

RESEARCH

Open Access



Bioinformatics-guided decoding of the *Ancylostoma duodenale* genome for the identification of potential vaccine targets

Mohibullah Shah^{1,2*}, Hira Anum^{1†}, Asifa Sarfraz¹, Md. Aktaruzzaman³, Al Riyad Hasan³, Muhammad Umer Khan⁴, Khaled Fahmi Fawy^{5,6}, Sarah A. Altwaim^{7,8}, Saeed M. N. Alasmari⁹, Abid Ali¹⁰, Umar Nishan¹¹ and Ke Chen^{12*}

Abstract

Ancylostoma duodenale, a parasitic nematode worm, is found to be involved in various infections, including intestinal blood loss, protein malnutrition, and anemia. Antimicrobial resistance to the available therapeutics has prompted the search for new drug and vaccine targets against *A. duodenale*. Despite significant advances in vaccine development against *A. duodenale*, no commercial and FDA-approved vaccine exists to safeguard humans from infections caused by this pathogen. In this investigation, a stringent bioinformatics analysis identified 36 unique essential and host-interacting proteins. Based on their subcellular localization, 6 proteins located in the extracellular space and outer membrane were categorized as vaccine targets, while the remaining proteins were predicted to act as potential drug candidates. These vaccine candidates were further assessed for antigenicity, allergenicity, and physicochemical analysis to determine their suitability for the designing of a multi-epitope vaccine. Two candidate proteins were chosen as optimal targets in the development of vaccine design. The identified T- and B-cell epitopes from these proteins were then combined with appropriate linkers and adjuvants to design chimeric vaccine constructs aimed at inducing both cellular and humoral immune responses. Molecular docking, molecular dynamic simulations, PCA analysis, DCCM analysis, and binding free energy calculations proved stable interactions of the designed vaccine with human immune cell receptors. Within a bacterial cloning system, the vaccine constructs demonstrated the ability to be cloned and expressed. The immunological stimulation elicited significant immunological responses to the proposed vaccine. Our investigation identified new therapeutic targets and developed a peptide-based multi-epitope vaccine against *A. duodenale* infection. Additional experimental verification will open up new therapeutic alternatives for this emerging resistant pathogen.

Keywords Antiparasitic, Anemia, Blood, Therapeutic agent, T-cell

[†]Mohibullah Shah and Hira Anum contributed equally to this work.

*Correspondence:

Mohibullah Shah

mohib@bzu.edu.pk; mohibusb@gmail.com

Ke Chen

chen_ke@swmu.edu.cn

Full list of author information is available at the end of the article



Introduction

Ancylostoma duodenale, commonly known as the Old World hookworm, is a parasitic nematode (roundworm) belonging to the family *Ancylostomatidae*. It was discovered in 1838 by James Paget, who identified hookworms in the intestines of a deceased individual during a post-mortem examination [1]. It is a parasitic, blood-feeding nematode that primarily inhabits the small intestine of its host, typically birds or mammals [2]. Hookworms are transmitted when eggs excreted in human feces mature in suitable soil conditions, hatching into infective larvae that penetrate the host's skin, typically through minor abrasions. *A. duodenale* can infect human hosts through both oral ingestion and percutaneous penetration. During its larval stage, the parasite may undergo arrested development within host tissues [3].

Hookworm infections, including those caused by *A. duodenale* remain a major public health concern. In 2015, hookworm infections accounted for approximately 1.8 million disability-adjusted life years (DALYs), underscoring their substantial global disease burden [1]. In 2016, *A. duodenale* accounted for a substantial proportion of global hookworm infections, affecting approximately 450.68 million people worldwide - a notable decrease from 1.297 billion cases reported in 1994. Approximately 1.5 billion people worldwide are at risk of *Ancylostoma* and other soil-transmitted helminth infections. The highest prevalence occurs in Asia and the Pacific region, where tropical climates, overcrowding, and poor sanitation conditions facilitate transmission [4]. Preschool and school-aged children, along with travelers returning from endemic regions, represent the highest-risk populations for *Ancylostoma* infection. In China and other *A. duodenale*-endemic areas, maternal hookworm infection may lead to vertical transmission to neonates through ingestion of third-stage larvae (L3) present in breast milk and colostrum [5]. *A. duodenale* caused significant outbreaks among workers during the construction of the St. Gotthard Tunnel in the 19th century. The parasite is clinically associated with iron-deficiency anemia, historically referred to as 'Egyptian chlorosis' in endemic regions [6].

A. duodenale causes ancylostomiasis [4], which primarily results in iron-deficiency anaemia [7], protein malnutrition, impaired cognitive development, stunted growth, and low physical fitness, especially in children [8]. Although blood loss through gastrointestinal bleeding is rare, approximately 0.05 to 0.3 ml blood is lost in adults, leading to varying degrees of anemia (mild, moderate, or severe) depending upon the parasitic load, i.e., the number of eggs eliminated per gram of feces [9]. These findings indicate that the disease burden is particularly high, especially among children and the populations residing

in tropical and subtropical regions, highlighting a need for an effective strategy to combat the infection.

Standard treatment for *A. duodenale* infection involves anthelmintic drugs including albendazole, pyrantel pamoate, ivermectin, and mebendazole. These medications effectively eliminate adult intestinal worms [10]. Albendazole exerts its anthelmintic effect by irreversibly inhibiting glucose uptake in parasites, resulting in energy depletion and subsequent death [11]. Clinical studies report mild, transient adverse effects including nausea, vomiting, abdominal pain, diarrhea, headache, and occasional hepatic transaminase elevation following albendazole administration [12]. Pyrantel pamoate acts as a depolarizing neuromuscular blocking agent, causing spastic paralysis in helminths. Its adverse effect profile may include gastrointestinal disturbances, neurological symptoms, and hepatotoxicity [13]. While these drugs demonstrate high efficacy against *A. duodenale* infections, their clinical application requires careful risk-benefit analysis due to: (1) the potential for adverse effects and (2) emerging drug resistance patterns. Notably, increasing albendazole resistance among intestinal helminths - particularly evident in pediatric populations - poses significant challenges for mass drug administration programs, as documented in Peru's national deworming initiative [14].

Despite the global impact of *A. duodenale*, no commercially licensed vaccine exists currently to prevent the infection. However, research is underway to produce vaccines that can prevent hookworm infections. Vaccines for *Ancylostoma duodenale*, including Na-APR-1 and Na-GST-1, target hookworm proteins but struggle with low production yields, adjuvant requirements, and multiple injections [15, 16]. Promising peptide candidates like A291Y-GCN4 and p3-P25 face challenges with immunogenicity and validation in challenge models. Previous studies on *A. duodenale* have investigated multiple aspects of its biology and epidemiology, including: life cycle characteristics, transmission dynamics, therapeutic efficacy of anthelmintic drugs, host-parasite immune interactions, and disease control strategies. Additionally, molecular epidemiological studies have examined genetic diversity patterns across geographically distinct *A. duodenale* populations [17]. Notably, no published studies have focused on developing a prophylactic vaccine against *A. duodenale* that could simultaneously address both the pathogen's clinical burden and the limitations of current anthelmintic therapies. To bridge this critical gap, our study aims to design a novel multi-epitope vaccine candidate using advanced immunoinformatics approaches, incorporating carefully selected antigenic peptides from key parasite proteins.

Materials and methods

This research employs a comprehensive framework that utilizes cutting-edge computational tools and techniques (Fig. 1) to aid vaccine designing against *A. duodenale*.

Among the various tools employed, NCBI, BLASTp, HPPPI, and DEG were essential for identifying non-redundant, pathogen-specific, and essential proteins. KEGG pathways analysis prioritized proteins involved in pathogen unique metabolic pathways. For epitope prediction and filtering, IEDB tools were significant in B-cell and T-cell epitope prediction, while AllerTOP, VaxiJen, and ToxinPred guided the final selection based on allergenicity, antigenicity, and toxicity. The vaccine structure was modeled and refined by using trRosetta and Galaxy-Refine Server. The ClusPro server performed molecular docking, and Desmond (Schrödinger) was used for molecular dynamics simulations, while C-ImmSim was used to assess immune response potential.

Retrieval of data and identification of essential proteins

The *Ancylostoma duodenale* reference genome (GenBank: GCA_000816745.1) was downloaded from the National Center for Biotechnology Information (NCBI). BLASTp searches were performed against the DEG [18] and HPPPI [19] databases to determine essential and

host-pathogen interacting proteins, with cutoff values of bitscore > 100, query coverage > 35, E-value < 1e-20, and percentage identity > 35. These parameters were selected to ensure that only high-confidence matches with significant biological relevance and alignment quality are retrieved. Conversely, proteins that did not meet these criteria were considered non-homologous and removed from further analysis. This stringent criteria reduced the inclusion of non-specific or weak matches and are consistent with thresholds commonly reported in recent literature [20].

Identification of host non-homologous proteins

To minimize potential adverse effects and ensure therapeutic safety in the host, drug and vaccine targets should be distinct from human proteins. Therefore, the essential and host-interacting proteins were filtered against the human genome by standalone BLASTp to identify the human non-homologous proteins. Additional scanning of *A. duodenale* proteins was performed against the human gut flora. The human gut microbiome, containing 75,176 proteins, was obtained from the NCBI database. BLASTp was run using standard values such bitscore ≤ 100, qcovs and pident ≤ 35, and E-value ≥ 1e-20 [21–23].

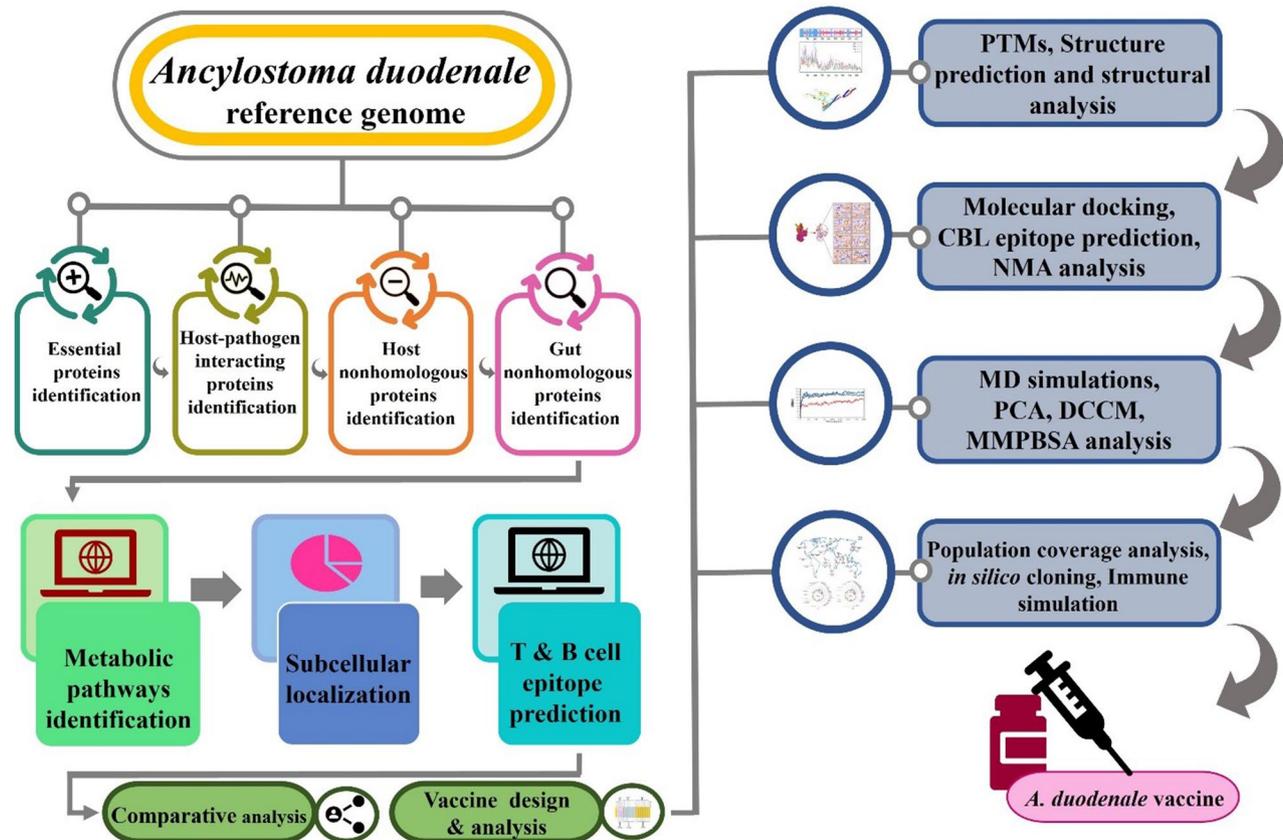


Fig. 1 Schematic illustration of approaches used to develop potential vaccines against *A. duodenale* infection

Metabolic pathways analysis

The Kyoto Encyclopedia of Genes and Genomes (KEGG) database provides access to the metabolic pathways of living organisms [24]. It was used to analyze the short-listed proteins to determine their associated metabolic pathways. The metabolic pathways of both humans and the pathogen were carefully analyzed to identify pathways unique to the pathogen. KEGG automatic annotation server (KAAS) [25] assigned KEGG Orthology (KO) identifiers to the proteins categorizing the KEGG independent and dependent proteins [24]. The proteins engaged in the pathogen-particular pathways were identified and subjected to further analysis.

Subcellular localization of proteins

All proteins require a specific localization to function optimally [26]. To comprehend the protein functions in the cells, it is essential to understand their subcellular locations. Subcellular localization of both KEGG-independent and dependent proteins was assessed by employing tools such as DeepLoc [27], BUSCA [28], and CELLO [29]. Identifying subcellular localization is crucial for identifying possible vaccine and drug targets.

Selection of vaccine candidates

The shortlisted vaccine targets were assessed for their potential as vaccine candidates using a variety of criteria. VaxiJen 2.0 [30] was utilized to analyze antigenicity, and the allergenic profile was assessed using the Allertop server [31]. Different servers such as ProtparamExPasy, TMHMM [32], and SolPro [33] servers were used to evaluate the physicochemical characteristics, topology, and water solubility of the vaccine targets, respectively. Vaccine targets were chosen for their high antigenicity and non-allergenic properties.

T-cell-binding epitope prediction

The regions of antigens that immune cells recognize are referred to as epitopes. T lymphocytes recognize MHC molecules displayed by antigen-presenting cells (APCs) as epitope peptides [34]. Hence, T-cell epitope prediction is a critical step in vaccine design, owing to their central role in coordinating the adaptive immune response. The MHC class I molecules generally present the antigenic peptides to cytotoxic T lymphocytes (CTLs) with the help of CD8 receptors, while the MHC class II molecules via the CD4 receptors interact with helper T lymphocytes (HTLs) [35]. For the identification of MHC binding epitopes, the Immune Epitope Database (IEDB) tool was accessed, followed by ranking the epitopes based on their percentile scores. The lowest ranked 10 epitopes were selected and evaluated for their antigenicity and toxicity using the VaxiJen v2.0 server [30] and ToxinPred server [36], respectively. The IEDB immunogenicity tool was

accessed to specifically predict the immunogenic potential of the MHC class I molecules. The MHC class II epitopes were examined for additional cytokine-inducing properties, including interleukin-4 (IL-4), interleukin-10 (IL-10), and interferon-gamma (IFN- γ) as they are essential for modulating the immune responses and triggering B-cell activation. Based on the comprehensive evaluation of all of these parameters, the resulting epitopes were classified as potent and versatile T-cell epitopes.

Antigenic B-cell epitope prediction

B lymphocytes play a crucial role in the immune system by secreting antibodies that contribute to long-term immunity. They recognize B-cell epitopes, which are essential for antibody-mediated immune responses. Linear B-cell epitope prediction was performed using the ABCpred server [37]. Predicted epitopes were then assessed through Allertop 2.0, VaxiJen, and Toxinpred servers to determine their allergenicity, antigenicity, and toxicity with a preferred length of 10 to 40 mer.

Chimeric vaccine construct

The lead CTL, HTL, and B-cell-binding epitopes were used for developing a multi-epitope vaccine construct and combined to form a fusion peptide using linkers. AAY linkers were used for CTL epitopes facilitating proteasomal cleavage and TAP binding to support efficient presentation of MHC-I. GPGPG linkers were utilized for linking the HTL epitopes together due to their role in the reduction of junctional immunogenicity and enhanced stimulation of T_H cells [38, 39]. Moreover, KK linkers were employed for joining the B-cell epitopes together. These KK linkers are known as cleavage sites for cathepsin B, which is a lysosomal protease involved in MHC-II antigen processing. This cleavage enhances the individual B-cell epitopes' accessibility to B-cell receptors and improves their presentation to T_H cells, helping in effective generation of antibodies [40].

The final vaccine formulation process involved the addition of six various adjuvants, i.e., beta-defensin-3 (Uniprot ID: Q5U7J2), L7/L12 ribosomal protein (Uniprot ID: A0A246AKK7), flagellin (Uniprot ID: P06179), HBHA (NCBI ID: A1KFU9), HBHA-conserved, and granulocyte macrophage colony (Uniprot ID: P04141), followed by T-cell and B-cell epitopes joined via different linkers to design six distinct vaccine constructs. The HBHA and HBHA conserved adjuvants differ in length, the full-length HBHA protein and a shortened version containing the conserved immunostimulatory domain only. The conserved version is less variable and more broadly immunogenic, making it a potential candidate for vaccine targeting [41, 42]. Another EAAAK linker was incorporated into the C-terminal of the vaccine and

at the N-terminal to attach the adjuvant to the construct [43].

Structural evaluation of vaccine construct

The ProtParam server was utilized to assess the physicochemical characteristics of the vaccine construct, which included the molecular weight (MW), Grand Average of Hydropathicity (GRAVY), alphabetic index (AI), theoretical isoelectric point (theoretical PI), instability index (II) [44]. Furthermore, the antigenicity and the allergenicity were predicted by employing the VaxiJen v2.0 [30], AntigenPro, and the AllerTop server. The SOLpro was used for predicting solubility of the vaccine upon expression in *E. coli* [33].

Determination of post-translational modifications

Post-translational modifications (PTMs) significantly enhance proteome diversity and play crucial roles in protein stability and function [45]. The vaccine constructs were acetylated and phosphorylated using NetAct 1.0 and NetPhos 3.0 servers [46].

Structure prediction

Prediction of secondary structure (random coiling, alpha helix, and beta-sheet) of vaccine constructs was done using SOPMA server, which is an accurate technique for 2D structures, using two feed-forward neural networks to assess PSI-BLAST results [47]. The GlobPlot 2.3 server was used to identify globular and disordered regions. The trRosetta server with residual deep neural networks was utilized to model the 3D structures of the vaccine constructs [48].

Refinement and validation of the tertiary structure

Tertiary structures predicted by trRosetta server were improved in quality by refinement. GalaxyRefine server refines protein models and improves local structure quality with a probability of over 50% [49]. The refined tertiary structure of the construct was validated using the ERRAT score, which verifies protein structures by representing non-bonded interactions. This evaluation is important for tracking the structural refinement and accuracy of crystallographic models. The optimal ERRAT quality factor value is $\geq 0.5\%$, with higher scores indicating better protein quality. Further quality validation of the vaccine construct was evaluated by the Ramachandran plot using PROCHECK [50]. The Ramachandran plot validates protein structural models, whether generated computationally or experimentally. The number of residues in the favored regions should be more than 90% for optimal protein structure [21].

Molecular Docking of chimeric vaccine with Toll-like receptors

An effective immune response relies on the interaction between an antigen and a specific immune receptor. Molecular docking is a technique used to predict the binding affinity between a ligand and its receptor. The ClusPro 2.0 server was used to perform docking studies to confirm the interaction of vaccines with the human immune receptors TLR4 (PDB ID: 48GA) and TLR5 (PDB ID: 3J0A). The structures were uploaded in PDB format to the ClusPro 2.0 server, a protein-protein docking tool that uses the Fourier correlation algorithm to clarify models using electrostatic and desolvation energies [51]. Docked complexes were identified based on their binding capacity, and their interactions were investigated using the PyMOL program. The complexes with the highest binding affinity and several interactions were subjected to normal mode analysis (NMA) and molecular dynamic simulations to check their stability [52].

Measurement of bond distances

PyMOL visualizes and measures the angles, lengths, and dihedrals between atoms, residues, or molecular structures. PyMOL was used to assess the bond distances between the receptor and vaccine. The determination of bond lengths between receptors and vaccines is crucial for understanding interactions.

NMA evaluation of the vaccine-receptor complex

The top interaction complex, identified based on its significant docking score, was further analyzed using dynamic modeling approaches that simulate molecular behavior in a natural environment. The iMODs server was used to assess the prioritized vaccine designs and their interactions with TLR4 and TLR5. This tool predicts the range and orientation of protein-protein complex basic movements based on variance, elastic network model, B-factor, covariance, eigenvalues, and deformability. The server was utilized for normal mode analysis (NMA) in nucleic acid structures and structural dynamics of proteins [53].

Conformational B-cell epitopes

The folding of protein can produce conformational B-cell epitopes by arranging residues in close contact. Over 90% of B-cell epitopes are considered to be conformational [54]. The vaccines prioritized in the docking step were assessed for the presence of these epitopes using the ElliPro server. This server uses Thornton's technique and clustering algorithms to compute the 3D structure and the protrusion index (PI) value of each predicted epitope [54].

MD simulations of vaccine-receptor complexes

MD simulations were performed for vaccine-receptor complexes to analyze their dynamic behavior under near-physiological conditions. The simulations were conducted using Desmond (Schrödinger Suite 2023-4) on a Linux workstation equipped with an NVIDIA RTX 4090 GPU (128 multiprocessors, 24 GB RAM). Each simulation spanned 200 nanoseconds with a time step of 2 femtoseconds [55]. Before starting the MD simulations, the vaccine was subjected to docking with the receptors, providing an initial static prediction of binding orientations. The docking analysis was critical for putting the vaccine molecules in the receptor's active regions, which provided a foundation for dynamic simulation. During the molecular dynamics simulations, Newton's classical equations of motion [56] were integrated to represent an individual atom's movements throughout time. The system setup and minimization were performed using Schrödinger's Maestro [57], including force-field-based optimization (OPLS5). The complexes were solvated in an orthorhombic box using the TIP3P (transferable intermolecular potential with 3 points) water model, and physiological conditions were simulated by adding counter ions (Na^+ and Cl^-) at a concentration of 0.15 M to neutralize the system [58]. The simulation utilized the NPT ensemble with 1 atm pressure (MTK barostat, $\tau = 2.0$ ps) and 300 K temperature (Nosé-Hoover thermostat, $\tau = 1.0$ ps). Electrostatics were handled using the u-series method with a 9 Å cutoff. An eight-stage relaxation protocol was implemented via the Multisim module. Stage 1 involved system recognition and validation, followed by Stages 2–6 comprising Brownian dynamics, NVT, and NPT simulations at 10 K with positional restraints to gradually equilibrate the system. The production run (Stage 7) was conducted under an NPT ensemble at 300 K for 200 nanoseconds (100,000 steps) with a 2-femtosecond time step. Stage 8 focused on post-simulation analysis using Maestro's Simulation Interaction Diagram module. Trajectory data were recorded every 200 ps, and snapshots were generated every 120 ps for detailed analysis. The initial 12 ps were considered equilibration, after which the vaccine-receptor complexes' stability and flexibility were evaluated using the root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) evaluations, which provided detailed insights into their structural dynamics throughout simulations.

PCA analysis

Principal Component Analysis (PCA) was conducted to analyze MD trajectories of each vaccine-receptor complex using the bio3d R package [59]. This statistical method uses the covariance matrix of alpha-carbon ($\text{C}\alpha$) atoms to detect large protein conformational changes. The primary components, which indicate the system's

dominant modes of motion, were chosen from the eigenvectors with the biggest eigenvalues. By focusing on these principal components, the audit offered useful information about the dynamic behavior of the protein complexes during the simulations.

DCCM analysis

Dynamic cross-correlated matrix (DCCM) analysis was carried out on 100 ns MD simulation trajectories to check the correlated movements into the V1-TLR4 and V2-TLR5 docked complexes. Simulation data was used to determine the cross-correlation coefficients between the fluctuations of each residue pair. The generated dynamic cross-correlated was as a heatmap, with values that range from -1.0 (anti-correlated motion) to 1.0 (positively correlated motion). This audit enabled the recognition of regions within the complexes that demonstrated significant anti-correlated or correlated motions, providing intuition into the potential functional linkages and dynamic linkage between the vaccine-receptor complexes.

Analysis of population coverage

T-cell epitopes and their associated alleles were examined using the Population Coverage tool from the IEDB database, ensuring that the selected epitopes provide coverage for the majority of the world's population. This server calculates the average coverage of epitopes across populations using the distribution of MHC-binding alleles. Epitopes were analyzed globally due to the worldwide impact of *A. duodenale*.

Codon optimization and in-silico expression

The vaccine's codon usage was modified employing the JCAT (Java Codon Adaptation Tool) based on the host strain *E. coli* K12 [60]. The amino acid sequence of vaccine was back-translated into DNA and then altered for codon usage in *E. coli*. The modification was based on Codon Adaptation Index (CAI) values calculated using an algorithm. Codon optimization maximizes efficiency and yields while sustaining the correct sequence of amino acids for vaccine protein production in the host organism. The optimized codon sequence was then cloned with the SnapGene program [61]. The designed vaccine gene was inserted into a pET-28a (+) plasmid vector for host expression due to its broad use in expression studies [62].

Immune simulations

To examine the immune reaction of our best-docked vaccine-receptor complex, C-ImmSim server was utilized [63]. Three injections were injected at intervals of 1, 84, and 168 h, with parameters kept default. This tool evaluates the levels of helper T-cells ($\text{T}_{\text{H}1}$ & $\text{T}_{\text{H}2}$), as well

as the induction of cytokines, antibodies, and interferons against the constructed vaccine. The immune response following vaccine administration was also predicted.

Results

Essential and host-interacting proteins

The Zhejiang strain of *Ancylostoma duodenale* containing 27,484 proteins was selected for in this study due to its genetic stability and importance in drug resistance and vaccine development studies. Pathogen survival is dependent on essential genes that are involved in structural organization, nutrition uptake, and pathogenicity. These genes allow for adaptability, resource acquisition, and infection establishment in host settings [64]. The results of the DEG blast result identified 524 essential proteins in *A. duodenale*. Infectious illnesses are facilitated by HP-PPI (host-pathogen protein-protein interactions), which occur through molecular cross-talk between pathogens and hosts. A total of 1392 host-pathogen-interacting proteins were predicted in *Ancylostoma duodenale*. Thus, 1761 DEG and HPPI combined proteins were filtered for the next analysis.

Identification of non-homologous protein

Filtered proteins were analyzed using standalone BLASTp against the human proteome to eliminate those identical to host proteins, thereby reducing the risk of potential deleterious effects. A total of 108 non-homologous pathogen proteins were retrieved. Microbes in the human body play important roles in metabolic processes. The 108 proteins were subjected to screening against human gut microbiota, resulting in 36 non-homologous proteins.

Metabolic pathway prediction

Metabolic pathways for *A. duodenale* (138 pathways) and the human host (356 pathways) were retrieved from the KEGG server and compared for common and unique pathways. The comparison found 136 pathways common to humans and *A. duodenale*, with two pathways unique to the parasite. The KAAS server (KO assignment) was used to perform functional annotation on 36 important pathogen proteins identified using comparative subtractive genomics for metabolic pathway enrichment. Ten

of these proteins were KEGG-independent, while the remaining 26 were engaged in the unique pathways of pathogen.

Subcellular localization of proteins

Subcellular localization is a key factor in determining the biological functions of many proteins. The accessibility of these proteins to therapeutics largely depends on their subcellular location. The subcellular localization data indicated that 28 KEGG-dependent and 8 KEGG-independent were either cytoplasmic, extracellular, nuclear, membrane, or mitochondrial proteins. Six membrane proteins were predicted as vaccine proteins (Table 1). The extracellular and membrane proteins were selected as potential vaccine candidates for chimera vaccine construction because surface-secreted and secreted exposed proteins are thought to be effective vaccine candidates since immune cells may quickly recognize them.

Vaccine target screening

The six membrane proteins were evaluated for the parameters including the GRAVY value, instability index, aliphatic index, topology, antigenicity, allergenicity, water solubility, number of amino acids, molecular weight, theoretical Pi, virulence, and half-life. Topology values of 0 or 1 were favored, with the GRAVY value negative. The vaccine target should also be stable. Among the six proteins, two (KIH49483.1 & KIH58423.1) indicated the potential for vaccine construction (Table 2).

T-cell epitopes prediction

T-cell-binding epitopes were determined by using the IEDB database. The top 10 non-overlapping MHC-I and MHC-II binding epitopes, each with the lowest ranks from both proteins, were selected. Even though the anticipated T-cell epitopes had a high capacity for MHC-I and MHC-II molecules, this does not necessarily suggest that the epitopes may elicit an immunological response; thus, several characteristics were utilized to investigate the predicted epitopes. The toxicity (as determined by Toxinpred) and antigenicity (assessed by VaxiJen 2.0) were also estimated, with non-toxic features for Toxinpred and a threshold of >0.4 for VaxiJen. The investigation utilized both the IEDB epitope immunogenicity tool for MHC-I and other cytokine induction techniques for MHC-II epitopes [65]. As a result, one MHC-I and four MHC-II binding epitopes were selected from KIH49483.1 (Table S1 & S2), while two MHC-I and three MHC-II binding epitopes were finalized from the KIH58423.1 protein (Table S3 & S4).

Prioritization of B-cell epitopes for chimeric vaccine

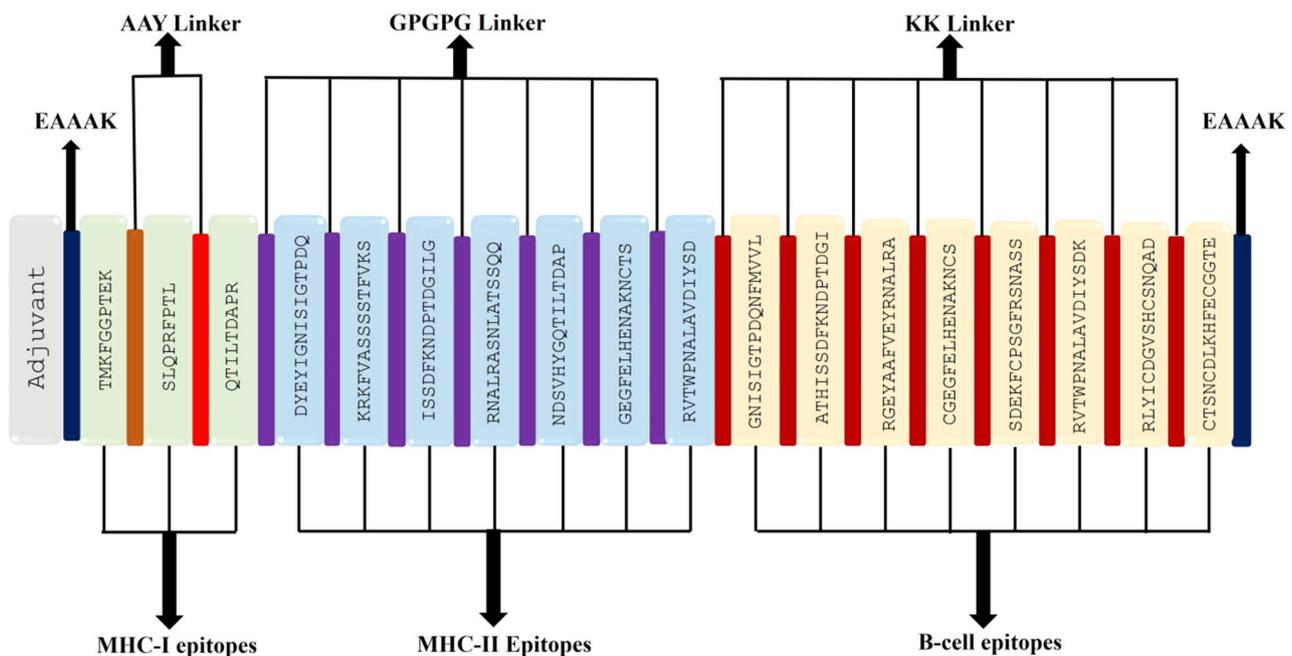
Linear B-cell epitopes were predicted using two vaccine target proteins. ABCpred server determined a total

Table 1 Subcellular localization of proteins predicted as vaccine candidates

Sr. No.	Accession numbers	Localization
1	KIH49483.1	Extracellular, Plasma membrane
2	KIH57671.1	Extracellular, Plasma membrane
3	KIH58423.1	Extracellular
4	KIH59417.1	Extracellular, Mitochondrion
5	KIH68051.1	Extracellular
6	KIH68881.1	Plasma Membrane

Table 2 Physicochemical characteristics of potential vaccine targets

Parameters	KIH49483.1	KIH57671.1	KIH58423.1	KIH59417.1	KIH68051.1	KIH68881.1
No. of AA	220	840	272	209	162	1110
Topology Value	0	0498-517i	0	0	0	0
Molecular Weight (Da)	24125.65	94484.84	29827.58	23891.56	17724.97	123053.36
Theoretical Pi	8.93	7.1	5.04	5.9	4.6	9.27
Aliphatic Index	86	79.94	76.1	81.15	55.93	94.43
Gravy	-0.056	-0.306	-0.232	-0.524	-0.386	-0.177
Antigenicity	Antigen	Antigen	Antigen	Non-Antigen	Antigen	Non-Antigen
Allergenicity	Non-Allergen	Non-Allergen	Non-Allergen	Non-Allergen	Allergen	Non-Allergen
Instability Index	Stable	Unstable	Stable	Stable	Stable	Unstable
Water Solubility	Poor	Good	Good	Poor	Good	Good
Half Life	30 h	5.5 h	30 h	4.4 h	30 h	30 h

**Fig. 2** Vaccine design built with LBL, CTL, and HTL epitopes and adjuvants coupled via different linkers

of 3 epitopes for KIH49483.1 and 5 B-cell epitopes for KIH58423.1. Of these, Vaxijen predicted 6 epitopes of KIH49483.1 and 9 of KIH58423.1 as antigenic peptides (Table S5). Linear epitopes were determined by combining biological factors such as surface accessibility, hydrophilicity, hydrophobicity, and amino acid composition using the IEDB server.

Multi-epitope vaccine designing

Six vaccine constructs were created by combining selected epitopes. The vaccine constructs were named V1 to V6 after being connected with different adjuvant sequences: beta-defensin, L7/L12 ribosomal protein, flagellin, HBHA, HBHA conserved, and granulocyte macrophage colony, respectively. Linkers are often employed to enhance antigen presentation activity in multi-epitope

vaccines. All six vaccine constructs were connected using AAY, EAAAK, GPGPG, and KK linkers (Fig. 2).

Chimeric vaccine construct analysis

Designed vaccine constructs (V1-V6) were evaluated for chemical and physical properties. The aliphatic index indicates the amount of space occupied by aliphatic chains, whereas the instability index (stability score < 40.0) evaluates the stability of the vaccine structure. GRAVY is an important element that is suitable with just negative ratings. Its calculation involves determining the hydrophobicity of amino acids using the mean formula. To facilitate in vitro vaccine development, final vaccine constructs must have a molecular weight of 41–59 kDa and amino acid content of 300–800 residues. ANTIGENpro (scoring > 0.90) and Vaxijen v2.0 (score > 0.75) evaluated vaccine constructs antigenicity.

Allergenicity was assessed employing the AllergenFP server, which identifies possibly allergenic zones in vaccines. The SOLpro server with a threshold of >0.5 was used to determine vaccine construct solubility, which is crucial for vaccine manufacturing (Table 3).

Prediction of post-translational modifications

The vaccine construct was evaluated for two types of post-translational modifications. Phosphorylation modification by the NetPhos-3.0 server predicted 43 (Threonine: 14, Serine: 23, Tyrosine: 6) and 58 phosphorylation sites (Threonine: 19, Serine: 33, Tyrosine: 6) within the constructs V2 and V3, respectively (Figure S1). Phosphorylated epitopes have a high affinity for cytotoxic T lymphocytes and serves an essential role in promoting particular and strong immune responses. Acetylation is a change that can dynamically alter the function of a protein by changing its characteristics, including hydrophobicity, surface properties, and solubility which all may impact protein conformation and interactions with substrates, cofactors, and other macromolecules [66]. NetAcet-1.0 server predicted a 0.514 acetylation score for construct V2 and a 0.453 acetylation score for construct V3.

2D structure prediction

The SOPMA server was utilized to predict the 2D structures of the vaccine constructs (V1 to V6). The results showed that the V1 vaccine construct contained 16.35% alpha-helices, 6.97% beta-strands, 23.06% extended strands, and 53.62% random coils. V2 construct contained 30.29% alpha-helix, 7.57% beta-strands, 18.26% extended strands, and 43.88% random coils. V3 construct contains 38.49% alpha-helix, 6.12% beta-strands, 16.19% extended strands, and 39.21% random coils. V4 construct contains 39.84% alpha-helix, 4.52% beta strands, 14.78% extended strands, and 40.86% random coils. V5 construct contains 40.59% alpha helix, 4.60% beta strands, 15.06% extended strands, and 39.75% random coils. V6 construct contains 29.72% alpha helix, 5.10% beta strands, 16.99% extended strands, and 48.20% random coils.

Globular and disordered region prediction

The GlobPlot2 server predicted two globular regions in the V2 vaccine construct: 3-155 and 290-371, and eight (8) disordered regions: 1-5, 156-190, 194-230, 235-289, 302-309, 326-336, 372-389, and 422-428 residues. The first globular region spans the adjuvant region and is located at start of the MHC II binding region, while the second globular region lies into the B-cell epitopes. The predicted disordered region starts at the end of the MHC-I epitope and ends within the B-cell epitopes. Moreover, three globular regions were identified in the V3 vaccine construct: 1-175, 310-391, and 410-556, and eight (8) discarded regions were identified: 176-210, 214-250, 255-309, 322-329, 346-356, 392-409, 442-448, and 547-552. The first globular region spans the whole adjuvant and covers almost all epitopes of MHC-I. Meanwhile, the second globular region starts from the last epitope of MHC-II and extends to the final EAAAK linkers, also covering the B-cell epitopes.

3D prediction, refinement, and validation

The 3D structures of the vaccine constructs were generated using the trRosetta webserver (Fig. 3). GalaxyRefine server improved and refined the functionality of vaccine 3D models. All models were inspected and validated for quality using ERRAT values and Ramachandran plots in the SAVES server v6.0. The ERRAT scores of 87.83, 96.15, 95.87, 91.18, 85, and 93.80 for V1, V2, V3, V4, V5, and V6, respectively, reflect minimal errors in protein folding and structure, indicating high-quality vaccine models. Higher ERRAT scores correspond to improved structural reliability in vaccine design. Ramachandran plot results showed 93.3%, 95.1%, 95.2%, 91.7%, 91.8%, and 94.8% residues in the favored region of the vaccines (Fig. 4). A standard model is characterized by having >90% of residues in favored regions, indicating that all vaccine models are suitable for docking and molecular dynamics simulation studies.

Molecular Docking of the designed vaccine with Toll-Like receptors

Using the ClusPro 2.0 webserver, the interactions among the vaccines' 3D models and immune receptors were

Table 3 Summary of the physicochemical characteristics of vaccine constructs

Vaccine construct design	Solpro	Antigenpro	Vaxijen 2.0	Allergenicity	Molecular weight	Theoretical PI	Gravy value	Ali-phatic index	Instability index	Amino Acids	Topology
V1	0.75	0.92	Antigen	Non-Antigen	39829.82	9.26	-0.705	54.75	Stable	373	0
V2	0.78	0.94	Antigen	Non-Antigen	47160.99	6.63	-0.51	67.48	Stable	449	0
V3	0.72	0.95	Antigen	Non-Antigen	59087.6	8.68	-0.575	71.17	Stable	556	0
V4	0.90	0.93	Antigen	Non-Antigen	52297.27	6.19	-0.637	65.46	Stable	487	0
V5	0.93	0.938	Antigen	Non-Antigen	51179.05	5.89	-0.619	67.28	Stable	478	0
V6	0.69	0.92	Antigen	Non-Antigen	50850.23	7.85	-0.565	62.42	Stable	471	0

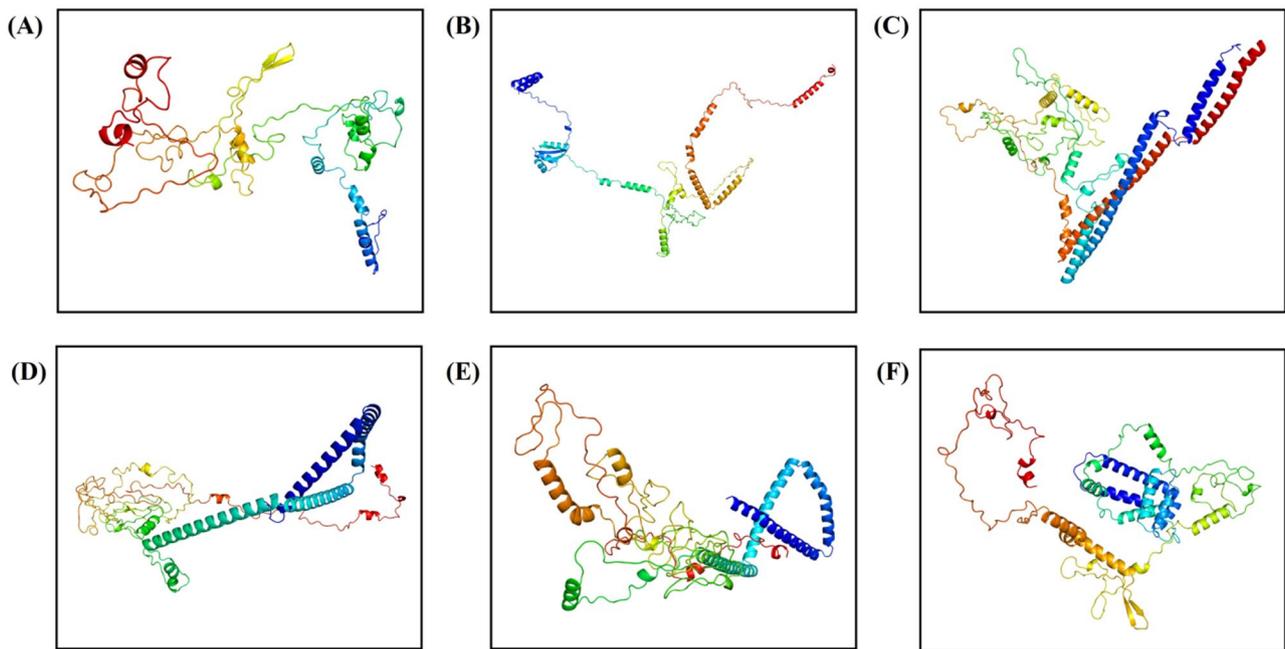


Fig. 3 3D structure of the vaccine constructs. (A) V1 (B) V2 (C) V3 (D) V4 (E) V5 (F) V6

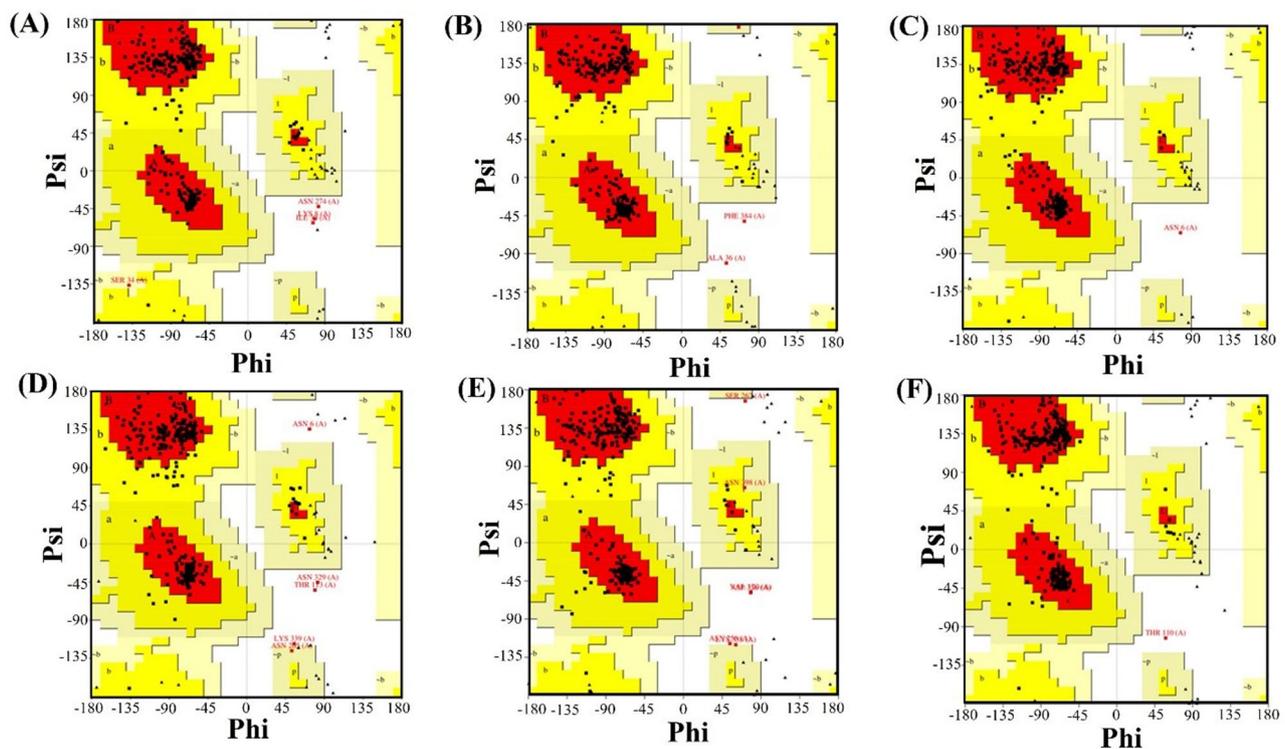


Fig. 4 Ramachandran plot of vaccines formed by the SAVES server (A) V1 (B) V2 (C) V3, (D) V4 (E) V5, and (F) V6

predicted. The vaccines were docked to human immune receptors, specifically TLR-4 (PDB ID: 48GA) and TLR5 (PDB ID: 3J0A). The top 10 complexes were produced by the server for every vaccine, along with their binding energies. One model was chosen for each vaccine

based on the lowest binding energy, as lower binding energy scores reflect a vaccine’s stronger binding affinity to immune receptors. The docking scores of the V1, V2, V4, V5, and V6 were predicted to be -1362.3, -1315.2, -1338.3, -1090.1, and -1330.5 respectively, with the

receptor TLR4, while with TLR5, the docking score of V3 was found to be 1789.4, respectively. The analysis of interactions between vaccine constructs (V1-6) and TLR 4 & 5 provided critical insights into immune response regulation and implications for vaccine design. Strong interactions were identified between vaccine models and TLR4 and TLR5 receptors. The interactions of V1, V2, V3, V4, V5, and V6 were determined to be 19, 35, 39, 20, 13, and 19 respectively. The docking scores and interactions of the vaccine construct in complex with TLR-4 and TLR-5 revealed that constructs V2 and V3 with the lowest binding energy and highest levels of receptor interactions. Therefore, the V2-TLR4 (Fig. 5) and V3-TLR5 (Fig. 6) docked complexes were further carried out to molecular dynamics (MD) simulation analysis to assess their stability and binding affinity.

Measurement of bond distances

Bond distance values in Å between vaccines (V2 and V3) and receptors (TLR4 and TLR5) were measured using

PyMOL. All the bond distances of interacting residues were less than 3.4 indicating stronger bonds among them. The interacting residues of the final L7/L12 ribosomal protein adjuvant vaccine with receptor TLR4 showed the interacting residues LYS-136 & THR-499, LYS-136 & GLU-474, THR-134 & GLN-547, THR-134 & GLN-547, ASN-114 & GLN-240, SER-237 & ASN-49, ASN-49 & ASN-237, ASN-49 & ASN-233, ARG-68 & TYR-167, ASN-212 & SER-27, ASN-26 & ASN-212, ARG-226 & GLU-53, ARG-226 & GLU-53, ARG-226 & GLU-53, ARG-226 & SER-28, GLY-263 & ASP-50, GLY-265 & ASP-50, LYS-47 & GLY-266, LYS-47 & GLY-266, LYS-47 & THR-257, LYS-47 & THR-257, GLN-91 & SER-248, GLN-91 & ASP-247, ARG-67 & GLN-253, ARG-67 & HIS-250, TYR-251 & THR-92, THR-92 & TYR-251, LYS-39 & THR-156, LYS-39 & ASP-157, TYR-42 & ARG-160, ARG-160 & GLU-144, ARG-160 & MET-147, TYR-151 & GLU-144, GLN-152 & GLU-136 had distance value 2.7, 2.0, 1.8, 2.0, 2.2, 1.8, 2.6, 2.3, 1.8, 2.0, 2.0, 2.0, 1.9, 2.2, 2.5, 2.1, 2.5, 1.6, 3.5, 1.6, 1.7, 2.0, 2.7, 2.0, 1.8, 1.8, 2.6, 2.2, 1.8,

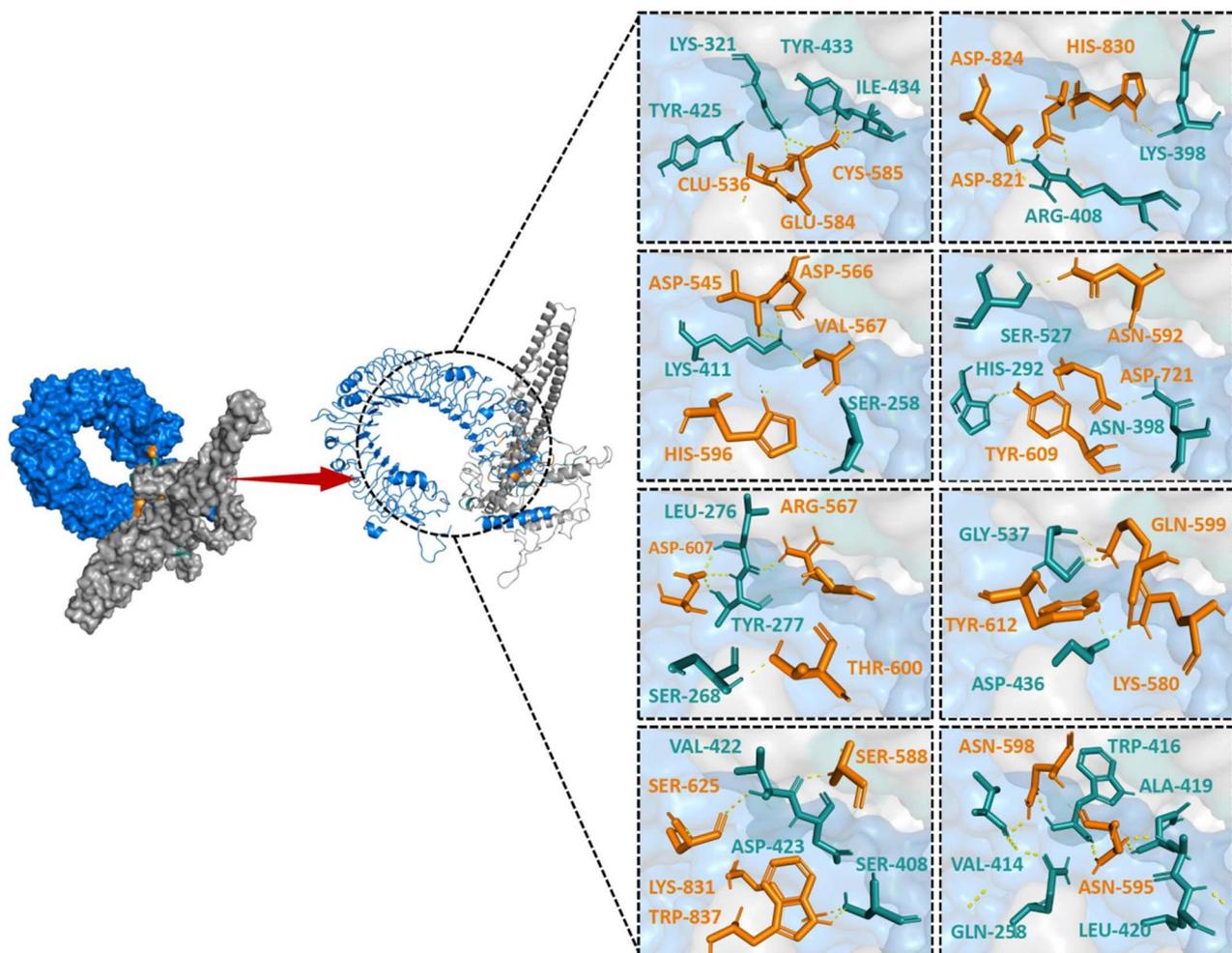


Fig. 5 A cartoon illustration of the docked complex V1-TLR4 between the designed vaccine construct V2 (grey) and TLR-4 (blue), including their interacting residues of TLR4 (orange) and vaccine residues (cyan)

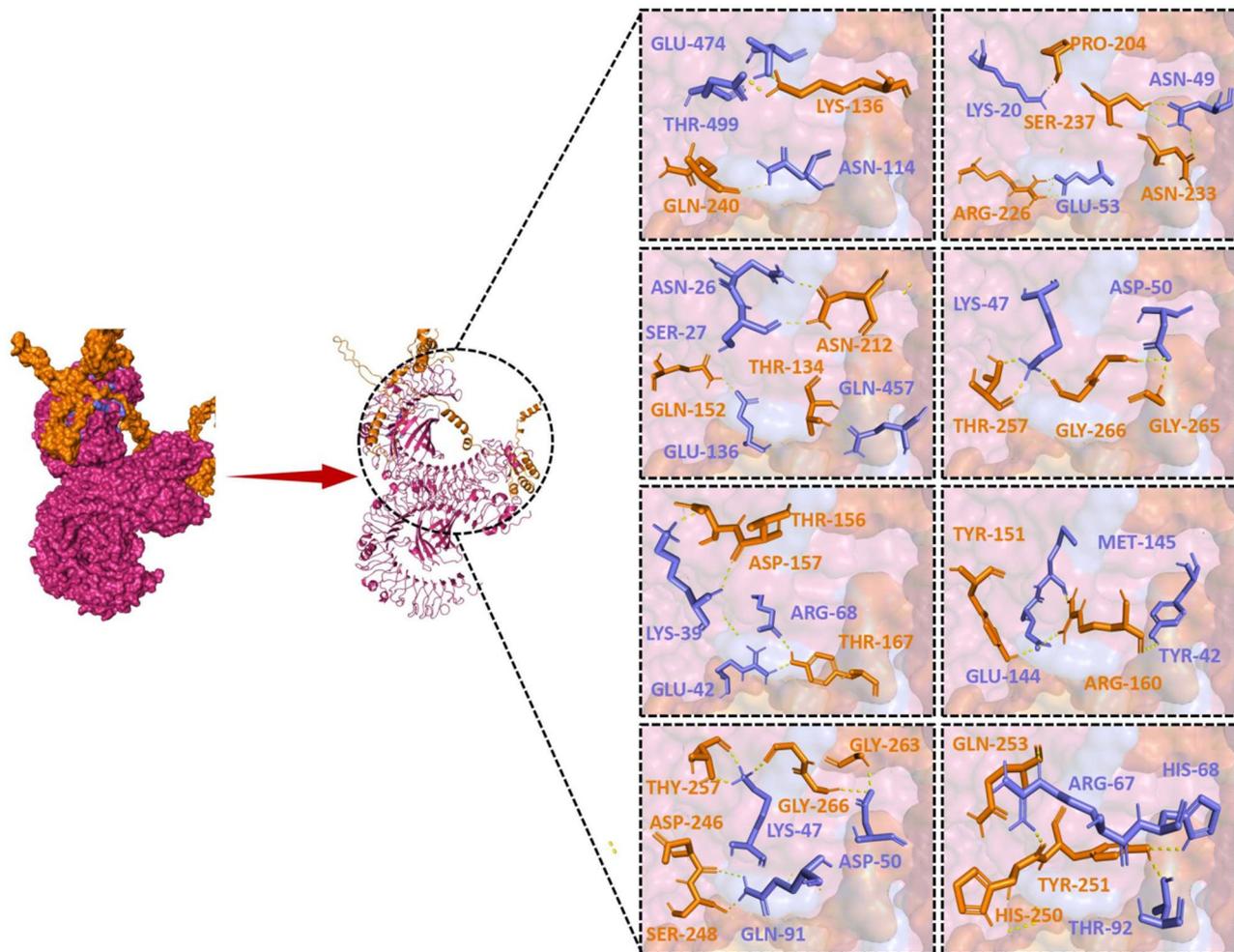


Fig. 6 A cartoon illustration of the docked complex V2-TLR5 between the designed vaccine construct V3 (orange) and TLR-5 (pink), including their interacting residues (lavender)

2.4, 2.3, 1.7, 2.0, 2.2 while flagellin adjuvant vaccine with receptor TLR5 showed the interacting residues ASP-7721 & ASN-389, HIS-292 & TYP-609, ASP-821 & ASP-824, LYS-398 & HIS-830, SER-402 & TRP-827, SER-402 & LYS-831, LYS-411 & ASP-545, VAL-567 & LYS-411, ASP-566 & LYS-411, LYS-411 & ASP-345, GLN-273 & ALA-602, SER-625 & VAL-422, SER-588 & 422, LYS-231 & GLU-584, LYS- 231 & GLU-584, TYR-425 & CYS-585, ALA-425 & GLU-586, ASN-592 & SER- 257, SER- 258 & HIS-596, and ALA-419 & ASN-595 had distance value 1.9, 2.5, 1.8, 2.0, 2.2, 1.7, 1.7, 1.7, 2.8, 2.0, 2.5, 1.8, 2.6, 1.8, 2.4, 2.1, 1.9, 3.4, 2.0, and 1.9.

Conformational B-cell epitopes

The conformational B-cell epitopes are surface-exposed and easily exposed to the solvent. These epitopes are typically produced through protein folding and are readily available to antibodies and B-cells, helping to stimulate a robust immune response. Using the ElliPro server with a score greater than 0.7, a total of 13, 26, 46, 12, 30, and 5

residues were predicted within the conformational B-cell epitopes of V2, while 5, 6, 7, 84, 76, and 66 residues were identified within the conformational B-cell epitopes of V3 respectively (Figure S2 and S3) (Tables 4 and 5). It is important to note that the 3D structures of the designed vaccines are different from the native proteins (KIH49483 and KIH58423). The goal of predicting conformational B-cell epitopes was to ensure that the designed multi-epitope constructs could form immunogenic and surface-accessible regions capable of being recognized by B-cell receptors. Therefore, the conformational B-cell epitopes identified in this study represent novel surface-accessible regions on the vaccine constructs rather than mimicking those of the native proteins. This strategy ensures that the engineered construct might be capable of simulating a robust humoral immune response.

NMA evaluation

The binding affinities of the top-docked V2 and V3 complexes with the immune receptors TLR4 and TLR5 were

Table 4 Conformational B-cell epitopes analysis of V2

No.	Residues	No. of residues	Score
1	A: H437, A: F438, A: E439, A: C440, A: G441, A: G442, A: T443, A: E444, A: E445, A: A446, A: A447, A: A448, A: K449	13	0.984
2	A: R411, A: L412, A: Y413, A: I414, A: C415, A: D416, A: G417, A: V418, A: S419, A: H420, A: C421, A: S422, A: N423, A: Q424, A: A425, A: D426, A: K427, A: K428, A: C429, A: T430, A: S431, A: N432, A: C433, A: D434, A: L435, A: K436	26	0.9
3	A: M1, A: S2, A: D3, A: L4, A: K5, A: N6, A: L7, A: A8, A: E9, A: T10, A: L11, A: V12, A: N13, A: L14, A: T15, A: V16, A: K17, A: D18, A: V19, A: N20, A: E21, A: L22, A: A23, A: A24, A: I25, A: L26, A: K27, A: D28, A: E29, A: Y30, A: G31, A: I32, A: E33, A: P34, A: A35, A: A36, A: A37, A: A38, A: V39, A: V40, A: M41, A: A42, A: G43, A: P44, A: G45, A: A46	46	0.84
4	A: V394, A: T395, A: W396, A: P397, A: N398, A: A399, A: L400, A: A401, A: V402, A: D403, A: I404, A: Y405	12	0.758
5	A: K211, A: N212, A: D213, A: P214, A: T215, A: D216, A: G217, A: I218, A: L219, A: G220, A: G221, A: P222, A: G223, A: P224, A: G225, A: R226, A: N227, A: A228, A: L229, A: R230, A: A231, A: S232, A: N233, A: L234, A: A235, A: T236, A: S237, A: S238, A: Q239, A: Q240	30	0.716
6	A: S406, A: D407, A: K408, A: K409, A: K410	5	0.707

Table 5 Conformational B-cell epitopes analysis of V3

No.	Residues	No. of residues	Score
1	A: C400, A: P401, A: S402, A: G403, A: F404	5	0.922
2	A: K394, A: S395, A: D396, A: E397, A: K398, A: F399	6	0.844
3	A: R405, A: S406, A: N407, A: A408, A: S409, A: S410, A: K411	7	0.827
4	A: M1, A: A2, A: Q3, A: V4, A: I5, A: N6, A: T7, A: N8, A: S9, A: L10, A: S11, A: L12, A: L13, A: T14, A: Q15, A: N16, A: N17, A: L18, A: N19, A: K20, A: S21, A: Q22, A: S23, A: A24, A: L25, A: G26, A: T27, A: A28, A: I29, A: E30, A: R31, A: L32, A: S33, A: S34, A: G35, A: L36, A: R37, A: I38, A: N39, A: S40, A: A41, A: K42, A: D43, A: D44, A: D517, A: S518, A: D519, A: Y520, A: A521, A: T522, A: E523, A: V524, A: S525, A: N526, A: M527, A: S528, A: R529, A: A530, A: Q531, A: I532, A: L533, A: Q534, A: Q535, A: A536, A: G537, A: T538, A: S539, A: V540, A: L541, A: A542, A: Q543, A: A544, A: N545, A: Q546, A: V547, A: P548, A: Q549, A: N550, A: V551, A: L552, A: S553, A: L554, A: L555, A: R556	84	0.748
5	A: G80, A: A81, A: N83, A: E84, A: I85, A: N86, A: N87, A: N88, A: L89, A: Q90, A: R91, A: V92, A: R93, A: E94, A: L95, A: A96, A: V97, A: Q98, A: S99, A: A100, A: N101, A: S102, A: T103, A: N104, A: S105, A: Q106, A: S107, A: D108, A: L109, A: D110, A: S111, A: I112, A: Q113, A: A114, A: E115, A: I116, A: Q117, A: Q118, A: R119, A: L120, A: N121, A: E122, A: I123, A: D124, A: R125, A: V126, A: S127, A: G128, A: Q129, A: T130, A: K136, A: L138, A: C453, A: D454, A: L455, A: K456, A: H457, A: F458, A: E459, A: C460, A: G461, A: G462, A: T463, A: E464, A: E465, A: A466, A: A467, A: A468, A: K469, A: L470, A: Q471, A: K472, A: I473, A: D474, A: A476, A: L477	76	0.625
6	A: N232, A: D233, A: P234, A: T235, A: D236, A: G237, A: I238, A: L239, A: G240, A: G241, A: P242, A: G243, A: P244, A: G245, A: R246, A: N247, A: A248, A: L249, A: A251, A: S252, A: N253, A: L254, A: A255, A: T256, A: S257, A: S258, A: Q259, A: Q260, A: G261, A: P262, A: G263, A: P264, A: G265, A: N266, A: D267, A: S268, A: V269, A: H270, A: Y271, A: G272, A: Q273, A: T274, A: G380, A: F381, A: E382, A: L383, A: H384, A: E385, A: N386, A: A387, A: K388, A: N389, A: C390, A: T391, A: S392, A: K393, A: K412, A: R413, A: V414, A: T415, A: D423, A: S426, A: D427, A: K428, A: K429, A: K430	66	0.623

verified using MD evaluation via the iMODs server, which performed Normal Mode Analysis for complex variance, B-factor, deformability, covariance, eigenvalue, and elastic network model evaluation of complexes (Figs. 7 and 8). The deformability of a complex is its capacity to change shape at its residues. The deformability of a complex refers to its ability to change shape at its residues. The graph's hinges indicate areas of strong deformability. The eigenvalue indicates the complex's motion stiffness. A lower eigenvalue correlates directly with the energy required for deformation, making it easier to distort the structure. The eigenvalues of V2 and V3 with TLR4 and TLR5 are 2.278430e-08 and 6.723409e-07, respectively. Normal mode variances and eigenvalue were shown to be inversely associated. The covariance matrix shows whether pairs of two residue have uncorrelated (white), correlated (red), or anti-correlated motions. The elastic network notion refers to pairs of atoms interconnected by springs. A single spring was used to

symbolize each couple of atoms in the graph. The color of the dots reflected the stiffness of the spring, with darker grays denoting stiffer springs and vice versa.

Molecular dynamic simulation

MD simulations were performed using Schrödinger LLC's Desmond program, with simulation times of 200 ns for the vaccine-TLR4 and vaccine-TLR complexes. The root-mean-square deviation (RMSD) was estimated to evaluate the vaccine-receptor complexes' conformational stability and dynamic behavior throughout the simulation period. Protein C α atoms were analyzed using RMSD histograms to determine deviations from their starting configuration to the end state. Minor fluctuations in the RMSD curves suggested that the docked complexes are stable.

In this case the vaccine-TLR4 complex, the RMSD value stabilized following initial fluctuations, which were primarily observed in the early phase of the simulation

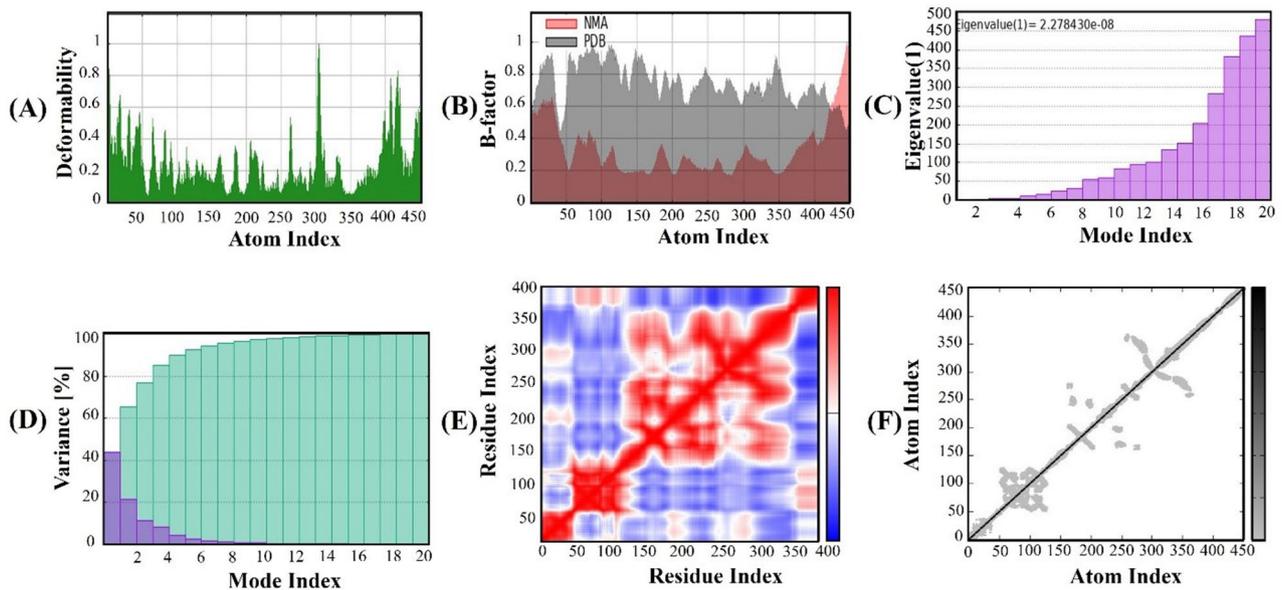


Fig. 7 Normal mode analysis of the TLR4 and vaccine complex using the iMODS server. **(A)** B-factor values are compared to regional mobility ratings ranging from 0 to 1. **(B)** Each residue in the complex's deformability index, with higher values indicating greater deformability. **(C)** Covariance map showing uncorrelated (white), correlated (red), and anti-correlated (blue) motions. **(D)** Eigenvalue: the smaller it is, the simpler it is to distort the complex. **(E)** Variance is inversely proportional to the eigenvalue. Green represents cumulative variances, whereas red represents individual deviations. **(F)** Elastic network, with darker grey patches indicating increased stiffness

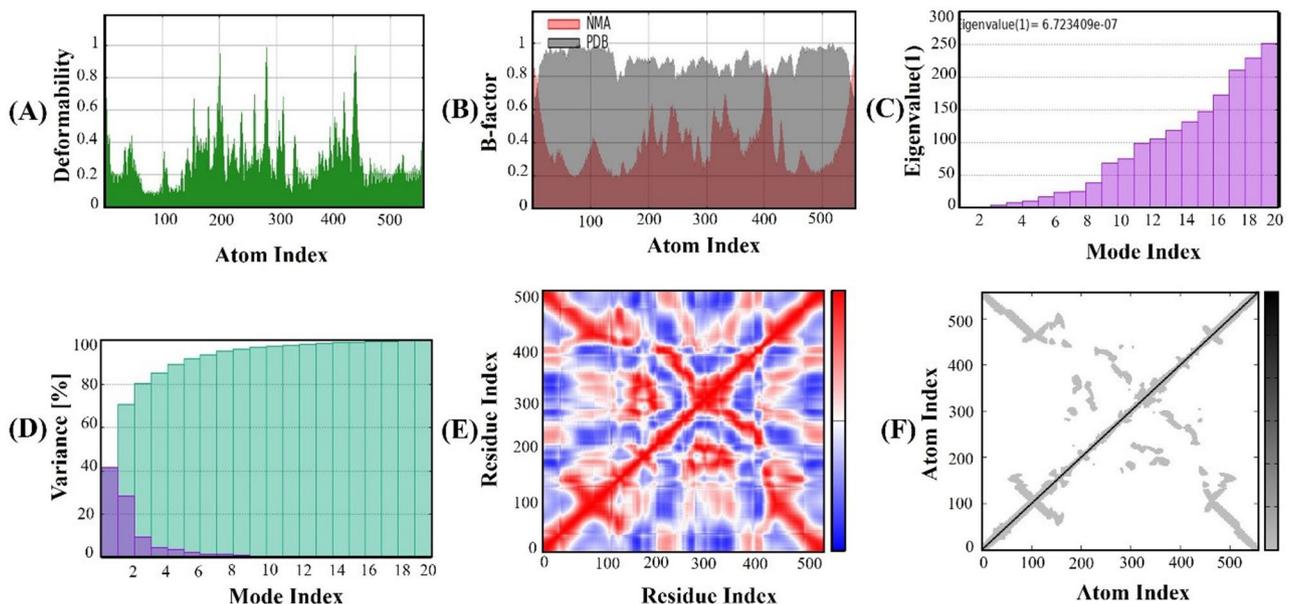


Fig. 8 Normal mode analysis of the TLR4 and vaccine complex using the iMODS server. **(A)** B-factor values are compared to regional mobility ratings ranging from 0 to 1. **(B)** Each residue in the complex's deformability index, with higher values indicating greater deformability. **(C)** Covariance map showing uncorrelated (white), correlated (red), and anti-correlated (blue) motions. **(D)** Eigenvalue: the smaller it is, the simpler it is to distort the complex. **(E)** Variance is inversely proportional to the eigenvalue. Green represents cumulative variances, whereas red represents individual deviations. **(F)** Elastic network, with darker grey patches indicating increased stiffness

(0–25 ns), with no significant variations thereafter. The average calculated RMSD for this complex was $20\text{Å} \pm 2\text{Å}$ (Fig. 9A). For the vaccine TLR5 complex, minor deviations were seen during the initial period (0–30 ns), followed by a stable trajectory for the rest of the simulation.

The RMSD for this complex was calculated as $12\text{Å} \pm 1\text{Å}$ (Fig. 9B). These fluctuation trends demonstrate the structural stability of both vaccine-receptor complexes. Observed fluctuations can be attributed to flexible regions in the vaccine, such as loop regions, linker

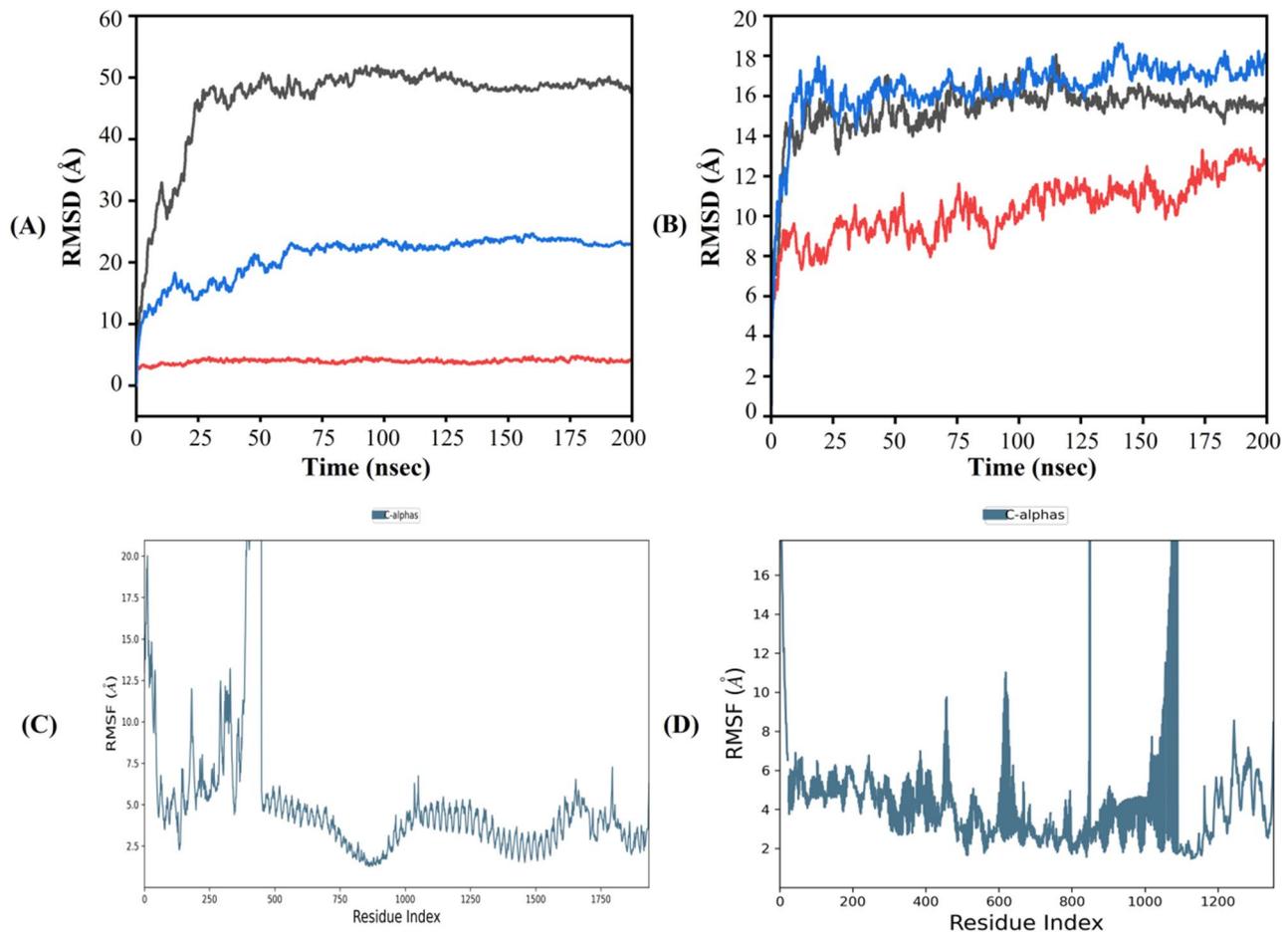


Fig. 9 RMSD and RMSF plots of protein-ligand complexes. **(A)** The RMSD plot of V1-TLR4 and **(B)** the RMSD plot of V2-TLR5 highlight the structural stability of both complexes during molecular dynamics simulations. **(C)** RMSF plot of V1-TLR4 **(D)** The RMSF plot of V2-TLR-5 demonstrates the residue-specific flexibility of each complex during molecular dynamics simulations

sequences, epitope insertions, the tPA signal peptide, and adjuvants [67]. These findings are consistent with previously reported findings [68, 69], indicating the dynamic behavior of the vaccine-receptor systems.

Root-mean-square fluctuation (RMSF) analysis was also conducted to predict local changes in protein chains. Peaks in the RMSD plots matched regions with greater movement, including flexibility within specific protein segments, confirming the stability of the vaccine-receptor interaction. The observed RMSF patterns are consistent with those reported in similar studies, reinforcing the stability and reliability of the docked complexes (Fig. 9C & D).

PCA analysis

PCA analysis of the molecular dynamics simulation data provides valuable insights into the structural and dynamic properties of the V1-TLR4 complex. The findings suggest that the complex has several stable conformations, as shown by the distinct clusters in the PCA plot. The first principal component (PC1) explains the

most substantial variation, implying that strong conformational modifications or dominating interactions generate key structural differences. The first three principal components account for approximately 92.9% of the total variation (Fig. 10), suggesting their significance in capturing the key aspects of the first component, indicating that PC1 explains most of the dataset's variability. These findings indicate that the V1-TLR4 complex is physically stable and has conformational flexibility, which could be significant for its biological functions. The PCA analysis of the molecular dynamics simulation indicates crucial structural dynamics for the V2-TLR5 complex. The first principal component, PC1, PC2, and PC3, accounts for approximately 93% of the overall variation (Fig. 11), indicating that substantial conformational shifts or dominating interactions are responsible for structural diversity in the V2-TLR5 complex. PCA plot clusters demonstrate that the V2-TLR5 complex also adopts several stable conformations during the simulations. The isolation along the PC1 axis indicates distinct primary conformational states, with variations along PC2 and PC3 providing

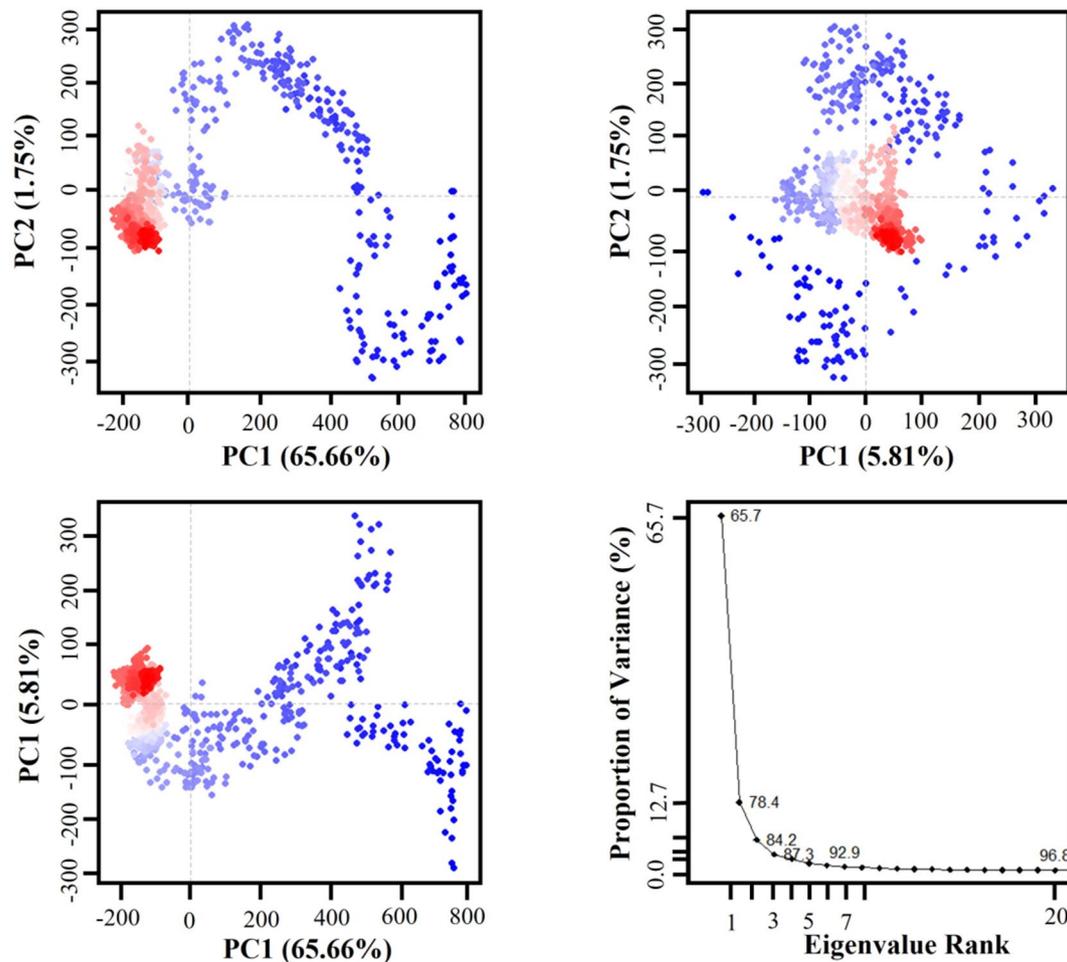


Fig. 10 Principal component analysis (PC1, PC2, and PC3) of the V1-TLR4 complex, reveals that substantial conformational shifts are responsible for structural diversity in the V1-TLR4 complex

information about small structural changes within these states. This analysis indicates that the V2-TLR5 complex maintains structural integrity and conformational flexibility, which is crucial for its potential biological activity.

DCCM analysis

The DCCM analysis of the V1-TLR4 and V2-TLR5 complexes over a 200 ns molecular dynamic simulation provides considerable insights into residue interaction and structural dynamics. The DCCM shows different patterns of positive and negative correlations on a color scale that ranges from -1.0 (green/yellow) to 1.0 . The large blue blocks along the diagonal and off-diagonal regions depict areas where certain domains of the complex and off-diagonal regions represent areas where specific domains or proportions of the complex move in highly coordinated ways for both complexes (Fig. 12). Specifically, regions surrounding residues 400 to 600 and 1000 to 1300 for the V1-TLR4 complex (Fig. 12A) and residues 100 to 400 and 600 to 800 in the V2-TLR5 complex (Fig. 12B) have substantial positive correlations, respectively, implying that

these regions travel together during the simulation. This may indicate functionally or structurally linked domains. Small green and yellow patches, on the other hand, emphasize anti-correlated motions in both complexes, which occur when residues move in opposite directions and may identify possible areas of functional antagonism. Although less common, anti-correlated motions are important for understanding the functional interplay and the complex's dynamic. Overall, the DCCM analysis sheds light on the dynamic behavior and residue interactions of the V1-TLR4 and V2-TLR5 complexes, offering guidance for further functional studies. Moreover, as the V1-TLR4 complex has more correlated movement as compared to V2-TLR5, the binding free energy analysis of V1-TLR4 was conducted, which predicted that the frames show significant fluctuations but remain consistently negative. The observed range of binding energies is approximately -59 to -111 kcal/mol, indicating strong binding affinity between the ligand and receptor (Fig. 13).

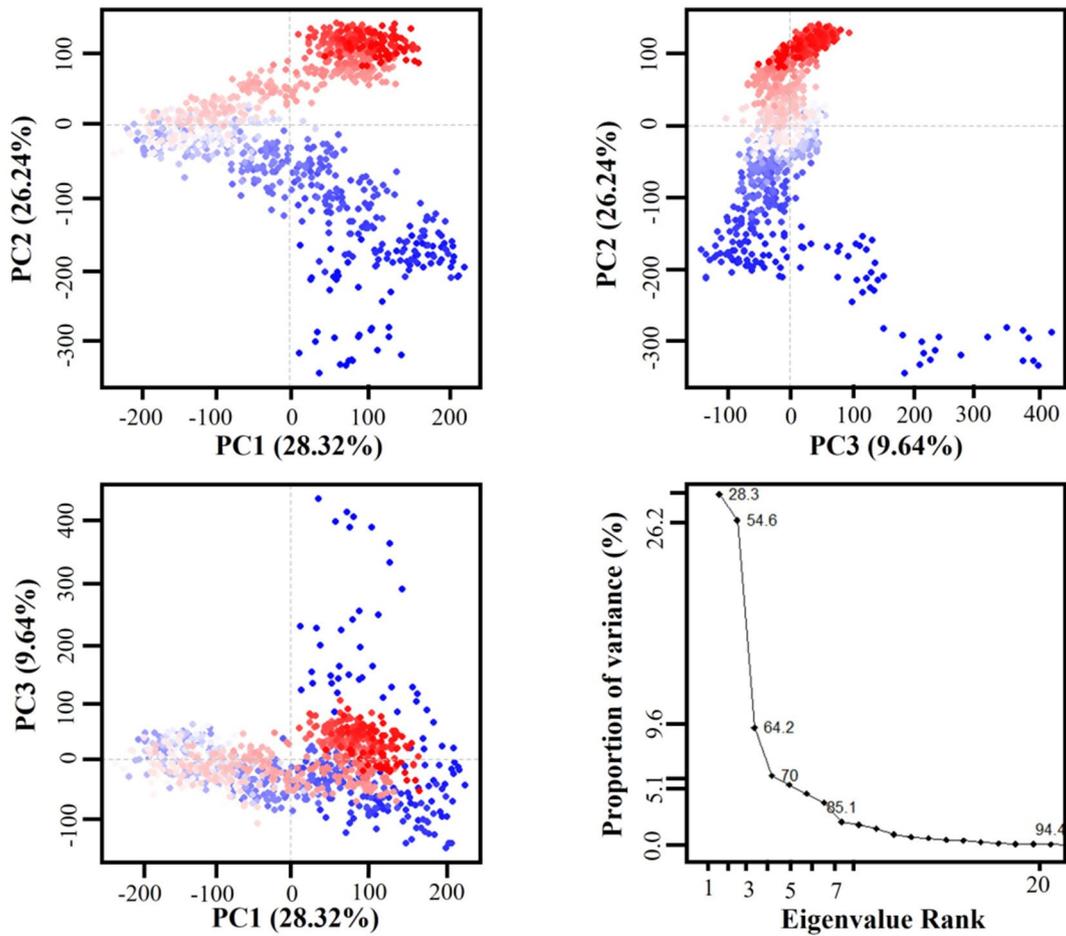


Fig. 11 Principal component analysis (PC1, PC2, and PC3) of the V1-TLR4 complex, reveals that substantial conformational shifts are responsible for structural diversity in the V2-TLR5 complex

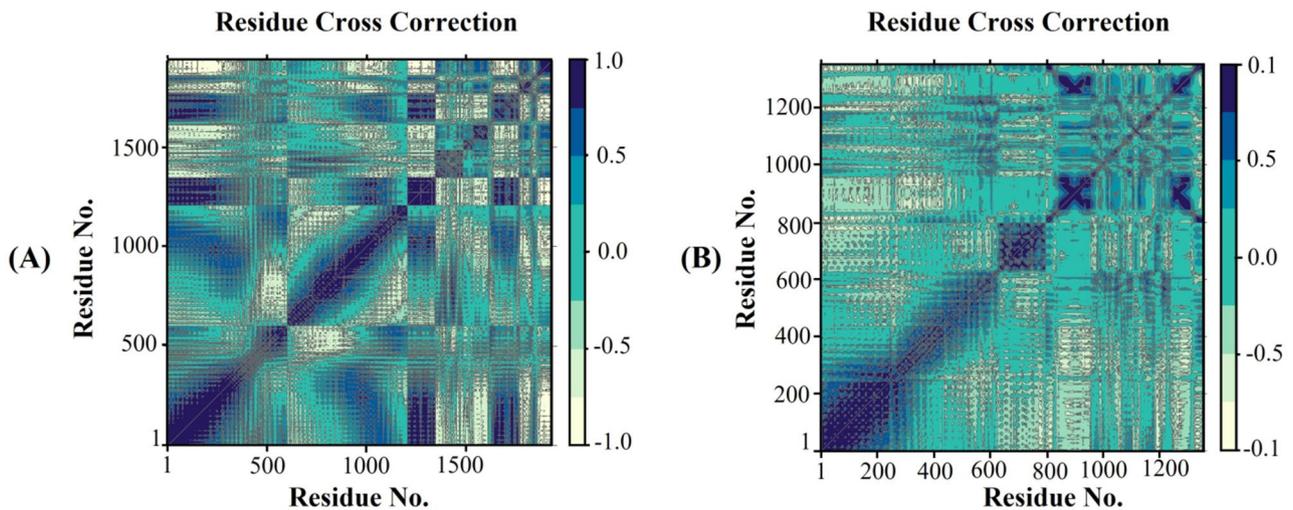


Fig. 12 Dynamic Cross-correlation Map (A) DCCM map for V1-TLR4 complex. (B) DCCM map for V2-TLR5 complex

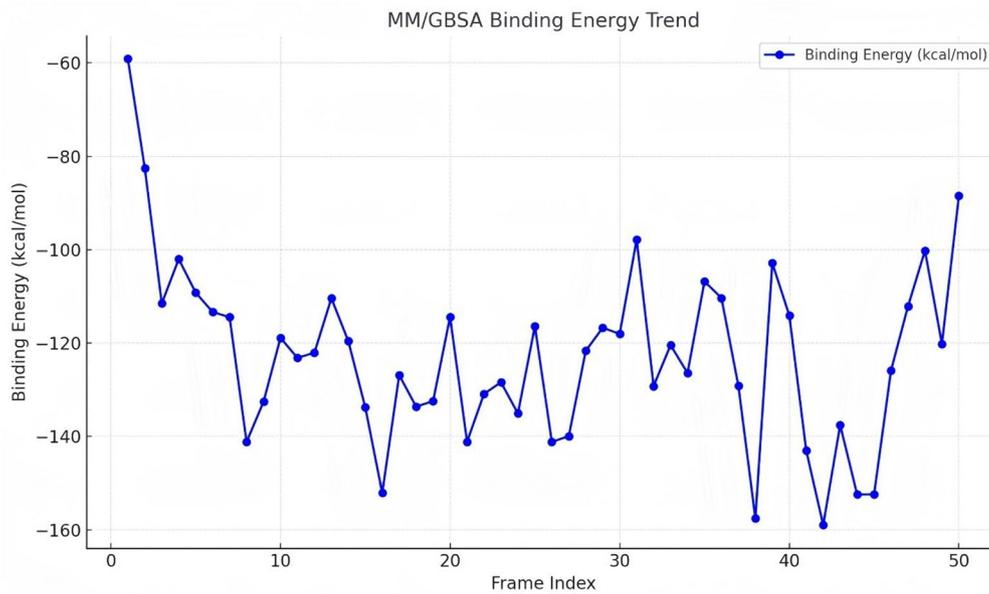


Fig. 13 MM/GBSA binding free energy trend of V1-TLR4 complex

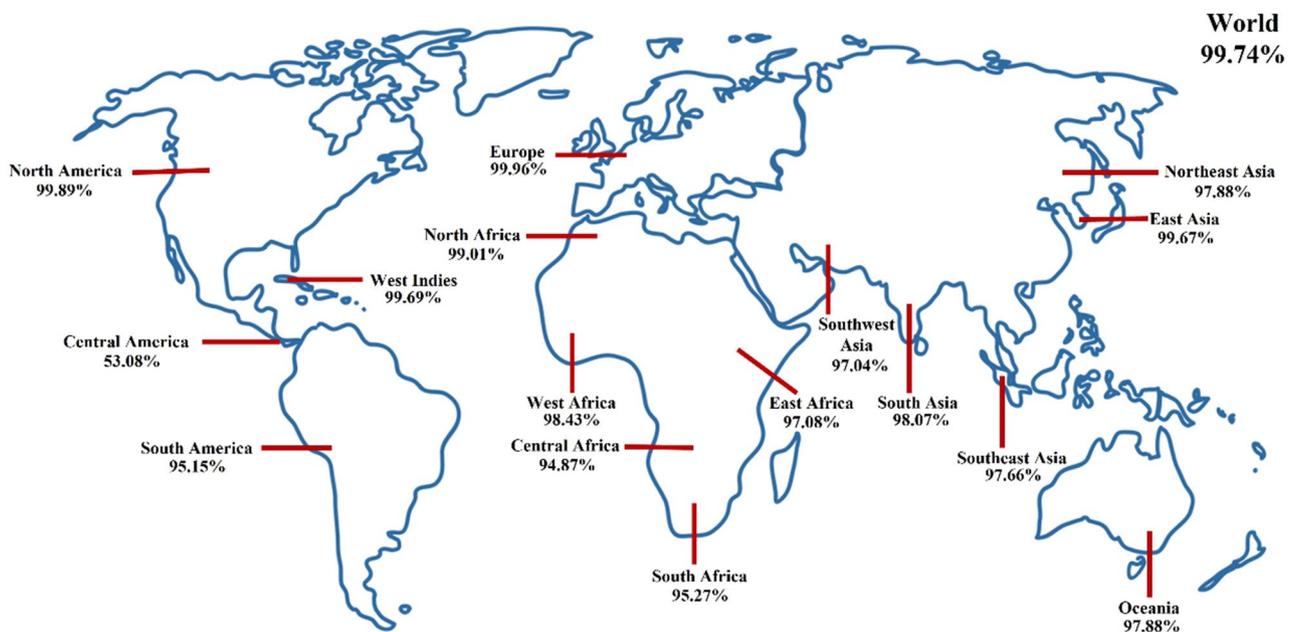


Fig. 14 Global population coverage map of the chosen T-cell epitopes

Population coverage analysis

HLA alleles vary among members of a group based on their distribution. Assessing the HLA compatibility of the population with the intended vaccine epitope is crucial for evaluating vaccination efficacy in infectious populations. The IEDB population coverage tool predicts that these epitopes are estimated to encompass 99.74% of the worldwide population (Fig. 14).

In Silico cloning and optimization of codon

Codon optimization is a technique that optimizes codons for bacterial genomes and results in higher expression rates. This method was utilized to increase the speed of manufacturing of vaccines in K12 strain of *E. coli*, as humans and the intended host use various codons. To maximize protein expression, the JCat tool was utilized to determine the best GC and CAI content. In this investigation, after sequence optimization, the CAI values were 0.72 and 0.75, and the GC content percentages were 5.26% and 51.16% for V2 and V3. The absolute GC

content ranges from 30 to 70%, and a CAI value of 0.8 is considered high. The findings align with the previous study indicating a similar range of data required for long-lasting vaccine expression. They also support the greater expression levels for the preferred vaccine. The SnapGene tool was utilized to create in silico clones of optimized V2 and V3 sequences. Sequences were inserted into the pET28a (+) vector at the PshA1 restriction site, yielding genomes of sizes 6716 and 7037 bps (Figure S4).

Immune response analysis

Immune simulation represents a critical phase in modern vaccine design, enabling the prediction of host immune responses prior to clinical testing. An effective vaccine should elicit a robust, pathogen-specific immune response while maintaining an optimal safety profile. The efficacy of a vaccine and its immunological responses are evaluated using the measurement of indicators such as antibodies, B cells, cytokines, NK cells, and T cells. To assess immunological responses elicited by V2 and V3, we conducted immune simulations. Each booster dose increased immune responses. Additionally, each dose resulted in a considerable increase in cytokine levels, which were equivalent across doses. The vaccine notably elicited cytokines production such as TGF- β , IL-8, IL-2, IFN- γ , and IL-10. V2 and V3 induced the generation of a range of targeted antibodies, including IgG, IgG1, IgG2, and IgM. Each administered dose resulted in a substantial reduction in antigen levels and a concomitant increase in antibody frequency. Following vaccination, a significant number of B cells were simulated, resulting in the formation of plasma cells, which are crucial for producing specific antibodies. After immunization, a marked increase in B-cell proliferation was observed. In addition, the T-cell count rose following vaccine administration, with a significant increase in activated T-cells and a corresponding decrease in inactive cells, indicating effective cellular activation post-vaccination. After immunization, B-cell proliferation was increased significantly. Moreover, the T-cell count rose following vaccine administration, with a significant increase in activated T-cells and a corresponding decrease in inactive cells, indicating effective cellular activation post-vaccination (Figure S5 and S6). The designed vaccines are expected to activate and interact with TLR-4 and TLR-5 expressed on B-cells, dendritic cells, macrophages, and monocytes. Through complex signaling pathways, this interaction will upregulate the expression of IL-6, IL-12, IL-8, IP-10, CD88, CD-80, MCP-1, and IFN genes. Additionally, this will indirectly activate other immune cells such as NK cells, monocytes, and T cells. Moreover, the HTL and CTL epitopes will bind with HLA Class I and HLA Class II molecules, producing epitope-HLA complexes that subsequently engage with CTLs and HTLs, leading to their activation

and proliferation. The IFN- γ epitopes will further stimulate IFN gene expression. Thus, the suggested vaccine has the potential to activate both the innate and adaptive immune systems (Fig. 15).

Discussion

Preventing multidrug-resistant *A. duodenale* infections presents a substantial public health challenge due to three key clinical consequences: (1) hookworm-associated maternal morbidity during pregnancy [70], (2) chronic intestinal helminthiasis with associated mucosal damage, and (3) iron-deficiency anemia resulting from blood loss. *A. duodenale*, due to its widespread distribution, significant global health impact, the emergence of drug resistance, economic burden, vaccine development potential, and the one health perspective, has drawn the attention of researchers to target this parasite. Despite noted disparities in antibiotic rates, resistance continues to increase, prompting researchers to seek alternate therapies. Research on hookworm vaccines, including *Ancylostoma duodenale* infection, indicates that live attenuated larvae offer better protection than dead larvae [71, 72]. However, they face issues like impracticality for widespread use, limited shelf life, and high production costs. Most recent studies focus on protein subunit vaccines, particularly aspartic protease-1 (Na-APR-1) and glutathione-S-transferase-1 (Na-GST-1), currently in progress for treating *A. duodenale* infection. Challenges with these vaccines include the need for adjuvants, multiple injections, and low production yield [15, 16]. Moreover, peptide candidates derived from Na-APR-1, such as A291Y-GCN4 and p3-P25, show potential but lack sufficient immunogenicity and validation in challenge models. Our study aims to tackle these issues using an in silico multi-epitope vaccine design approach. Bioinformatics is being used to design unique vaccines and drug targets against different parasites using in silico approaches [73, 74]. To locate an alternative therapeutic target, subtractive proteomics was paired with reverse vaccinology to uncover possible vaccine targets.

In this study, a reference genome of *A. duodenale* was used to extract the proteins of the parasite. Previous studies on combating *A. duodenale* infections involved surveys and experimental methods, but no methodologies were used for the formulation of an effective treatment against this pathogen. In silico analysis identified 36 essential proteins, human non-homologous, and host-interacting potentially serving as novel drug or vaccine candidates. During screening, host homologous proteins with significant identity such as KIH65223.1 (with a % identity of 100%) and KIH44842.1 (with a % identity of 98.6%) were identified. Such proteins were carefully considered and excluded from our vaccine design pipeline to minimize the risk of off-target effects and avoid the risk

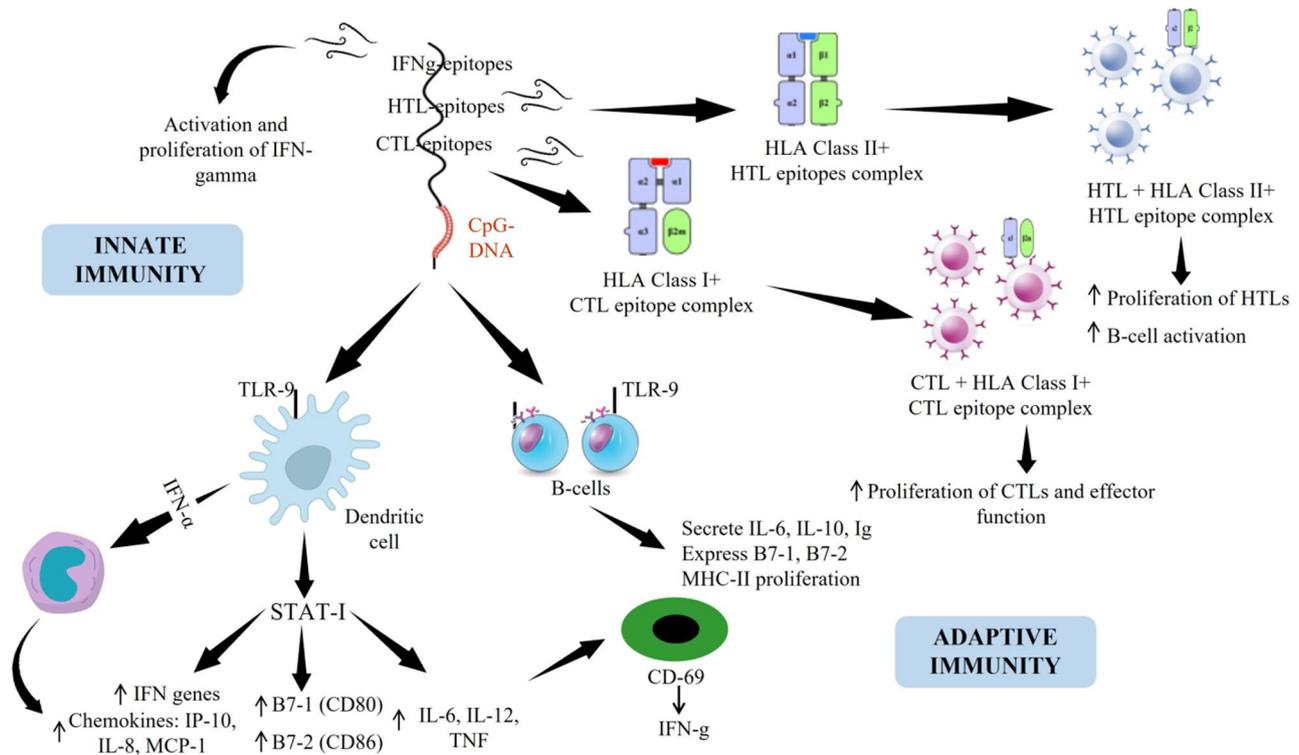


Fig. 15 The proposed vaccine mechanism might have the potential to stimulate both the adaptive and innate immune systems

of cross-reactivity. This approach is in line with recent studies in rational vaccine design, which emphasize the importance of identifying and excluding host-homologous proteins to ensure target specificity and reduce the risk of autoimmune responses, as demonstrated in nematode parasite vaccine development [75]. The reverse vaccinology approach prioritizes membrane proteins included in host-pathogen interaction as potential vaccine targets [76]. Six vaccine candidates were chosen for chimeric vaccine development following the analysis of various factors.

Eukaryotic aspartyl proteases (KIH49483.1) are enzymes found in all eukaryotes carrying out a range of functions, including protein maturation, signal transduction, and secretion of virulence factors [77]. Receptor L domain protein (KIH57671.1) is a transmembrane protein. The extracellular domain of a receptor protein is responsible for ligand binding. In contrast, the cytoplasmic domain of a receptor may have several roles. One of these roles may be to interact with the proteins responsible for the formation of the clathrin-coated pit [78]. Low-density lipoprotein receptor domain class A (KIH58423.1) plays an essential role in the transport of approximately 70% of LDL from plasma to liver and 75% of cholesterol and, hence, plays a vital role in lipoprotein metabolism [79]. N-acetylmuramoyl-L-alanine amidases (KIH59417.1) are cell wall-modifying enzymes that break the amide linkage between stem peptides and sugar

residues in peptidoglycan. Amidases play an important role in septal cell wall cleavage and assist separate daughter cells during cell division [80]. Calcium-binding EGF domain proteins (KIH68051.1) are found in a wide range of extracellular proteins with various functions, including cell adhesion, blood coagulation, connective tissue architecture, and determination of cell fate [81]. Beta-ketoacyl synthase protein (KIH68881.1) is also known as beta-ketoacyl-ACP synthase. The three types of beta-ketoacyl synthases in type II FAS are FabB, FabE, and FabH synthases. FabH catalyzes the classic ketoacyl synthase reaction involving malonyl ACP and acetyl CoA. FabB and FabF accelerate further related processes [82]. The survival and virulence of the parasite rely on proteins that differ from host proteins. Inhibiting these proteins may enhance the host's ability to develop an efficient immunological response, resulting in improved infection management.

Two proteins (KIH49483.1 and KIH58423.1), selected based on specific criteria, were identified as potential vaccine targets for epitope determination. The minimal immunogenic peptides within these proteins, known as epitopes, can activate either cytotoxic T-cells (CTLs) or helper T-cells (HTLs) [18]. T-cell interacting epitopes play a crucial role in aiding HTLs and CTLs to produce specific immunoglobulins and eliminate pathogens, thereby enhancing the efficacy of multi-epitope vaccines. To facilitate the development of an effective vaccine,

T- and B-cell epitopes were predicted from these protein targets and subsequently combined with adjuvants and linkers for the formulation of the chimeric vaccine constructs. MHC-I, MHC-II, and B-cell epitopes were joined using the Gly-Pro-Gly-Pro-Gly (GPGPG), bi-lysine (KK), and Ala-Ala-Tyr (AAY) linkers. Immunogenic linkers enhance the immunogenicity of the proposed vaccines. In addition, linkers help keep epitopes apart and prevent folding [83, 84]. The linker AAY connects two protein domains. This linker increase the immunological response to vaccine antigen by allowing immune cells more access to antigenic sites while maintaining the structural stability of vaccine antigens [85]. The GPGPG linker is a flexible component in vaccine design that allows for easy folding of the vaccine's antigens. It reduces antigen aggregation and can affect the immunogenicity and stability of the antigen [86]. The KK linker is commonly utilized to link carrier protein to protein antigens, such as a bacterial toxin. The positive charge of the linker helps maintain the connection among the antigen and carrier protein, leading to a greater immunological response [87]. Using an empirical-helical linker, EAAAK, enhanced the bi-functional catalytic activity, fusion protein's stability, and stiffness [85].

Six adjuvants, specifically beta-defensin, L7/L12 ribosomal protein, flagellin, HBHA, HBHA conserved, and granulocyte-macrophage colony, were employed in this analysis to design six vaccine constructs with different immunological profiles. The L7/L12 ribosomal protein is an adjuvant applied in vaccines that enhances immune responses by promoting dendritic cell (DC) maturation and T-cell activation [88]. Beta-defensins are naturally occurring antimicrobial peptides also employed as vaccine adjuvants, improving both cellular and humoral immune responses through innate immunity activation, via cytokine release [89]. Moreover, HBHA and its conserved form are powerful immunological adjuvants with the ability to induce an intense T_{H1} cell immune response. Another adjuvant called Flagellin is a potent immune activator that stimulates diverse biological effects and mediates both innate inflammatory responses and the development of adaptive immunity [90]. The granulocyte-macrophage colony as an adjuvant encourages the development and maturation of dendritic cells, stimulates the expression of molecules like major histocompatibility complex class II, and stimulates the proliferation and differentiation of B and T lymphocytes, which play an essential role in activating immune responses [91–93].

The vaccine candidates demonstrated optimal physicochemical characteristics and contained sufficient antigenic determinants to elicit a robust immune response. The evaluation results from multiple servers indicated that vaccine constructs were non-allergenic and soluble. The vaccine's structure analysis revealed no deformability

in its secondary structures. TrRosetta server created the 3D structures of the vaccine, which were subsequently refined by using the Galaxy web server. Protein-protein docking studies using immune receptors TLR-4 and TLR-5 were conducted to investigate the function of vaccine targets in inducing an innate immune response. Toll-like receptors are found on multiple immune cells, including macrophages, cancer cells NK cells. They discriminate between damage-associated patterns (DAMPs) and pathogen-associated patterns (PAMPs), activating the immune system. TLR-4 exhibits superior tumor antigen detection and can provoke both adaptive and innate immune reactions among the 10 classes of TLRs [94]. The extensive expression of TLR5 across multiple immune cell types contributes to flagellin's adjuvant properties. Flagellin can activate TLR, which is shown on the surface of immune cells [95, 96].

The ClusPro 2.0 results showed that V2 and V3 have a strong affinity for immunological receptors. Normal mode analysis and MD simulation were used to evaluate the binding affinity of the vaccine constructs with immune cell receptors and illustrate the improved stability of the V2-TLR4 and V3-TLR5 complexes. A discontinuous or conformational epitope is a cluster of antigenic residues that are isolated in the main sequence but come together spatially due to polypeptide folding. Conformational epitopes account for approximately 90% of all B-cell epitopes [97]. Therefore, the residues responsible for forming the conformational B-cell epitopes of the vaccines were also determined. V2 and V3 docked complexes are stable because of the low energy required for distortion, as seen by the lower eigenvalues in the NMA study. The analysis of B-factor, eigenvalue, deformability, covariance, the elastic network model, and variance revealed a significant relationship between vaccine models and immunological receptors (TLR4 and TLR5), indicating that vaccines may trigger an innate immune response [98, 99].

HLA allele distribution exhibits ethnic variation and significantly affects T-cell epitope recognition. To maximize population coverage, selected epitopes must bind multiple HLA alleles. We therefore prioritized MHC-I and MHC-II epitopes based on their corresponding HLA allele frequencies worldwide. Computational analysis demonstrated that the selected epitopes provide 99.77% global population coverage across diverse geographic regions. Moreover, to enhance codon usage by the host organism, both vaccines were codon-adapted and cloned into the pET28-a (+) vector. Immune simulations were performed using C-ImmSim, an online tool, to better understand the functions of vaccines in inducing the passive immune response. Server findings indicate that the vaccine can boost the host's immunoglobulin levels and maintain memory B cells effectively after three

doses. Additionally, the presence of IFN-gamma verifies our anticipation that vaccines can trigger an innate immune response. Bioinformatics techniques confirmed that the developed multi-epitope vaccines effectively elicit both acquired and innate immune responses, hence successfully eliminating the *A. duodenale*. The vaccine recommended by immunoinformatics is believed to be immunogenic, but its effectiveness against *A. duodenale* infections remains unknown. Immunoinformatics approaches enable direct lab investigations and in-silico research, saving expenses and time. The next step is to carry out in-vitro immunological testing to confirm the immunogenicity of vaccines and construct preclinical challenge-protection studies to verify the methodology. Previously, vaccines targeting various pathogens, including *Ebola virus*, *Marburg virus*, *Trypanosoma vivax*, and *Acinetobacter baumannii*, were designed using computational methods and suitably proven experimentally [100–102]. Therefore, the final vaccine constructs, V2 and V3, can be subjected to the wet lab for in vivo and in vitro investigation, potentially offering a treatment against *A. duodenale*.

Despite the promising outcomes of this in silico approach, several limitations should be acknowledged. Epitope prediction techniques are based on machine learning algorithms and existing datasets that may introduce biases and lead to false positives or negatives due to the antigenic regions' variability across different pathogens and populations [103]. Similarly, docking studies rely on simplified scoring functions and static 3D structures that may not accurately reflect the complex and dynamic nature of molecular interactions under physiological conditions [104]. MD simulations operate within theoretical force fields and constrained time scales and thus may not fully capture the intricate biological environment [105]. Additionally, computational predictions for allergenicity, antigenicity, and toxicity require experimental validation, as they cannot fully substitute for in vivo and in vitro validation. Therefore, while the computational pipeline offers a strong foundation for the design of a vaccine, experimental studies are indispensable for validating the immunogenicity, safety, and efficacy of the proposed vaccine candidates.

Conclusion

Ancylostoma duodenale presents a significant global public health challenge, prompting the development of an improved vaccine strategy. This analysis identifies potential broad-spectrum epitope candidates for constructing vaccines that might protect against infections caused by *A. duodenale*. Our research suggests that the proposed vaccine constructs might effectively induce immune responses. However, experimental validation of the final vaccine constructs is essential to assess their potential

for immunological reinforcement and therapeutic applications. The current research provides a foundation for developing epitope-based vaccines against *A. duodenale* through experimental approaches.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11652-4>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6

Acknowledgements

The authors express their appreciation to the Deanship of Scientific Research at King Khalid University, Saudi Arabia, for this work through a research group program under grant number RGP-2/623/45. The authors wish to thank the Bahauddin Zakariya University, Multan, Pakistan for providing the necessary infrastructure to perform this research.

Author contributions

MS: Conceptualization, Supervision, Resources, Writing–review & editing. HA: Formal analysis, Investigation, Writing– original draft. AS: Formal analysis, Investigation, Writing– original draft. MA: Formal analysis, Investigation, Software. ARH: Formal analysis, Investigation. MUK: Investigation, Software, Writing– original draft. KFF: Investigation, validation, data curation, funding acquisition, Writing–review & editing. SAA: Formal analysis, Investigation, Validation. SMNA: Visualization, data curation, Writing–review & editing. AA: Data curation, visualization. UN: Writing–review & editing. KC: Conceptualization, Resources, Writing–review & editing.

Funding

This study was funded by the Doctoral Research Fund, the affiliated hospital of southwest medical university, China, awarded to KC.

Data availability

The *Ancylostoma duodenale* reference genome (GenBank: GCA_000816745.1) was downloaded from the National Center for Biotechnology Information (NCBI). The data analyzed in this study is available in the tables and supplementary files of the manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Biochemistry, Bahauddin Zakariya University, Multan, Punjab 66000, Pakistan

²Department of Animal Science, Federal University of Ceara, Fortaleza, Brazil

³Department of Pharmacy, Faculty of Biological Science and Technology, Jashore University of Science and Technology, Jashore 7408, Bangladesh

⁴Institute of Molecular Biology and Biotechnology, The University of Lahore, Lahore, Pakistan

⁵Chemistry Department, Faculty of Science, King Khalid University, P.O. Box 9004, Abha 61413, Saudi Arabia

⁶Research Center for Advanced Materials Science (RCAMS), King Khalid University, P.O. Box 960, AlQura'a, Abha, Saudi Arabia

⁷Department of Clinical Microbiology and Immunology, Faculty of Medicine, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁸Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

⁹Department of Biology, Faculty of Science and Arts, Najran University, Najran 1988, Saudi Arabia

¹⁰Department of Zoology, Abdul Wali Khan University, Mardan, Khyber Pakhtunkhwa 23200, Pakistan

¹¹Department of Chemistry, Kohat University of Science & Technology, Kohat, Pakistan

¹²Department of Infectious Diseases, The Affiliated Hospital of Southwest Medical University, Luzhou 646000, China

Received: 19 March 2025 / Accepted: 29 April 2025

Published online: 12 May 2025

References

1. Chan MS, Medley GF, Jamison D, Bundy DAP. The evaluation of potential global morbidity attributable to intestinal nematode infections. *Parasitology*. 1994;109:373–87.
2. Hotez PJ, Bethony J, Bottazzi ME, Brooker S, Buss P. Hookworm. The great infection of mankind. *PLoS Med*. 2005;2:e67.
3. Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, et al. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *Lancet*. 2006;367:1521–32.
4. Garn JV, Wilkers JL, Meehan AA, Pfadenhauer LM, Burns J, Imtiaz R et al. Interventions to improve water, sanitation, and hygiene for preventing soil-transmitted helminth infection. *Cochrane Database of Systematic Reviews*. 2022;2022.
5. Yu Sen-Hai, Jiang Ze-Xiao, Xu Long-Qi. Infantile hookworm disease in China. A review. *Acta Trop*. 1995;59:265–70.
6. Peduzzi R, Piffaretti JC. *Ancylostoma duodenale* and the saint Gothard anaemia. *BMJ*. 1983;287:1942–5.
7. Stracke K, Jex AR, Traub RJ. Zoonotic ancylostomiasis: an update of a continually neglected zoonosis. *Am J Trop Med Hyg*. 2020;103:64–8.
8. Inpankaew T, Schär F, Dalsgaard A, Khieu V, Chimnoi W, Chhoun C et al. High prevalence of *Ancylostoma ceylanicum* hookworm infections in humans, Cambodia, 2012. *Emerg Infect Dis*. 2014;20.
9. Chen J-M, Zhang X-M, Wang L-J, Chen Y, Du Q, Cai J-T. Overt Gastrointestinal bleeding because of hookworm infection. *Asian Pac J Trop Med*. 2012;5:331–2.
10. Reynolds JA, Behne JM, Pallant LJ, Macnish MG, Gilbert F, Giles S, et al. Failure of pyrantel in treatment of human hookworm infections (*Ancylostoma duodenale*) in the Kimberley region of North West Australia. *Acta Trop*. 1997;68:301–12.
11. HORTON J. Albendazole: a review of anthelmintic efficacy and safety in humans. *Parasitology*. 2000;121:S113–32.
12. Samuel F, Degarege A, Erko B. Efficacy and side effects of albendazole currently in use against *Ascaris*, *trichuris* and hookworm among school children in Wondo Genet, Southern Ethiopia. *Parasitol Int*. 2014;63:450–5.
13. Ferrara P, Bersani I, Bottaro G, Vitelli O, Liberatore P, Gatto A, et al. Massive proteinuria: A possible side effect of pyrantel Pamoate? *Ren Fail*. 2011;33:534–6.
14. Curico G, Garcia-Bardales P, Pinedo T, Shapiama W, Moncada-Yaicate M, Romaina L, et al. Resistance to single dose albendazole and reinfection with intestinal helminths among children ages 2 to 11 years from the Peruvian Amazon region: a study protocol. *BMC Infect Dis*. 2022;22:528.
15. Diemert DJ, Zumer M, Campbell D, Grahek S, Li G, Peng J, et al. Safety and immunogenicity of the Na-APR-1 hookworm vaccine in infection-naïve adults. *Vaccine*. 2022;40:6084–92.
16. Diemert DJ, Pinto AG, Freire J, Jariwala A, Santiago H, Hamilton RG, et al. Generalized urticaria induced by the Na-ASP-2 hookworm vaccine: implications for the development of vaccines against helminths. *J Allergy Clin Immunol*. 2012;130:169–e1766.
17. Hoagland KE, Schad GA. Necator americanus and *Ancylostoma duodenale*: life history parameters and epidemiological implications of two sympatric hookworms of humans. *Exp Parasitol*. 1978;44:36–49.
18. Zhang R. DEG: a database of essential genes. *Nucleic Acids Res*. 2004;32:D271–272.
19. Dyer MD, Murali TM, Sobral BW. Computational prediction of host-pathogen protein–protein interactions. *Bioinformatics*. 2007;23:i159–66.
20. Rahman N, Shah M, Muhammad I, Khan H, Imran M. Genome-wide core proteome analysis of *Brucella melitensis* strains for potential drug target prediction. *Mini-Reviews Med Chem*. 2020;21:2778–87.
21. Aslam M, Shehroz M, Hizbullah, Shah M, Khan MA, Afridi SG, et al. Potential druggable proteins and chimeric vaccine construct prioritization against *Brucella melitensis* from species core genome data. *Genomics*. 2020;112:1734–45.
22. Shah M, Jaan S, Fatima B, Javed MS, Amjad A, Khan A, et al. Delineating novel therapeutic drug and vaccine targets for *Staphylococcus cornubiensis* NW1T through computational analysis. *Int J Pept Res Ther*. 2021;27:181–95.
23. Jaan S, Shah M, Ullah N, Amjad A, Javed MS, Nishan U et al. Multi-epitope chimeric vaccine designing and novel drug targets prioritization against multi-drug resistant *Staphylococcus pseudintermedius*. *Front Microbiol*. 2022;13.
24. Fatoba AJ, Okpeku M, Adeleke MA. Subtractive genomics approach for identification of novel therapeutic drug targets in *Mycoplasma genitalium*. *Pathogens*. 2021;10:921.
25. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res*. 2007;35(Web Server):W182–5.
26. Shurety W, Merino-Trigo A, Brown D, Hume DA, Stow JL. Localization and Post-Golgi trafficking of tumor necrosis Factor- α in macrophages. *J Interferon Cytokine Res*. 2000;20:427–38.
27. Thumuluri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res*. 2022;50:W228–34.
28. Savojardo C, Martelli PL, Fariselli P, Profti G, Casadio R. BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res*. 2018;46:W459–66.
29. Yu C-S, Cheng C-W, Su W-C, Chang K-C, Huang S-W, Hwang J-K, et al. CEL-LO2GO: A web server for protein subcellular localization prediction with functional gene ontology annotation. *PLoS ONE*. 2014;9:e99368.
30. Doytchinova IA, Flower DR. Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8:4.
31. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2—a server for in Silico prediction of allergens. *J Mol Model*. 2014;20:2278.
32. Kall L, Krogh A, Sonnhammer ELL. Advantages of combined transmembrane topology and signal peptide prediction—the phobius web server. *Nucleic Acids Res*. 2007;35:429–32. Web Server:W.
33. Magnan CN, Randall A, Baldi P. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics*. 2009;25:2200–7.
34. Jaan S, Zaman A, Ahmed S, Shah M. mRNA vaccine designing using Chikungunya virus E glycoprotein through Immunoinformatics-Guided approaches. *Vaccines (Basel)*. 2022;10:1476.
35. Shah M, Anwar A, Qasim A, Jaan S, Sarfraz A, Ullah R et al. Proteome level analysis of drug-resistant *Prevotella melaninogenica* for the identification of novel therapeutic candidates. *Front Microbiol*. 2023;14.
36. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GPS. Silico approach for predicting toxicity of peptides and proteins. *PLoS ONE*. 2013;8:e73957.
37. Saha S, Raghava GPS. Prediction Methods for B-cell Epitopes. 2007. pp. 387–94.
38. Kolla HB, Tirumalasetty C, Sreerama K, Ayyagari VS. An immunoinformatics approach for the design of a multi-epitope vaccine targeting super antigen TSST-1 of *Staphylococcus aureus*. *J Genetic Eng Biotechnol*. 2021;19:69.
39. Aram C, Alijanizadeh P, Saleki K, Karami L. Development of an ancestral DC and TLR4-inducing multi-epitope peptide vaccine against the Spike protein of SARS-CoV and SARS-CoV-2 using the advanced immunoinformatics approaches. *Biochem Biophys Rep*. 2024;39:101745.
40. Tarahimofrad H, Rahimnahl S, Zamani J, Jahangirian E, Aminzadeh S. Designing a multi-epitope vaccine to provoke the robust immune response against influenza A H7N9. *Sci Rep*. 2021;11:24485.
41. Guerrero GG, Feunou FP, Loch C. The coiled-coil N-terminal domain of the heparin-binding haemagglutinin is required for the humoral and cellular immune responses in mice. *Mol Immunol*. 2008;46:116–24.
42. Keshavam CC, Naz S, Gupta A, Sanyal P, Kochhar M, Gangwal A, et al. The heparin-binding hemagglutinin protein of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *J Biol Chem*. 2023;299:105364.

43. Alshammari A, Alharbi M, Alghamdi A, Alharbi SA, Ashfaq UA, Tahir ul Qamar M, et al. Computer-Aided Multi-Epitope vaccine design against Enterobacter xiangfangensis. *Int J Environ Res Public Health*. 2022;19:7723.
44. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the expasy server. *The proteomics protocols handbook*. Totowa, NJ: Humana; 2005. pp. 571–607.
45. Javadi M, Oloomi M, Bouzari S. In Silico design of a Poly-epitope vaccine for urinary tract infection based on conserved antigens by modern vaccinology. *Int J Pept Res Ther*. 2021;27:909–21.
46. Aarthy M, Pandiyan GN, Paramasivan R, Kumar A, Gupta B. Identification and prioritisation of potential vaccine candidates using subtractive proteomics and designing of a multi-epitope vaccine against *Wuchereria bancrofti*. *Sci Rep*. 2024;14:1970.
47. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16:404–5.
48. Simbulan AM, Banico EC, Sira EMJS, Odchimar NMO, Orosco FL. Immunoinformatics-guided approach for designing a pan-proteome multi-epitope subunit vaccine against African swine fever virus. *Sci Rep*. 2024;14:1354.
49. Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*. 2013;41:W384–8.
50. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 1993;26:283–91.
51. Kozakov D, Hall DR, Xia B, Porter KA, Padhorney D, Yueh C, et al. The cluspro web server for protein–protein Docking. *Nat Protoc*. 2017;12:255–78.
52. Ashley ESD, Lewis R, Lewis JS, Martin C, Andes D. Pharmacology of systemic antifungal agents. *Clin Infect Dis*. 2006;43 Supplement1:S28–39.
53. Lopéz-Blanco JR, Garzón JI, Chacón P. iMod: multipurpose normal mode analysis in internal coordinates. *Bioinformatics*. 2011;27:2843–50.
54. Ponomarenko J, Bui H-H, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*. 2008;9:514.
55. Bowers KJ, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE, Chow E et al. Molecular dynamics—Scalable algorithms for molecular dynamics simulations on commodity clusters. In: *Proceedings of the 2006 ACM/IEEE conference on Supercomputing - SC '06*. New York, New York, USA: ACM Press; 2006. p. 84.
56. Hildebrand PW, Rose AS, Tiemann JKS. Bringing molecular dynamics simulation data into view. *Trends Biochem Sci*. 2019;44:902–13.
57. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013;27:221–34.
58. Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J Chem Theory Comput*. 2010;6:1509–19.
59. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*. 2006;22:2695–6.
60. Grote A, Hiller K, Scheer M, Munch R, Nortemann B, Hempel DC et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res*. 2005;33 Web Server:W526–31.
61. Dey J, Mahapatra SR, Patnaik S, Lata S, Kushwaha GS, Panda RK, et al. Molecular characterization and designing of a novel multi-epitope vaccine construct against *Pseudomonas aeruginosa*. *Int J Pept Res Ther*. 2022;28:49.
62. Aslam M, Shehroz M, Ali F, Zia A, Pervaiz S, Shah M, et al. Chlamydia trachomatis core genome data mining for promising novel drug targets and chimeric vaccine candidates identification. *Comput Biol Med*. 2021;136:104701.
63. Rapin N, Lund O, Castiglione F. Immune system simulation online. *Bioinformatics*. 2011;27:2013–4.
64. Ahmad S, Verli H. In Silico identification of drug targets and vaccine candidates against *Bartonella Quintana*: a subtractive proteomics approach. *Mem Inst Oswaldo Cruz*. 2024;119.
65. Sanchez-Trincado JL, Gomez-Perosanz M, Reche PA. Fundamentals and methods for T- and B-Cell epitope prediction. *J Immunol Res*. 2017;2017:1–14.
66. Christensen DG, Xie X, Basisty N, Byrnes J, McSweeney S, Schilling B et al. Post-translational protein acetylation: an elegant mechanism for Bacteria to dynamically regulate metabolic functions. *Front Microbiol*. 2019;10.
67. Kaushik V, Jain P, Akhtar N, Joshi A, Gupta LR, Grewal RK, et al. Immunoinformatics-Aided design and *In vivo* validation of a Peptide-Based multi-epitope vaccine targeting canine circovirus. *ACS Pharmacol Transl Sci*. 2022;5:679–91.
68. Akhtar N, Magdalena JSL, Ranjan S, Wani AK, Grewal RK, Oliva R, et al. Secreted aspartyl proteinases targeted Multi-Epitope vaccine design for *Candida Dubliniensis* using immunoinformatics. *Vaccines (Basel)*. 2023;11:364.
69. Kaushik V, Gupta GSK, Kalra LR, Shaikh U, Cavallo AR. L. Immunoinformatics aided design and In-Vivo validation of a Cross-Reactive peptide based Multi-Epitope vaccine targeting multiple serotypes of dengue virus. *Front Immunol*. 2022;13.
70. Navitsky RC, Dreyfuss ML, Shrestha J, Khatri SK, Stoltzfus RJ, Albonico. *Ancylostoma duodenale* is responsible for hookworm infections among pregnant women in the rural plains of Nepal. *J Parasitol*. 1998;84:647.
71. Vinayak VK, Gupta NK, Chopra AK, Sharma GL, Kumar A. Efficacies of vaccines against canine hookworm disease. *Parasitology*. 1981;82:375–82.
72. Fujiwara RT, Loukas A, Mendez S, Williamson AL, Bueno LL, Wang Y, et al. Vaccination with irradiated *Ancylostoma Caninum* third stage larvae induces a Th2 protective response in dogs. *Vaccine*. 2006;24:501–9.
73. Sarfraz A, Wara TU, Sheheryar, Chen K, Ansari SH, Zaman A et al. Structural informatics approach for designing an epitope-based vaccine against the brain-eating *Naegleria fowleri*. *Front Immunol*. 2023;14.
74. Shah M, Sitara F, Sarfraz A, Shehroz M, Wara TU, Perveen A et al. Development of a subunit vaccine against the cholangiocarcinoma causing *Opisthorchis Viverrini*: a computational approach. *Front Immunol*. 2024;15.
75. Nebangwa DN, Shey RA, Shadrack DM, Shintouo CM, Yaah NE, Yengo BN, et al. Predictive immunoinformatics reveal promising safety and anti-onchocerciasis protective immune response profiles to vaccine candidates (Ov-RAL-2 and Ov-103) in anticipation of phase I clinical trials. *PLoS ONE*. 2024;19:e0312315.
76. Lu J, Yu Y, Zhu I, Cheng Y, Sun P. Structural mechanism of serum amyloid A-mediated inflammatory amyloidosis. *Proc Natl Acad Sci*. 2014;111:5189–94.
77. Bhattacharya S, Parveen S, Mukherjee B. Elaborating the role of aspartyl protease in host modulation and invasion in apicomplexan parasites *Plasmodium* and *Toxoplasma*. *Pathobiology of parasitic protozoa: dynamics and dimensions*. Singapore: Springer Nature Singapore; 2023. pp. 109–31.
78. Casem ML. *Endocytosis. Case studies in cell biology*. Elsevier; 2016. pp. 217–40.
79. Kaur A, Sharma S, Noor A, Dang S, Talegaonkar S. Potential targeting sites in brain and brain tumors. *Nanocarriers for Drug-Targeting brain tumors*. Elsevier; 2022. pp. 69–120.
80. Modi M, Thambiraja M, Cherukat A, Yennamalli RM, Priyadarshini R. Structure predictions and functional insights into Amidase_3 domain containing N-acetylmuramyl-L-alanine amidases from *Deinococcus indicus* DR1. *BMC Microbiol*. 2024;24:101.
81. Knott V, Downing KA, Cardy CM, Handford P. Calcium binding properties of an epidermal growth Factor-like domain pair from human Fibrillin-1. *J Mol Biol*. 1996;255:22–7.
82. Zhang Y-M, Rock CO. Membrane lipid homeostasis in bacteria. *Nat Rev Microbiol*. 2008;6:222–33.
83. Suleman M, Rashid F, Ali S, Sher H, Luo S, Xie L et al. Immunoinformatics-based design of immune-boosting multi-epitope subunit vaccines against Monkeypox virus and validation through molecular dynamics and immune simulation. *Front Immunol*. 2022;13.
84. Parvizpour S, Pourseif MM, Razmara J, Rafi MA, Omid Y. Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches. *Drug Discov Today*. 2020;25:1034–42.
85. Khan A, Khan S, Saleem S, Nizam-Uddin N, Mohammad A, Khan T, et al. Immunogenomics guided design of Immunomodulatory multi-epitope subunit vaccine against the SARS-CoV-2 new variants, and its validation through in Silico cloning and immune simulation. *Comput Biol Med*. 2021;133:104420.
86. Sami SA, Marma KKS, Mahmud S, Khan MAN, Albogami S, El-Shehawi AM, et al. Designing of a Multi-epitope vaccine against the structural proteins of Marburg virus exploiting the immunoinformatics approach. *ACS Omega*. 2021;6:32043–71.
87. Shah SZ, Jabbar B, Mirza MU, Waqas M, Aziz S, Halim SA, et al. An immunoinformatics approach to design a potent Multi-Epitope vaccine against Asia-1 genotype of Crimean–Congo haemorrhagic fever virus using the structural glycoproteins as a target. *Vaccines (Basel)*. 2022;11:61.
88. Lee SJ, Shin SJ, Lee MH, Lee M-G, Kang TH, Park WS, et al. A potential protein adjuvant derived from *Mycobacterium tuberculosis* Rv0652 enhances dendritic Cells-Based tumor immunotherapy. *PLoS ONE*. 2014;9:e104351.
89. Wang T, Wang Z, Mi J, Wang W, Li K, Qi X, et al. Recombinant avian β -Defensin produced by Food-Grade *Lactococcus* as a novel and potent immunological enhancer adjuvant for avian vaccine. *Probiotics Antimicrob Proteins*. 2021;13:1833–46.
90. Cui B, Liu X, Fang Y, Zhou P, Zhang Y, Wang Y. Flagellin as a vaccine adjuvant. *Expert Rev Vaccines*. 2018;17:335–49.

91. Yu T-W, Chueh H-Y, Tsai C-C, Lin C-T, Qiu JT. Novel GM-CSF-based vaccines: one small step in GM-CSF gene optimization, one giant leap for human vaccines. *Hum Vaccin Immunother*. 2016;12:3020–8.
92. Hasan MS, Agosti JM, Reynolds KK, Tanzman E, Treanor JJ, Evans TG. Granulocyte macrophage Colony-Stimulating factor as an adjuvant for hepatitis B vaccination of healthy adults. *J Infect Dis*. 1999;180:2023–6.
93. Chang DZ, Lomazow W, Joy Somberg C, Stan R, Perales M-A. Granulocyte-Macrophage colony stimulating factor: an adjuvant for Cancer vaccines. *Hematology*. 2004;9:207–15.
94. Javaid N, Choi S. Toll-like receptors from the perspective of Cancer treatment. *Cancers (Basel)*. 2020;12:297.
95. Rhee JH, Khim K, Puth S, Choi Y, Lee SE. Deimmunization of Flagellin adjuvant for clinical application. *Curr Opin Virol*. 2023;60:101330.
96. Song L, Liu G, Umlauf S, Liu X, Li H, Tian H, et al. A rationally designed form of the TLR5 agonist, Flagellin, supports superior immunogenicity of influenza B globular head vaccines. *Vaccine*. 2014;32:4317–23.
97. Ansari H, Raghava GP. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immunome Res*. 2010;6:6.
98. Pinlaor S, Tada-Oikawa S, Hiraku Y, Pinlaor P, Ma N, Sithithaworn P, et al. *Opisthorchis viverrini* antigen induces the expression of Toll-like receptor 2 in macrophage RAW cell line. *Int J Parasitol*. 2005;35:591–6.
99. Ninlawan K, O'Hara SP, Splinter PL, Yongvanit P, Kaewkes S, Surapaitoon A, et al. *Opisthorchis viverrini* excretory/secretory products induce toll-like receptor 4 upregulation and production of Interleukin 6 and 8 in cholangiocyte. *Parasitol Int*. 2010;59:616–21.
100. Bazhan SI, Antonets DV, Karpenko LI, Oreshkova SF, Kaplina ON, Starostina EV, et al. In Silico designed Ebola virus T-Cell Multi-Epitope DNA vaccine constructions are Immunogenic in mice. *Vaccines (Basel)*. 2019;7:34.
101. Hasan M, Azim KF, Begum A, Khan NA, Shammi TS, Imran AS, et al. Vaccinomics strategy for developing a unique multi-epitope monovalent vaccine against Marburg marburgvirus. *Infect Genet Evol*. 2019;70:140–57.
102. Guedes RLM, Rodrigues CMF, Coatnoan N, Cosson A, Cadioli FA, Garcia HA, et al. A comparative in Silico linear B-cell epitope prediction and characterization for South American and African trypanosoma Vivax strains. *Genomics*. 2019;111:407–17.
103. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct*. 2013;8:30.
104. Plewczynski D, Łażniewski M, Augustyniak R, Ginalski K. Can we trust Docking results? Evaluation of seven commonly used programs on PDBbind database. *J Comput Chem*. 2011;32:742–55.
105. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron*. 2018;99:1129–43.

Publisher's note

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