duplication of CD4+ T regulatory cells compared to the other two proteins (Fig. 8C). Moreover, the level of antigen internalization by macrophages and by the B-cell lymphocytes showed a spike in ALT-2, but the lower level of antigen presentation was observed (Fig. 8D, E). In summary, while ALT-2 demonstrated favorable secondary and tertiary immune responses, the simultaneous decrease in IL-2 levels and alterations in regulatory T cell populations raises questions about the regulatory mechanisms associated with ALT-2, potentially influencing the balance of immune responses. These findings indicate the potential superiority of the fused protein over ALT-2. However, VAH showed a similar response pattern as observed in the fused protein based on this analysis.

Conclusion

In conclusion, our study sheds light on the promising role of in silico approaches in the identification of vaccine candidates for the management of filariasis. The in-silico analysis with the goal to identify therapeutic targets offers several advantages. Firstly, it reduces the time required to identify potential targets for medications and vaccines, allowing for a more efficient development process with fewer trials. Additionally, it helps in reducing costs associated with experimental testing. While we acknowledge the unfortunate limitation of any in silico approach, which inherently relies on predictions, the significance of this research lies in its potential to advance our understanding of the disease and provide valuable insights for further experimental studies. The candidates we have identified through computational analysis represent a crucial starting point in the quest for an effective vaccine against filariasis. It is essential to emphasize that the transition from in-silico predictions to real-world vaccine development is a complex and multifaceted process that demands rigorous experimental validation. However, the candidate identified here provide a valuable foundation upon which future research and vaccine development efforts can be built. Future research should focus on extensive experimental studies, immunogenicity assessments in diverse populations, optimization of vaccine formulations, and long-term efficacy evaluations. Additionally, collaborative efforts, epidemiological studies, and exploration of combination strategies will further enhance the prospects of developing a successful filariasis vaccine. However, this holistic approach ensures that the findings from in-silico predictions contribute substantively to the advancement of developing effective preventive measures against filariasis.

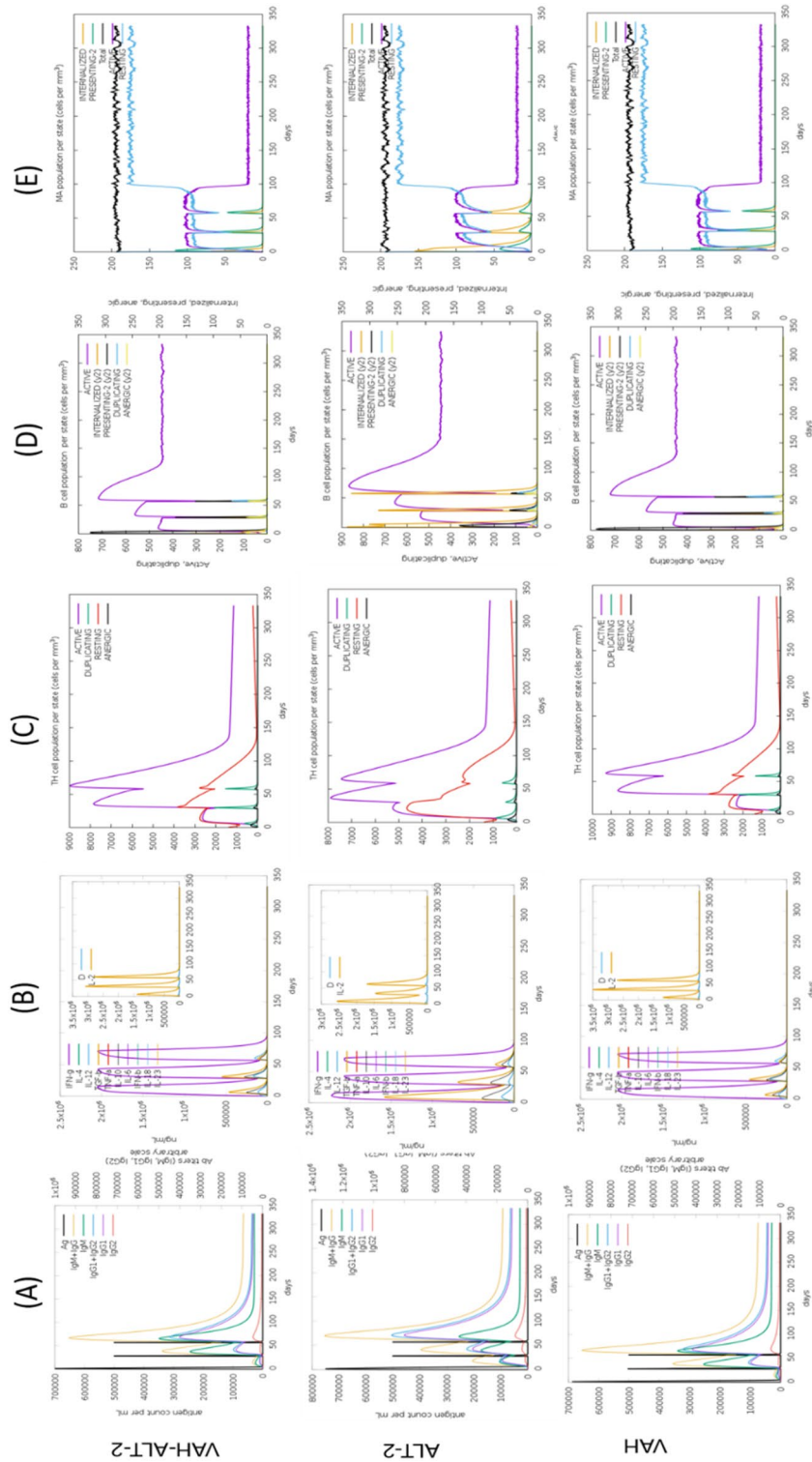


Figure 8. In-silico immune simulation analysis of the candidate vaccine targets, ALT-2, VAH and fused protein (VAH-ALT-2). The simulation was performed for three injections at time steps of 1, 84, and 170. (A) Antigen and immunoglobulins (antibodies are sub-divided per isotype); (B) concentration of cytokines and interleukins. D in the inset plot is danger signal; (C) CD4 T-helper lymphocyte count sub-divided per entity (i.e. active, resting, anergic and duplicating); (D) B-lymphocytes population per entity active, presenting on class-II, internalized the antigen, duplicating and anergic; (E) macrophage cell population.

Data availability

The datasets generated and/or analyzed during the current study are included in the text.

Received: 26 June 2023; Accepted: 18 January 2024

Published online: 23 January 2024

References

- Pi-Bansa, S. *et al.* Potential factors influencing lymphatic filariasis transmission in “hotspot” and “control” areas in Ghana: The importance of vectors. *Infect. Dis. Poverty* **8**, 9. <https://doi.org/10.1186/s40249-019-0520-1> (2019).
- Small, S. T. *et al.* Population genetics of the filarial worm *Wuchereria bancrofti* in a post-treatment region of Papua New Guinea: Insights into diversity and life history. *PLoS Negl. Trop. Dis.* **7**, e2308. <https://doi.org/10.1371/journal.pntd.0002308> (2013).
- Shenoy, R. K. Clinical and pathological aspects of filarial lymphedema and its management. *Korean J. Parasitol.* **46**, 119–125. <https://doi.org/10.3347/kjp.2008.46.3.119> (2008).
- Ichimori, K. *et al.* Global programme to eliminate lymphatic filariasis: The processes underlying programme success. *PLoS Negl. Trop. Dis.* **8**, e3328. <https://doi.org/10.1371/journal.pntd.0003328> (2014).
- Organization, W. H. *Ending the Neglect to Attain the Sustainable Development Goals—A Road Map for Neglected Tropical Diseases 2021–2030*. Vol. 177 (World Health Organization, 2020).
- King, C. L. *et al.* A trial of a triple-drug treatment for lymphatic filariasis. *N. Engl. J. Med.* **379**, 1801–1810. <https://doi.org/10.1056/NEJMoa1706854> (2018).
- Yajima, A. & Ichimori, K. Progress in the elimination of lymphatic filariasis in the Western Pacific Region: Successes and challenges. *Int. Health* **13**, S10–S16. <https://doi.org/10.1093/inthealth/ihaa087> (2020).
- Gyapong, J. O., Owusu, I. O., da-Costa Vroom, F. B., Mensah, E. O. & Gyapong, M. Elimination of lymphatic filariasis. *Res. Rep. Trop. Med.* **9**, 25–33. <https://doi.org/10.2147/RRTM.S125204> (2018).
- Cross, J.H. Filarial nematodes. In *Medical Microbiology* (Baron S, ed.), 4th ed. (1996).
- Laman, M. *et al.* Mass drug administration of ivermectin, diethylcarbamazine, plus albendazole compared with diethylcarbamazine plus albendazole for reduction of lymphatic filariasis endemicity in Papua New Guinea: A cluster-randomised trial. *Lancet Infect. Dis.* **22**, 1200–1209. [https://doi.org/10.1016/S1473-3099\(22\)00026-3](https://doi.org/10.1016/S1473-3099(22)00026-3) (2022).
- Cobo, F. Determinants of parasite drug resistance in human lymphatic filariasis. *Rev. Esp. Quimioter.* **29**, 288–295 (2016).
- Noroës, J. *et al.* Assessment of the efficacy of diethylcarbamazine on adult *Wuchereria bancrofti* in vivo. *Trans. R. Soc. Trop. Med. Hyg.* **91**, 78–81. [https://doi.org/10.1016/S0035-9203\(97\)90405-3](https://doi.org/10.1016/S0035-9203(97)90405-3) (1997).
- Eberhard, M. L., Lammie, P. J., Dickinson, C. M. & Roberts, J. M. Evidence of nonsusceptibility to diethylcarbamazine in *Wuchereria bancrofti*. *J. Infect. Dis.* **163**, 1157–1160. <https://doi.org/10.1093/infdis/163.5.1157> (1991).
- Esterne, P., Plichart, C., Sechan, Y. & Nguyen, N. L. The impact of 34 years of massive DEC chemotherapy on *Wuchereria bancrofti* infection and transmission: The Maupiti cohort. *Trop. Med. Int. Health* **6**, 190–195. <https://doi.org/10.1046/j.1365-3156.2001.00682.x> (2001).
- Schwab, A. E., Boakye, D. A., Kyelem, D. & Prichard, R. K. Detection of benzimidazole resistance-associated mutations in the filarial nematode *Wuchereria bancrofti* and evidence for selection by albendazole and ivermectin combination treatment. *Am. J. Trop. Med. Hyg.* **73**, 234–238 (2005).
- Eberhard, M. L., Lowrie, R. C. Jr. & Lammie, P. J. Persistence of microfilaremia in bancroftian filariasis after diethylcarbamazine citrate therapy. *Trop. Med. Parasitol.* **39**, 128–130 (1988).
- Osei-Atweneboana, M. Y. *et al.* Phenotypic evidence of emerging ivermectin resistance in *Onchocerca volvulus*. *PLoS Negl. Trop. Dis.* **5**, e998. <https://doi.org/10.1371/journal.pntd.0000998> (2011).
- Andre, F. E. *et al.* Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull. World Health Organ.* **86**, 140–146. <https://doi.org/10.2471/blt.07.040089> (2008).
- Kalyanasundaram, R., Khatri, V. & Chauhan, N. Advances in vaccine development for human lymphatic filariasis. *Trends Parasitol.* **36**, 195–205. <https://doi.org/10.1016/j.pt.2019.11.005> (2020).
- Chauhan, N. *et al.* Improving the efficacy of a prophylactic vaccine formulation against lymphatic filariasis. *Parasitol. Res.* **116**, 2821–2830. <https://doi.org/10.1007/s00436-017-5593-9> (2017).
- Madanagopal, P., Muthusamy, S., Pradhan, S. N. & Prince, P. R. Construction and validation of a multi-epitope in silico vaccine model for lymphatic filariasis by targeting *Brugia malayi*: A reverse vaccinology approach. *Bull. Natl. Res. Center* **47**, 47. <https://doi.org/10.1186/s42269-023-01013-0> (2023).
- Shey, R. A. *et al.* In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci. Rep.* **9**, 4409. <https://doi.org/10.1038/s41598-019-40833-x> (2019).
- Das, N. C. *et al.* Designing of a novel multi-epitope peptide based vaccine against *Brugia malayi*: An in silico approach. *Infect. Genet. Evol.* **87**, 104633. <https://doi.org/10.1016/j.meegid.2020.104633> (2021).
- Gorai, S. *et al.* Designing efficient multi-epitope peptide-based vaccine by targeting the antioxidant thioredoxin of bancroftian filarial parasite. *Infect. Genet. Evol.* **98**, 105237. <https://doi.org/10.1016/j.meegid.2022.105237> (2022).
- Sharma, O. P., Vadlamudi, Y., Kota, A. G., Sinha, V. K. & Kumar, M. S. Drug targets for lymphatic filariasis: A bioinformatics approach. *J. Vector Borne Dis.* **50**, 155–162 (2013).
- Das, N. C., Gupta, P. S. S., Panda, S. K., Rana, M. K. & Mukherjee, S. Reverse vaccinology assisted design of a novel multi-epitope vaccine to target *Wuchereria bancrofti* cystatin: An immunoinformatics approach. *Int. Immunopharmacol.* **115**, 109639. <https://doi.org/10.1016/j.intimp.2022.109639> (2023).
- Rawal, K. *et al.* Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. *Sci. Rep.* **11**, 17626. <https://doi.org/10.1038/s41598-021-96863-x> (2021).
- Fadaka, A. O. *et al.* Immunoinformatics design of a novel epitope-based vaccine candidate against dengue virus. *Sci. Rep.* **11**, 19707. <https://doi.org/10.1038/s41598-021-99227-7> (2021).
- Goodswen, S. J., Kennedy, P. J. & Ellis, J. T. A state-of-the-art methodology for high-throughput in silico vaccine discovery against protozoan parasites and exemplified with discovered candidates for *Toxoplasma gondii*. *Sci. Rep.* **13**, 8243. <https://doi.org/10.1038/s41598-023-34863-9> (2023).
- Albutti, A. An integrated computational framework to design a multi-epitopes vaccine against *Mycobacterium tuberculosis*. *Sci. Rep.* **11**, 21929. <https://doi.org/10.1038/s41598-021-01283-6> (2021).
- Margaroni, M. *et al.* Immunoinformatics approach to design a multi-epitope nanovaccine against leishmania parasite: Elicitation of cellular immune responses. *Vaccines (Basel)* <https://doi.org/10.3390/vaccines11020304> (2023).
- Anand, S. B. *et al.* Comparison of immunogenicity, protective efficacy of single and cocktail DNA vaccine of *Brugia malayi* abundant larval transcript (ALT-2) and thioredoxin peroxidase (TPX) in mice. *Acta Trop.* **107**, 106–112. <https://doi.org/10.1016/j.actatropica.2008.04.018> (2008).
- Anand, S. B. *et al.* Immune response studies with *Wuchereria bancrofti* vespider allergen homologue (WbVAH) in human lymphatic filariasis. *Parasitol. Res.* **101**, 981–988. <https://doi.org/10.1007/s00436-007-0571-2> (2007).

34. Dakshinamoorthy, G. *et al.* Biochemical characterization and evaluation of a *Brugia malayi* small heat shock protein as a vaccine against lymphatic filariasis. *PLoS One* **7**, e34077. <https://doi.org/10.1371/journal.pone.0034077> (2012).
35. Veerapathran, A., Dakshinamoorthy, G., Gnanasekar, M., Reddy, M. V. & Kalyanasundaram, R. Evaluation of *Wuchereria bancrofti* GST as a vaccine candidate for lymphatic filariasis. *PLoS Negl. Trop. Dis.* **3**, e457. <https://doi.org/10.1371/journal.pntd.0000457> (2009).
36. Dakshinamoorthy, G., Samyikutty, A. K., Munirathinam, G., Reddy, M. V. & Kalyanasundaram, R. Multivalent fusion protein vaccine for lymphatic filariasis. *Vaccine* **31**, 1616–1622. <https://doi.org/10.1016/j.vaccine.2012.09.055> (2013).
37. Chauhan, N., Khatri, V., Banerjee, P. & Kalyanasundaram, R. Evaluating the vaccine potential of a tetravalent fusion protein (rBm-HAXT) vaccine antigen against lymphatic filariasis in a mouse model. *Front. Immunol.* **9**, 1520. <https://doi.org/10.3389/fimmu.2018.01520> (2018).
38. Khatri, V. *et al.* Prospects of developing a prophylactic vaccine against human lymphatic filariasis—Evaluation of protection in non-human primates. *Int. J. Parasitol.* **48**, 773–783. <https://doi.org/10.1016/j.ijpara.2018.04.002> (2018).
39. Kalyanasundaram, R. & Balumuri, P. Multivalent vaccine formulation with BmVAL-1 and BmALT-2 confer significant protection against challenge infections with *Brugia malayi* in mice and jirds. *Res. Rep. Trop. Med.* **45–56**, 2011. <https://doi.org/10.2147/RRTM.S13679> (2011).
40. Sanches, R. C. O. *et al.* Immunoinformatics design of multi-epitope peptide-based vaccine against *Schistosoma mansoni* using transmembrane proteins as a target. *Front. Immunol.* **12**, 621706. <https://doi.org/10.3389/fimmu.2021.621706> (2021).
41. Zhang, L. Multi-epitope vaccines: A promising strategy against tumors and viral infections. *Cell Mol. Immunol.* **15**, 182–184. <https://doi.org/10.1038/cmi.2017.92> (2018).
42. Zhang, R., Ou, H. Y. & Zhang, C. T. DEG: A database of essential genes. *Nucleic Acids Res.* **32**, D271–272. <https://doi.org/10.1093/nar/gkh024> (2004).
43. Horton, P. *et al.* WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **35**, W585–587. <https://doi.org/10.1093/nar/gkm259> (2007).
44. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423. <https://doi.org/10.1038/s41587-019-0036-z> (2019).
45. Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G. & Brunak, S. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* **17**, 349–356. <https://doi.org/10.1093/protein/gzh037> (2004).
46. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **305**, 567–580. <https://doi.org/10.1006/jmbi.2000.4315> (2001).
47. Doytchinova, I. A. & Flower, D. R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **8**, 4. <https://doi.org/10.1186/1471-2105-8-4> (2007).
48. Magnan, C. N. *et al.* High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics* **26**, 2936–2943. <https://doi.org/10.1093/bioinformatics/btq551> (2010).
49. Larsen, M. V. *et al.* Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinform.* **8**, 424. <https://doi.org/10.1186/1471-2105-8-424> (2007).
50. Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339–D343. <https://doi.org/10.1093/nar/gky1006> (2019).
51. Saha, S. & Raghava, G. P. Prediction methods for B-cell epitopes. *Methods Mol. Biol.* **409**, 387–394. https://doi.org/10.1007/978-1-60327-118-9_29 (2007).
52. SaR Saha, G. P. S. BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In *Artificial Immune Systems: Third International Conference, ICARIS 2004, Catania, Sicily, Italy. Proceedings 3*. 197–204, September 13–16, 2004 (2004).
53. Jespersen, M. C., Peters, B., Nielsen, M. & Marcatili, P. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **45**, W24–W29. <https://doi.org/10.1093/nar/gkx346> (2017).
54. Saha, S. & Raghava, G. P. AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res.* **34**, W202–209. <https://doi.org/10.1093/nar/gkl343> (2006).
55. Nguyen, M. N. *et al.* AllerCatPro 2.0: A web server for predicting protein allergenicity potential. *Nucleic Acids Res.* **50**, W36–W43. <https://doi.org/10.1093/nar/gkac446> (2022).
56. Maurer-Stroh, S., Eisenhaber, B. & Eisenhaber, F. N-terminal N-myristoylation of proteins: Refinement of the sequence motif and its taxon-specific differences. *J. Mol. Biol.* **317**, 523–540. <https://doi.org/10.1006/jmbi.2002.5425> (2002).
57. Eisenhaber, B., Bork, P. & Eisenhaber, F. Prediction of potential GPI-modification sites in proprotein sequences. *J. Mol. Biol.* **292**, 741–758. <https://doi.org/10.1006/jmbi.1999.3069> (1999).
58. UniProt, C. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489. <https://doi.org/10.1093/nar/gkaa1100> (2021).
59. Zhou, X. *et al.* I-TASSER-MTD: A deep-learning-based platform for multi-domain protein structure and function prediction. *Nat. Protoc.* **17**, 2326–2353. <https://doi.org/10.1038/s41596-022-00728-0> (2022).
60. Laskowski, R. A. *et al.* PDBsum: A web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.* **22**, 488–490. [https://doi.org/10.1016/s0968-0004\(97\)01140-7](https://doi.org/10.1016/s0968-0004(97)01140-7) (1997).
61. Choe, J., Kelker, M. S. & Wilson, I. A. Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* **309**, 581–585. <https://doi.org/10.1126/science.1115253> (2005).
62. Park, B. S. *et al.* The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **458**, 1191–1195. <https://doi.org/10.1038/nature07830> (2009).
63. Yan, Y., Tao, H., He, J. & Huang, S. Y. The HDock server for integrated protein–protein docking. *Nat. Protoc.* **15**, 1829–1852. <https://doi.org/10.1038/s41596-020-0312-x> (2020).
64. Yan, Y., Zhang, D., Zhou, P., Li, B. & Huang, S. Y. HDock: A web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* **45**, W365–W373. <https://doi.org/10.1093/nar/gkx407> (2017).
65. Romero-Molina, S. *et al.* PPI-Affinity: A web tool for the prediction and optimization of protein–peptide and protein–protein binding affinity. *J. Proteome Res.* **21**, 1829–1841. <https://doi.org/10.1021/acs.jproteome.2c00020> (2022).
66. Ganne, A. *et al.* In silico analysis of TUBA4A mutations in Amyotrophic Lateral Sclerosis to define mechanisms of microtubule disintegration. *Sci. Rep.* **13**, 2096. <https://doi.org/10.1038/s41598-023-28381-x> (2023).
67. Grote, A. *et al.* JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* **33**, W526–531. <https://doi.org/10.1093/nar/gki376> (2005).
68. Rapin, N., Lund, O., Bernaschi, M. & Castiglione, F. Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One* **5**, e9862. <https://doi.org/10.1371/journal.pone.0009862> (2010).
69. Bendtsen, J. D., Kiemer, L., Fausboll, A. & Brunak, S. Non-classical protein secretion in bacteria. *BMC Microbiol.* **5**, 58. <https://doi.org/10.1186/1471-2180-5-58> (2005).
70. Popot, J.-L. Integral membrane protein structure: Transmembrane α -helices as autonomous folding domains. *Curr. Opin. Struct. Biol.* **3**, 8 (1993).
71. Lin, S. H. & Guidotti, G. Purification of membrane proteins. *Methods Enzymol.* **463**, 619–629. [https://doi.org/10.1016/S0076-6879\(09\)63035-4](https://doi.org/10.1016/S0076-6879(09)63035-4) (2009).

72. Frand, A. R., Russel, S. & Ruvkun, G. Functional genomic analysis of *C. elegans* molting. *PLoS Biol.* **3**, e312. <https://doi.org/10.1371/journal.pbio.0030312> (2005).
73. McNulty, S. N., Mitreva, M., Weil, G. J. & Fischer, P. U. Inter and intra-specific diversity of parasites that cause lymphatic filariasis. *Infect. Genet. Evol.* **14**, 137–146. <https://doi.org/10.1016/j.meegid.2012.11.002> (2013).
74. Gottschalk, A. *et al.* Identification and characterization of novel nicotinic receptor-associated proteins in *Caenorhabditis elegans*. *EMBO J.* **24**, 2566–2578. <https://doi.org/10.1038/sj.emboj.7600741> (2005).
75. Simmer, F. *et al.* Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain reveals novel gene functions. *PLoS Biol.* **1**, E12. <https://doi.org/10.1371/journal.pbio.0000012> (2003).
76. Sonnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469. <https://doi.org/10.1038/nature03353> (2005).
77. Jmel, M. A. *et al.* Tick salivary Kunitz-type inhibitors: Targeting host hemostasis and immunity to mediate successful blood feeding. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms24021556> (2023).
78. Ranasinghe, S. L., Duke, M., Harvie, M. & McManus, D. P. Kunitz-type protease inhibitor as a vaccine candidate against schistosomiasis mansonii. *Int. J. Infect. Dis.* **66**, 26–32. <https://doi.org/10.1016/j.ijid.2017.10.024> (2018).
79. Ascenzi, P. *et al.* The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): A milestone protein. *Curr. Protein Pept. Sci.* **4**, 231–251. <https://doi.org/10.2174/1389203033487180> (2003).
80. Silvano, L. *et al.* A vaccine based on Kunitz-type molecule confers protection against *Fasciola hepatica* challenge by inducing IFN-gamma and antibody immune responses through IL-17A production. *Front. Immunol.* **11**, 2087. <https://doi.org/10.3389/fimmu.2020.02087> (2020).
81. Bozas, S. E. *et al.* Characterisation of a novel Kunitz-type molecule from the trematode *Fasciola hepatica*. *Mol. Biochem. Parasitol.* **74**, 19–29. [https://doi.org/10.1016/0166-6851\(95\)02478-6](https://doi.org/10.1016/0166-6851(95)02478-6) (1995).
82. Ranasinghe, S. L., Fischer, K., Zhang, W., Gobert, G. N. & McManus, D. P. Cloning and characterization of two potent Kunitz type protease inhibitors from *Echinococcus granulosus*. *PLoS Negl. Trop. Dis.* **9**, e0004268. <https://doi.org/10.1371/journal.pntd.0004268> (2015).
83. Hernandez-Goenaga, J. *et al.* Peptides derived of Kunitz-type serine protease inhibitor as potential vaccine against experimental schistosomiasis. *Front. Immunol.* **10**, 2498. <https://doi.org/10.3389/fimmu.2019.02498> (2019).
84. Islam, M. K. *et al.* The Kunitz-like modulatory protein haemangin is vital for hard tick blood-feeding success. *PLoS Pathog.* **5**, e1000497. <https://doi.org/10.1371/journal.ppat.1000497> (2009).
85. Pandey, R. K., Bhatt, T. K. & Prajapati, V. K. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. *Sci. Rep.* **8**, 1125. <https://doi.org/10.1038/s41598-018-19456-1> (2018).
86. Magnan, C. N., Randall, A. & Baldi, P. SOLpro: Accurate sequence-based prediction of protein solubility. *Bioinformatics* **25**, 2200–2207. <https://doi.org/10.1093/bioinformatics/btp386> (2009).
87. Wilkins, M. R. *et al.* Protein identification and analysis tools in the Expasy server. *Methods Mol. Biol.* **112**, 531–552. <https://doi.org/10.1385/1-59259-584-7:531> (1999).
88. Zarling, A. L. *et al.* Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. *J. Exp. Med.* **192**, 1755–1762. <https://doi.org/10.1084/jem.192.12.1755> (2000).
89. Dimitrov, I., Flower, D. R. & Doytchinova, I. AllerTOP—A server for in silico prediction of allergens. *BMC Bioinform.* **14**(6), S4. <https://doi.org/10.1186/1471-2105-14-S6-S4> (2013).
90. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **86**, 2.9.1–2.9.37. <https://doi.org/10.1002/cpps.20> (2016).
91. Ross, S. H. & Cantrell, D. A. Signaling and function of interleukin-2 in T lymphocytes. *Annu. Rev. Immunol.* **36**, 411–433. <https://doi.org/10.1146/annurev-immunol-042617-053352> (2018).
92. Sanjabi, S., Oh, S. A. & Li, M. O. Regulation of the immune response by TGF-beta: From conception to autoimmunity and infection. *Cold Spring Harb. Perspect. Biol.* <https://doi.org/10.1101/cshperspect.a022236> (2017).

Acknowledgements

The authors would like to thank Indian council of Medical Research (ICMR), New Delhi for extramural funding.

Author contributions

M.A. performed Data retrieval, analysis and writing. G.N.P. performed data analysis, writing and visualization, R.P. carried out manuscript review, A.K. performed manuscript review and provided resources, B.G. performed conceptualization, study designing, data analysis, fund acquisition, overall project monitoring and management, manuscript writing, review and visualization.

Funding

This study has been funded by Indian Council of Medical Research (ICMR), New Delhi under extramural Adhoc Projects No./9-7(247)2020/ECD-II.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-52457-x>.

Correspondence and requests for materials should be addressed to B.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024