Bioinformatics analysis and immunogenicity assessment of the novel multi-stage DNA vaccine W541 against Mycobacterium tuberculosis

Yourong Yang[1], Yan Liang[1], Yong Xue[1], Xiaoou Wang[2], Lan Wang[1], Jie Wang[3], Junxian Zhang[2], and Xueqiong Wu[1]

[1]Chinese PLA General Hospital
[2]The Eighth Medical Center of Chinese PLA General Hospital
[3]Tuberculosis Prevention and Control Key Laboratory, Beijing Key Laboratory of Ne beijing, Beijing, CN 100091

April 25, 2024

Abstract

Background: Vaccination is one of the effective measures to prevent latent tuberculosis infection (LTBI) from developing into active tuberculosis (TB). Applying bioinformatics methods to pre-evaluate the biological characteristics and immunogenicity of vaccines can improve the efficiency of vaccine development. Objectives: To evaluate the immunogenicity of tuberculosis vaccine W541 and explore the application of bioinformatics technology in tuberculosis vaccine research. Methods: This study concatenated the immunodominant sequences of Ag85A, Ag85B, Rv3407, and Rv1733c to construct the W541 DNA vaccine. Then, bioinformatics methods were used to analyze the physicochemical properties, antigenicity, allergenicity, toxicity, and population coverage of the vaccine, identify its epitopes, and perform molecular docking with MHC alleles and Toll-like receptor 4 (TLR4) of the host. Finally, the immunogenicity of the vaccine was evaluated through animal experiments. Results: the W541 vaccine protein is a soluble cytoplasmic protein with a half-life of 1.1 hours in vivo and an instability index of 45.37. It has good antigenicity and wide population coverage without allergenicity and toxicity. It contains 138 HTL epitopes, 73 CTL epitopes, 8 linear and 14 discontinuous epitopes of B cells, and a strong affinity for TLR4. Immune simulations showed it could effectively stimulate innate and adaptive immune responses. Animal experiments have confirmed that the W541 DNA vaccine could effectively activate the Th1- and Th17-type immune responses, producing high levels of IFN-γ and IL-17A, but could not significantly increase antibody levels. Conclusion: the W541 DNA vaccine can induce strong cellular immune responses. However, further optimization of the vaccine design is needed to make the expressed protein more stable in vivo. Bioinformatics analysis could reveal vaccines' physicochemical and immunological information, which is critical for guiding vaccine design and development.

Bioinformatics analysis and immunogenicity assessment of the novel multi-stage DNA vaccine W541 against *Mycobacterium tuberculosis*

Yourong Yang+,Yan Liang+,Yong Xue, Xiaoou Wang, Lan Wang, Jie Wang, Junxian Zhang, Yinping Liu,Xueqiong Wu*

Beijing Key Laboratory of New Techniques of Tuberculosis Diagnosis and Treatment, Institute of Tuberculosis Research, Senior Department of Tuberculosis, the Eighth Medical Center of PLA General Hospital, Beijing 100091, China

+These authors contributed equally to this study.

*Corresponding author.E-mail:xueqiongwu@139.com.

Running title : Analysis and Verification of DNA vaccine

Abstract

Background: Vaccination is one of the effective measures to prevent latent tuberculosis infection (LTBI) from developing into active tuberculosis (TB). Applying bioinformatics methods to pre-evaluate the biological characteristics and immunogenicity of vaccines can improve the efficiency of vaccine development. **Objectives:** To evaluate the immunogenicity of tuberculosis vaccine W541 and explore the application of bioinformatics technology in tuberculosis vaccine research.**Methods:** This study concatenated the immunodominant sequences of Ag85A, Ag85B, *Rv3407* , and *Rv1733c* to construct the W541 DNA vaccine. Then, bioinformatics methods were used to analyze the physicochemical properties, antigenicity, allergenicity, toxicity, and population coverage of the vaccine, identify its epitopes, and perform molecular docking with MHC alleles and Toll-like receptor 4 (TLR4) of the host. Finally, the immunogenicity of the vaccine was evaluated through animal experiments.**Results:**the W541 vaccine protein is a soluble cytoplasmic protein with a half-life of 1.1 hours in vivo and an instability index of 45.37. It has good antigenicity and wide population coverage without allergenicity and toxicity. It contains 138 HTL epitopes, 73 CTL epitopes, 8 linear and 14 discontinuous epitopes of B cells, and a strong affinity for TLR4. Immune simulations showed it could effectively stimulate innate and adaptive immune responses. Animal experiments have confirmed that the W541 DNA vaccine could effectively activate the Th1- and Th17-type immune responses, producing high levels of IFN-γ and IL-17A, but could not significantly increase antibody levels. **Conclusion:** the W541 DNA vaccine can induce strong cellular immune responses. However, further optimization of the vaccine design is needed to make the expressed protein more stable in vivo. Bioinformatics analysis could reveal vaccines' physicochemical and immunological information, which is critical for guiding vaccine design and development.

Keywords: Tuberculosis; DNA vaccine; Bioinformatic analysis; Simulated immunization; Immunogenicity

1. Introduction

Latent tuberculosis infection (LTBI) is characterized by the presence of specific immune responses to *Mycobacterium tuberculosis*(*M.tb*) previously infected without clinical evidence of active tuberculosis (TB)(1). Currently, about 23% of the world's population is in an LTBI state, in which 5-15% of those with LTBI may develop into active TB in their lifetime; LTBI has become an essential source of active TB(2). According to the 2015 WHO Guidelines for the Management of Latent Tuberculosis Infection, individuals with LTBI can take anti-TB drugs to avoid developing active TB(2). Considering that LTBI has no clinical symptoms, it appears that vaccination-based preventive treatment is more acceptable than chemotherapy. However, the Bacillus Calmette-Guérin (BCG) vaccine, widely used for tuberculosis prevention, has a poor preventive effect on LTBI(3). M72/AS01E, which was in phase IIb clinical trial and developed by GlaxoSmithKline Plc., has only a 54.0% protective efficacy against LTBI developing into active pulmonary TB(4). The phase III clinical trial of the *M.vaccae* vaccine produced by Anhui ZhiFeiLongKeMa Biopharmaceutical Co., Ltd showed a protective efficacy of 54.7% against LTBI(5). These data suggest that developing an effective LTBI preventive and therapeutic vaccine has broad prospects.

According to research reports, the Ag85 complex is the main secretory protein of*M.tb* , consisting of three proteins: Ag85A, Ag85B, and Ag85C. It accounts for 30% of the total secreted protein In the *M.tb H37Rv* strain and can be isolated from early cultures. It has mycobacterial acid transferase activity, allowing trehalose to transfer and deposit on the cell wall of*M.tb* , playing an essential role in the final stage of *M.tb*cell wall synthesis(6, 7). Ag85A and Ag85B contain multiple human T-cell epitopes, and $CD4^+$ T cells from TB patients could respond to the whole Ag85A or Ag85B polypeptides to produce interferon-gamma (IFN-γ)(8). Our previous animal experimental studies(9-13)and the clinical trials reported(8, 13-25) have shown that Ag85A and Ag85B had high immunogenicity, could induce Th1-type responses and cytotoxic T lymphocytes, reduce bacterial loads in lung and other tissues, alleviate lung lesions, and had better protective or therapeutic effects on TB or mouse model with latent tuberculosis infection (LTBI). At present, many new TB vaccines internationally chose Ag85A and/or Ag85B protein as vaccine antigens(15, 16), among which multiple vaccines have entered clinical trials, such as AERAS-402 (including Ag85A, Ag85B, and TB10.4)(17), MVA85A \ Ad5Ag85A \ ChAdOx1 85A (all including Ag85A)(8, 13-20), TB/FLU-04L (including Ag85A

and ESAT6)(21), GamTBvac (including Ag85A, ESAT6, and CFP10)(22), H1/IC31 (including Ag85B and ESAT6)(23), H4: IC31 (including Ag85B and TB10.4) [], H56: IC31 (including Ag85B, ESAT6, and *Rv2660c*)(24), AEC/BC02 (including Ag85B, ESAT6, and CFP10)(25).

Rv3407 is a protein consisting of 99 amino acids, specifically expressed during the *M.tb* transition from dormancy to reactivation. It may be a kind of antitoxin, only slightly expressed in*M.tb* virulent strains and not expressed in BCG strains(26-28). Schuck D et al. have revealed that the *Rv3407* protein could induce abundant IFN-γ and robust Th1-type cell-mediated immune responses in individuals with LTBI and was notably deficient in active TB patients, indicating that the *Rv3407*protein may confer significant protection against dormant *M.tb*infection in susceptible populations(29). Reece et al. engineered the*rv3407* gene into the BCG vaccine and immunized mice with the recombinant BCG vaccine, and found that this modified BCG vaccine stimulated high levels of IFN-γ production in mice and markedly enhanced protection against TB(30). Our research group also found through animal experiments that the mice immunized with the *rv3407* DNA vaccine could produce higher levels of antigen-specific IFN-γ in the culture supernatant of splenic lymphocytes, had more Th1 cells and an increased Th1/Th2 cells ratio in the whole blood, could reduce the bacterial load in the lungs of mouse models with acute infection or LTBI, and alleviate the degree of lung lesions(31, 32).

*Rv1733c*is a major dormancy antigen highly expressed by latent*M.tb*and can be well recognized by T cells from individuals with LTBI(33). Zhang W et al. immunized mice with a DNA vaccine encoding *Rv1733c*and exhibited higher splenocyte stimulation index and IFN-γ, IL-2, and IL-4 levels than those injected with saline(34). Our research group used animal experiments to compare the preventive and therapeutic effects of MTB ag85ab and 7 types of LTBI DNA vaccines on a mouse LTBI model, and it showed that the ag85ab, *rv2659c* , and*rv1733c* DNA vaccines reduced the bacterial load and degree of lung lesions in the mouse LTBI model(32). Additionally, Coppola M et al. immunized mice with synthetic *Rv1733c* long peptides (28 amino acid sequences located at positions 57-84, IPFAAAAGTAVQDSRSHVYAHQAQTRHP) and exhibited significantly increased expression of IFN-γ, TNF-α, and specific antibody, and reduced the pulmonary*M.tb* load. The findings suggest that *Rv1733c* has the potential for the prevention or treatment of TB(33).

Based on this, we chose the full-length amino acid sequence of Ag85A and 308 amino acids of Ag85B as the vaccine backbone, added 51 amino acids of*Rv3407* protein (including 15 amino acids at positions 16-30 and 36 amino acids at positions 61-96) and 28 amino acids of *Rv1733c*protein (located at positions 57-84), and then concatenated them to construct a new TB DNA vaccine, named W541 based on the number of recombinant plasmids been constructed by our research group over the years, aiming to elicit synergistic protective immunity on TB and LTBI.

The development of bioinformatics and the application of big data analytics have provided convenient conditions for the design and development of vaccines, allowing researchers to have the opportunity to understand vaccine-related information in advance, thereby gaining a deeper understanding of vaccine characteristics and making corresponding optimizations to improve the efficiency of vaccine development (35-38). In this study, we used bioinformatics techniques to analyze various physicochemical properties and immunological characteristics of the W541 DNA vaccine. Then, we verified the immunogenicity of the W541 vaccine through animal experiments, exploring the feasibility of employing bioinformatics analysis methods as a means of preliminary assessment during tuberculosis vaccine development to aid vaccine research.

Material and method

The flow chart of the study design was shown in figure 1.

2.1 Bioinformatics analysis

2.1.1Servers and Databases

The bioinformatics analysis servers and databases used in this study are shown in Table 1.

2.1.2 The amino acid sequence of the W541 vaccine protein

The W541 vaccine protein contains immunodominant sequences of four antigens: ag85A, ag85B, *Rv3407* , and *Rv1733c* . The amino acid sequence of the vaccine protein is shown in figure 2.

2.1.3 The physicochemical and basic biological characteristics of the W541vaccine protein

The physicochemical properties of the vaccine proteins, including amino acid composition, molecular weight, theoretical isoelectric point (pI), instability index (II), aliphatic index (AI), and grand average of hydropathicity (GRAVY), were predicted using the ProtParam server(39). The SignaIP 4.1 server (40), MHMM 2.0 server, and Cell-PLoc 2.0 server(41) were used to predict the presence of signal peptide sequences, transmembrane regions, and subcellular localization of the vaccine proteins after expression within host cells, respectively. The antigenicity, allergenicity, and toxicity of the vaccine proteins were predicted using the VaxiJen server(42), AllerTop server(39), and ToxinPred server(43), respectively. Human homology analysis on the vaccine proteins was performed using the BLAST server(44).

2.1.4 The spatial structure analyses of W541 vaccine protein and molecular docking with Toll-like receptor (TLR)4

Using the SOPMA serve predicted the secondary structure of the W541 vaccine protein (45). Utilize the bKunyun supercomputing platform to execute the AlphaFold2(46)to predict the tertiary structure of W541, which was then authenticated using the Prosa server(47).

The GRAMM server can systematically assess a series of docking postures between proteins and ligands and predict the most stable docking conformation(48). Firstly, the PDB files of TLR4 were downloaded from the PDB database(49) and uploaded to the GRAMM server along with the W541 PDB files predicted by AlphaFold to predict the docking status between the TLR4 and W541 vaccine. Finally, the PDBePISA server (50) was used to calculate the detailed docking data of the docking complexes, such as the interaction surface areas and binding energies.

2.1.5 Prediction of B-cell epitopes in the W541 vaccineprotein

The linear and discontinuous B-cell epitopes of W541 vaccine proteins were predicted using the ElliPro server(51) with default parameters by uploading the AlphaFold-predicted W541 PDB file.

2.1.6 Prediction of HTL and CTL epitopes in the W541 vaccine protein

The helper T lymphocytes (HTL) epitopes were predicted by the MHC-II Binding Predictions server(52) using the "IEDB recommended method" specified for the Full HLA reference set with default parameters, which can predict all the 15 amino acid residues in the W541 vaccine protein. Later the epitopes(excluding epitopes containing linker amino acids) with IC50 values below 500 nM were further analyzed by the VaxiJen server with a threshold of 0.4, the IFN epitope server(53), IL4pred server (54), IL6pred server(55), and IL-10pred server (56) to predict their antigenicity, the abilities to stimulate IFN-γ, IL-4, IL-6, and IL-10 secretion with default parameters.

The cytotoxic T lymphocytes (CTL) epitopes were predicted by the MHC-I Binding Predictions server(57) using the "recommended epitope predictor" specified for the default HLA allele reference set, which can predict all the 9 and 10 amino acid residues CTL epitopes of the vaccine. Further, the epitopes (excluding epitopes containing linker amino acids) with IC50 values below 500 nM were analyzed using the Class I Immunogenicity server with default parameters to predict their immunogenicity (58)and using the VaxiJen server to predict their antigenicity with default parameters.

At last, all the HTL and CTL epitopes of the W541 vaccine protein were submitted to AllerTOP v. 2.0 serve and ToxinPred serve to predict their allergenicity and toxicity with default parameters.

2.1.7 Analysis of population coverage and molecular docking of T-cell epitopes with MHC molecules for the W541 vaccine

The population coverage of the vaccine was predicted by the IEDB Population Coverage tool(59). The molecular docking between the vaccine epitopes and MHC molecules was performed using the GRAMM

docking server. Firstly, the PDB files of the MHC molecules were downloaded from the RCSB PDB database and processed with Pymol software 2.0 (an open-source tool) to remove unnecessary ligands. Meanwhile, the structures of all docking epitopes were predicted using the PEP-FOLD 3.5 server (60). The PDB files of MHC molecules and corresponding epitope structures were then submitted to the GRAMM docking server using the "free docking" mode. The docking results were analyzed using the PDBePISA server.

2.1.8 Immune simulation in silico

The immune responses to the W541 vaccine were simulated using the C-ImmSim server(61). In this study, we performed three rounds of in silico immunization with the W541 vaccine at 0, 14, and 28 days to simulate the immune effects of the W541 vaccine.

2.2 Experiment validation

2.2.1Preparation and characterization of W541 DNA vaccine and its corresponding recombinant protein W540.

According to the genetic code rules, the amino acid sequence of the W541 vaccine protein was translated into a DNA sequence. The DNA sequence was then codon-optimized to generate sequences specifically suited for eukaryotic and prokaryotic expression. The optimized eukaryotic expression DNA sequence and the pVAX1 vector plasmid were digested with restriction enzymes NheI and EcoRI, respectively. The two fragments were ligated using DNA ligase to construct the recombinant plasmid. The recombinant plasmid was then transformed into Escherichia coli (*E. coli*) DH5α competent cells. Subsequently, the transformed *E. coli* DH5α cells were cultured in the LB medium. After culturing, the W541 recombinant plasmid was isolated and purified using the plasmid extraction kit (Qiagen, Germany). Similarly, the optimized prokaryotic expression DNA sequence and the pET28a vector plasmid were digested with restriction enzymes NheI and EcoRI, respectively. Then, the obtained fragments were ligated using DNA ligase to construct the W540 recombinant protein plasmid and transformed into *E. coli*BL21(DE3) competent cells. The transformed E.coli BL21(DE3) cells were cultured on LB medium and induced by 0.1 mM IPTG at 37 for 3 hours before purifying the recombinant fusion protein(62). The W541 recombinant plasmid was identified by sequencing and enzyme digestion. The W540 recombinant plasmid was identified by sequencing, and the protein expression of transformed *E.coli* BL21(DE3)was identified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (63).

2.2.2 Experimental animals

Female BALB/c mice aged 56-62 days and weighing 18-20g were obtained from Vital River Laboratory Animal Technology Co. Ltd (Beijing, China). The animal experiments were conducted following the Regulations on Management of Experimental Animals formulated by the Ministry of Science and Technology of China. The Animal Ethics Committee of the Eighth Medical Center of PLA General Hospital has approved the experimental plan.

2.2.3 Immunization of the Animals

Mice were divided into three groups (8 mice per group) and treated as follows: 1) Control group: each mouse was injected intramuscularly with 100 μl of saline; 2) pVAX1 group: each mouse was injected intramuscularly with 100 μg pVAX1 plasmid diluted in 100 μl saline; 3) Vaccine group: each mouse was injected intramuscularly with 100 μg W541 DNA vaccine diluted in 100 μl saline (32). Mice received injections every two weeks for a total of 3 injections. At five weeks after the last immunization, the mice were euthanized. Blood samples were collected in heparin lithium anticoagulant tubes, and then the plasma was separated. The spleens of the mice were also harvested.

2.2.4 Detection of specific antibodies against W540

The blood sample of each mouse was separated from the plasma and stored at -20. The enzyme-linked immunosorbent assay (ELISA) was employed to individually detect specific antibodies IgG and its subtypes IgG1 and IgG2a against W540 protein in the plasma samples obtained from all 24 mice in 3 groups. The

method is briefly described as follows: The ELISA plate (Costar, China) was coated with 100 μl of 10 μg/ml W540 protein in the coating buffer (0.05M $NaHCO_3$, pH 9.6) and incubated overnight at 4. The plate was then washed with PBS-T (1×Phosphate-Buffered Saline, pH7.4, containing 0.05% Tween-20) and blocked with 200 μl/well of 2% (w/v) ovalbumin (Coolaber, China) in PBS buffer and incubated for 3 hours at 37. Plasma samples were diluted 1:100 in assay buffer (1% w/v BSA in PBS-T) and added in duplicate at 100 μl/well, followed by incubation for 2 hours at 37. The coating buffer without W540 protein was added to each plate as a negative control. Goat anti-mouse IgG, IgG1, and IgG2a conjugated to horseradish peroxidase (HRP) (Abcam, England) diluted 1:10,000 in assay buffer was added at 100 μl/well and incubated for 1 hour at 37. After washing, 100 μl of 3,3',5,5'-Tetramethylbenzidine(TMB) substrate (BD, America) was added to each well, and plates were incubated in the dark at room temperature for 5 minutes. The enzymatic reaction was stopped by adding 50 μl of 1M H_2SO_4, and then the absorbance at 450nm was measured using a microplate reader (Thermo, America).

2.2.5 Δετεςτ τηε νυμβερ οφ μουσε σπλενοςψτες σεςρετινγ ΙΦΝ-γ βψ ενζψμε-λινκεδ Ιμμυνοσποτ ασσαψ (ΕΛΙΣΠΟΤ)

Equal amounts of mice spleen cells from the same group were mixed to reduce individual variations, and 3 independent experiments were subsequently conducted with the Mouse IFN-γ ELISpotPLUS (HRP) assay kit (MABTECH AB, Sweden)to assess the immune response of mouse splenocytes to the W540 protein. The detailed process follows: 3×10^5mixed cells were seeded in each well of the filter membrane plate pre-packaged with IFN-γ capturing antibody (MABTECH AB, Sweden). The negative control wells were added with RPMI 1640 complete medium, while the positive control wells were added phytohemagglutinin (PHA) (Sigma, USA) at a final concentration of 20mg/ml. Following the kit's instructions, the number of splenocytes secreting IFN-γ was measured after stimulation with W540 protein at a concentration of 30μg/ml in a CO_2 incubator at 37 for 20 hours. The number of spots in each well was detected using CTL ImmunoSpot®S5 Micro equipment (Cellular Technology, America).

2.2.6 Cytokine analysis

Spleen cell suspensions from each group of mice were mixed in equal proportion to reduce individual variations, and then 3 repeated experiments were conducted. The mixed spleen cell concentration was adjusted to 3×10^6 cells/ml. After stimulation with W540 protein at a final concentration of 30μg/ml for 24h, the supernatant from spleen cell cultures was subjected to analyze the expression levels of cytokine IFN-γ, IL-2, IL-4, IL-6, IL-10, and IL-17A using the CBA assay kit (BD Biosciences, USA) following the instructions provided. The measurements were performed in triplicate.

2.2.7 Statistical analysis

The antibody detection, ELISPOT assay, and cytokine detection data were analyzed using GraphPad Prism 8 software. Quantitative data of normal distribution were analyzed using one-way analysis of variance (ANOVA), and pairwise comparisons were conducted using Tukey' s multiple comparisons test. The results were presented as mean±standard error, and a p-value of less than 0.05 was considered a statistically significant difference.

3 Results

3.1 Bioinformatics analysis results

3.1.1 Physicochemical characteristics and basic biological characteristics of W541vaccine protein

The W541 vaccine protein sequence contains 704 amino acid residues with a molecular weight of 74.6 kDa and a theoretical isoelectric point of 6.29. It exhibits an average hydrophobicity of -0.294, an aliphatic index of 68.85, an instability index of 45.37, and possesses 57 phosphorylation sites and 7 glycosylation sites. The expressed vaccine protein lacks signal peptide and transmembrane domains. Upon expression in human cells, it localizes to lysosomes and the cytoplasm. It had a half-life of 1.1 hours in mammalian reticulocytes, 3

minutes in yeast, and 2 minutes in *E. coli* . These data indicate that the W541 vaccine protein is a relatively large soluble protein prone to degradation. The W541 vaccine protein also displays commendable antigenicity, non-allergenicity, non-toxicity, and no homology with human proteins (Table 2).

3.1.2 The spatial structures of the W541 vaccine protein and docking with TLR4

According to the analysis conducted by the SOPMA server, the secondary structure of the W541 protein exhibited a distribution of α-helices, β-sheet, β-turns, and random coils, accounting for 26.99%, 19.03%, 11.51%, and 42.47% of the total sequence, respectively. These four secondary structures were arranged alternately throughout the overall structure but had one distinct feature: the α-helices primarily exist in the central and posterior parts of the protein (figure 3A). Due to their inherently flexible nature, the random coils and β-turns tend to be situated on the protein surface, showcasing a prominent structure. This region is usually enriched with epitopes that are advantageous for B-cell recognition.

For the W541 vaccine protein, AlphaFold2 predicted five possible tertiary structures exhibiting high similarities. Both Ag85A and Ag85B sequences were folded separately, while the three C-terminal epitopes exist independently as α-helices and random coils. The Prosa server scored respectively the five predicted structures as -10.99, -10.96, -10.65, -10.60, and -10.97, which fall within the expected range of the server's protein model database. This result indicates a high reliability of the predicted results (figure 3B).

The top-ranking W541 vaccine protein A structure was selected to study the interaction with TLR4 to evaluate the immune system activation effect of the vaccine. The central region of the W541 vaccine protein showed a remarkable docking affinity with the B chain of TLR4, with a binding energy of -10.7 kcal/mol and an interface area of 2209.7Å^2 (figure 3C). The lower the binding energy indicates the closer vaccine-TLR binding, the more likely it is to activate TLR4 and induce the production of cytokines and chemokines to enhance immune responses.

3.1.3 B-cell epitopes of W541 vaccine protein

B-cell epitopes can bind to B cell receptors (BCRs), activating downstream signaling cascades that regulate B-cell activation and antibody production. Activated B cells can also act as antigen-presenting cells. The prediction results from the ElliPro server showed that the W541 vaccine protein had 15 linear epitopes with prediction scores ranging from 0.511 to 0.755 and 11 discontinuous epitopes with predicted scores ranging from 0.543 to 0.987 (shown in Supplementary Table 1).

3.1.4 HTL and CTL epitopes of W541 vaccine protein

Based on the selection criteria, the W541 vaccine protein contained a total of 138 HTL epitopes (Supplementary Table 2) and 73 CTL epitopes (Supplementary Table 3), in which 101, 63, 3, and 19 HTL epitopes could induce IFN-γ, IL-4, IL-6, and IL-10, respectively. The W541 protein contained 10 HTL toxic epitopes and 6 CTL toxic epitopes. These toxic epitopes were concentrated within two discreet regions of the amino acid chain, totaling 21 residues (Supplementary Table 4).

3.1.5 Population coverage analysis and molecular docking of T lymphocyte epitopes with MHC for W541 vaccine protein

The population coverage analyses on the HTL and CTL epitopes of the W541 vaccine protein and their corresponding MHCs were performed using the Population Coverage server. The results revealed that the population coverage rates of HTL and CTL epitopes were 99.68% and 98.36%, respectively. 138 HTL epitopes that stimulate IFN-γ secretion (Supplementary Table 2) and 73 CTL epitopes (Supplementary Table 3) could recognize 24 MHC class II molecules and 26 MHC class I molecules, respectively. Some of these epitopes had a binding affinity for a single MHC molecule, while others could bind to multiple MHC molecules. From this epitope library, 18 HTL epitopes (antigenicity threshold > 0.8, inducing only IFN, IC50 < 500) and corresponding MHC molecules are displayed in Table 3A; 27 CTL epitopes (antigenicity threshold > 0.8, class I immunogenicity > 0.1, IC50 < 500) and corresponding MHC molecules are shown in

Table 3B. Molecular docking results of two selected T-cell epitopes with their corresponding MHC molecules are shown in figure 4.

3.1. 6 Immune simulation in silico

3.1.6.1 Activation of innate immune cells by

the W541 DNA vaccine

Macrophages and dendritic cells are crucial to the body's innate immune responses. The C-ImmSim server provides a simulation immune platform for evaluating the immune response after vaccination. After the first immunization, the number of activated macrophages significantly increased. Each subsequent injection further enhanced macrophage activation, and the macrophage count declined around day 50 until it returned to pre-immunization levels on day 70 (figure 5A). Dendritic cells were activated after the first immunization, and the number of dendritic cells in the presenting state increased after each immunization until it returned to pre-immunization levels on day 90 (figure 5B). The simulation results indicate a significant activation of macrophages and dendritic cells after three immunizations of the W541 DNA vaccine.

3.1.6.2 Activation of adaptive immune cells by theW541 DNA vaccine

The adaptive immune cells mainly include TH, TC, and B cells, which play a crucial role in anti-TB immunity. After the initial vaccination, TH, TC, and B cells were activated. With each subsequent immunization, their activation was enhanced and reached their peaks on day 35, day 50, and day 30, respectively. It is worth noting that the activation levels of the TH and B cells decreased before the subsequent immunization, while the TC cells remained in a sustained activated state (figure 5C-E). Following the first vaccination, IFN-γ levels rapidly increased. With the second immunization, IFN-γ levels were further elevated. After the third immunization, the IFN-γ levels remained relatively stable compared to the second immunization and then rapidly declined, eventually returning to pre-immunization levels around day 50 (figure 5F). A slight increase in IgG and IgM was observed after the initial immunization. After the second immunization, the expression of IgG and IgM increased rapidly and further increased after the third immunization. IgM peaked around day 40, followed by a decline, and dropped to the plateau phase around day 90. IgG peaked around day 50, followed by a similar decline, eventually entering a steady phase around day 150 (figure 5G). Simulation results demonstrated that the W541 DNA vaccine could effectively activate TH, TC, and B cells, leading to elevated levels of corresponding cytokines such as IFN-γ and antibodies IgG and IgM.

3.2 Experimental validation results

3.2.1 Identification of W541 DNA vaccine and W540 recombinant protein

The sequencing findings indicate that the DNA sequences inserted in the W541 vaccine and W540 plasmid are congruent with the design. Upon induction with IPTG, the transformed *E. coli* BL21(DE3) with W540 plasmid expressed protein with an approximate molecular weight of 70kDa on SDS-PAGE (figure 6).

3.2.2 Antibody production induced by W541 DNA vaccine

After three immunizations, there was no significant difference in the antibody IgG, IgG1, and IgG2a levels among the W541 vaccine group, Control group, and pVAX1 group ($P>0.05$) (figure 7).

3.2.3Δετεςτ νυμβερ οφ σπλενοςψτεσσεςρετινγ ΙΦΝ-γ υσινγ ΕΛΙΣΠΟΤ ασσαψ

The ELISPOT assay results showed that the number of spots produced by splenocytes secreting IFN-γ in the vaccine group (26±10) was significantly higher than that in the normal control group (1.2±0.6, $P<0.01$) and the pVAX1 vector control group (2±1, $P<0.01$) (figure 8). These findings indicate that the W541 DNA vaccine effectively activated the splenic lymphocytes in mice; upon stimulation with the W540 protein, the activated splenocytes exhibited a rapid and powerful secondary immune response.

3.2.4 Analyses of Th1, Th2, and Th17 cytokines in splenocyte culture supernatants of mice in each group

For Th1-type and Th17-type cytokines, the IFN-γ and IL17 levels were significantly higher in the supernatants of splenocyte cultures from mice immunized with the W541 DNA vaccine compared to the normal control and pVAX1 vector groups ($P<0.05$, figure 9A&F). However, there were no significant differences in IL-2 levels among the vaccine and control groups ($P>0.05$, figure 9B). For Th2-type cytokines, comparing the W541 vaccine group to the normal control group revealed that the W541 vaccine decreased expression of IL-4 and IL-6 by murine splenocytes ($P<0.05$, figure 9C&D). There were no significant differences in IL-10 expression among the three groups($P>0.05$, figure 9E). These results demonstrate that the W541 DNA vaccine effectively activated a Th1-type immune response in the host.

4. Discussion

In recent years, the rapid development of bioinformatics methods and the application of Internet databases have provided practical approaches for selecting protein-dominant epitopes and the pre-evaluation of vaccines (64-66), accelerating the vaccine research and development process. In this study, we employed bioinformatic methods to analyze the physicochemical properties, structure, safety, and immunological functions of a multi-stage DNA vaccine W541 constructed by tandem immunodominant sequences of ag85A, ag85B, *Rv1733c* , and*Rv3407* antigens. Furthermore, we evaluated its immunogenicity through a murine model to verify the results of bioinformatics analysis and lay a foundation for further optimization of the vaccine.

The bioinformatic analysis revealed that the W541 vaccine protein was a soluble protein with a molecular weight of 74 kDa, which exhibited excellent antigenicity and broad population coverage without allergenicity or toxicity. These fundamental data strongly support the suitability of W541 as a vaccine candidate based on the essential criteria.

There is a close relationship between the secondary structure of a protein and B-cell epitopes. The structures of irregular coiling and β-turns, being more loosely arranged, are prone to distortion and spiralization, leading to their exposure to the protein surface. These regions typically harbor a greater abundance of B-cell epitopes. The predictive analysis revealed that within the W541 vaccine protein, the sequence length of irregular coiling and β-turns constitutes 53.98% of the total sequence length, most located on the surface of protein molecules. Furthermore, the protein harbors 8 linear B-cell epitopes and 14 discontinuous epitopes. These structures and epitopes provide the foundation for eliciting effective humoral immune responses in the host. The immune stimulation in silico revealed that the W541 vaccine protein could induce a heightened antibody response. However, in animal experiments, most mice in the W541 DNA vaccine group did not exhibit a significant increase in antibody levels. This phenomenon may be attributed to the subcellular localization (in the cytoplasm and lysosomes) and poor stability of the W541 protein, which consequently affects the vaccine's efficacy in eliciting a robust humoral immune response. The W541 vaccine protein expressed in the cytoplasm and lysosomes lost the opportunity to bind to BCR. In addition, it had a short half-life and poor stability, indicating that the vaccine protein is prone to degradation, leading to loss of conformational epitopes, which is detrimental to B cell activation and antibody production(65). Emerging evidence suggests that B cells and humoral immunity can regulate various immune responses of the intracellular pathogen *M. tuberculosis* (67). In addition, unstable proteins may undergo degradation before being captured by macrophages and DC cells, compromising their antigen presentation effectiveness. Ultimately, this could also impact the activation of TH cells (66). Therefore, optimizing the vaccine by increasing signal peptides to alter the subcellular localization of vaccine proteins or optimizing sequences to enhance their stability may help improve the immunogenicity of the vaccine (68-70). Studies have shown that lysine residues exposed on protein surfaces can bind to ubiquitin and promote protein degradation by proteases (71, 72). Additionally, the glycine, lysine, arginine, and cysteine residues at the N-terminus and middle portion of proteins have marked effects on protein stability (73). In the future, we will focus on fine-tuning the structure of the W541 vaccine to enhance the immunogenicity of the vaccine protein.

TLRs play a crucial role in activating the anti-TB immune responses, and TLR agonists are considered a promising class of vaccine adjuvants. TLR4 is an essential receptor involved in MTB recognition within cells, expressed in both immune and non-immune cells, and its structure includes the extracellular domain, transmembrane domain, and intracellular domain(74). TLR4 forms homodimers to recognize pathogen-associated

molecular patterns such as lipopolysaccharides (LPS), lipoteichoic acid (LTA), dsRNA, etc. After TLR4 activation, it can initiate innate immune responses and regulate the migration, maturation, and function of antigen-presenting cells while facilitating adaptive immune responses(75, 76).*M.tb*dormancy-related proteins *Rv2659c* and *Rv1738* can mediate the production of inflammatory cytokines through the TLR4 pathways(77). Furthermore, the activated TLR4 can limit the survival of *M.tb* by inducing cellular autophagy(78). These studies demonstrate the beneficial inclusion of TLR4 agonists in the design of TB vaccines. The relatively stable binding between TLR and ligand is fundamental for TLR activation. Molecular docking simulations in this study have demonstrated that the expressed W541 protein in cells can establish stable interactions with TLR4, thereby possessing the potential to activate TLRs. Simulation immune results also proved that the W541 vaccine can effectively trigger the body's innate immune responses.

The prominent immune characteristics of TB patients include impaired Th1 cell-mediated immune function or imbalanced Th1/Th2 cell immune responses, which represent the primary risk factors of TB(79, 80). Hence, effectively activating the host's Th1-type cell-mediated immune response is essential for TB vaccines to exert their protective effects against *M.tb* (81). IFN-γ plays a crucial role in the defense against *M.tb* infection by promoting the proliferation and differentiation of Th0 cells into Th1 cells and activating macrophages(82). The bioinformatics analysis of this study showed that the W541 vaccine protein was an antigen with a majority of T-cell epitopes, containing a total of 138 HTL epitopes, in which 101 HTL epitopes could induce the production of IFN-γ. In addition, it also included a smaller number of HTL epitopes that could induce the production of IL-4, IL-6, and IL-10. The simulated immunization with the W541 vaccine demonstrated that the W541 DNA vaccine could effectively activate TH cells and elicit a robust release of IFN-γ and small amounts of TGF-β, IL-10, and IL-12 brief secretion. In contrast, the secretion of other cytokines (including IL-4 and IL-6) was not observed. Animal experiments have confirmed that mice in the vaccine group exhibited significantly higher levels of IFN-γ secretion in spleen cells compared to the control group, and there was no significant increase in the secretion levels of Th2-type cytokines (IL-4, IL-6, and IL-10). These results collectively validate the consistency between most of the immunoinformatics analysis results and animal experimentation results of the vaccine, demonstrating a favorable structure-function relationship.IL2 is also a representative multifunctional cytokine mainly secreted by $CD4^+$ T cells in the Th1-type immunity, which can activate T cells and promote cytokine production, activate macrophages, enhance the killing activity of NK cells, and promote the production of immunoglobulins by B cells, playing an essential role in the body's anti-TB immunity. In this study, immunoinformatics prediction showed a significant increase in IL2, but animal experiments exhibited that the W541 immunization could not effectively induce the production of IL2. The possible reasons for this are: (1) W541 immunization expressed high levels of antigens in mice, mainly inducing effector memory T cells (TEM) and producing predominant high IFN-γ/ low IL-2 reaction. After weakened vaccine expression, it may mainly induce central memory T cells (TCM), possibly with a predominant low IFN- γ/ high IL2 reaction. (2) It may also be because mouse spleen cells were stimulated by the specific antigen W540 for a shorter time (24 hours), which generally requires culture for 72 hours to induce high levels of IL2 production(83). In addition, the experiment results showed that the level of IL-17A in the vaccine group was significantly higher compared to the control group. IL-17A was produced by activated T cells and mediated the production of inflammatory molecules, chemokines, antimicrobial peptides, and remodeling proteins (84), playing an essential role in the immune response to *M.tb* (85). The W541 vaccine protein contains 138 CTL epitopes, which can effectively activate cytotoxic T cells, confirmed in immune simulations. Cytotoxic T cells are crucial in clearing*M.tb*infection by promoting target cell apoptosis or clearing infected target cells and persisting bacteria through the perforin-granzyme pathway(86-89). Numerous HTL and CTL epitopes within the W541 vaccine protein have been experimentally validated. For instance, KLIANNTRV has been identified as an HLA-A2-specific $CD8^+$ immunodominant antigen peptide (90), PBMCs from TB patients exhibit a strongly proliferative response to a peptide (DQSGLSVVMPVGGQSSFY) derived from Ag85(91).

5. Conclusions

In summary, although the results of bioinformatics analysis may not encompass all biological effects elicited by vaccines in vivo, the results of the existing analysis showed a strong positive correlation with animal

experiments. The research findings indicate that the W541 DNA vaccine composed of antigen Ag85A, Ag85B, *Rv3407* , and*Rv1733c* sequences contains a large number of HTL and CTL epitopes, which can activate TH cells and TC cells, mainly inducing Th1 and Th17 immune responses of the body. However, the protein expressed by the vaccine in vivo was not stable and could not effectively induce humoral immune responses. Therefore, we will optimize the vaccine design to address the issues with W541 and further evaluate the immunogenicity and protective efficacy of the vaccine in a mouse LTBI model.

Funding: This study was supported by a grant from the National Key R&D Program of China (2022YFA1303500-003).

Institutional Review Board Statement: The animal study was reviewed and approved by The Animal Ethical Committee of the Eighth Medical Center of the Chinese PLA General Hospital (Approved Number: 309202108250973).

Data Availability Statement: The original contributions presented in the study are included in the article and Supplementary Material. Further inquiries can be directed to the corresponding author.

Conflict of interest: The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Getahun H, Matteelli A, Abubakar I, Aziz MA, Baddeley A, Barreira D, et al. Management of latent Mycobacterium tuberculosis infection: WHO guidelines for low tuberculosis burden countries. The European respiratory journal. 2015;46(6):1563-76.

2. Houben RM, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. PLoS medicine. 2016;13(10):e1002152.

3. Alvarez AH, Flores-Valdez MA. Can immunization with Bacillus Calmette-Guérin be improved for prevention or therapy and elimination of chronic Mycobacterium tuberculosis infection? Expert review of vaccines. 2019;18(12):1219-27.

4. Van Der Meeren O, Hatherill M, Nduba V, Wilkinson RJ, Muyoyeta M, Van Brakel E, et al. Phase 2b Controlled Trial of M72/AS01(E) Vaccine to Prevent Tuberculosis. The New England journal of medicine. 2018;379(17):1621-34.

5. Organization WH. Global tuberculosis report 2019. 2019:179.

6. Tang X, Deng W, Xie J. Novel insights into Mycobacterium antigen Ag85 biology and implications in countermeasures for M. tuberculosis. Critical reviews in eukaryotic gene expression. 2012;22(3):179-87.

7. Karbalaei Zadeh Babaki M, Soleimanpour S, Rezaee SA. Antigen 85 complex as a powerful Mycobacterium tuberculosis immunogene: Biology, immune-pathogenicity, applications in diagnosis, and vaccine design. Microbial pathogenesis. 2017;112:20-9.

8. Huygen K. The Immunodominant T-Cell Epitopes of the Mycolyl-Transferases of the Antigen 85 Complex of M. tuberculosis. Frontiers in immunology. 2014;5:321.

9. Liang Y, Wu X, Zhang J, Li N, Yu Q, Yang Y, et al. The treatment of mice infected with multidrug-resistant Mycobacterium tuberculosis using DNA vaccines or in combination with rifampin. Vaccine. 2008;26(35):4536-40.

10. Liang Y, Wu X, Zhang J, Xiao L, Yang Y, Bai X, et al. Immunogenicity and therapeutic effects of Ag85A/B chimeric DNA vaccine in mice infected with Mycobacterium tuberculosis. FEMS immunology and medical microbiology. 2012;66(3):419-26.

11. Liang Y, Bai X, Zhang J, Song J, Yang Y, Yu Q, et al. Ag85A/ESAT-6 chimeric DNA vaccine induces an adverse response in tuberculosis-infected mice. Molecular medicine reports. 2016;14(2):1146-52.

12. Liang Y, Zhang J, Yang Y, Bai X, Yu Q, Li N, et al. Immunogenicity and therapeutic effects of recombinant Ag85AB fusion protein vaccines in mice infected with Mycobacterium tuberculosis. Vaccine. 2017;35(32):3995-4001.

13. Liang Y, Cui L, Xiao L, Liu X, Yang Y, Ling Y, et al. Immunotherapeutic Effects of Different Doses of Mycobacterium tuberculosis ag85a/b DNA Vaccine Delivered by Electroporation. Frontiers in immunology. 2022;13:876579.

14. Wang C, Fu R, Chen Z, Tan K, Chen L, Teng X, et al. Immunogenicity and protective efficacy of a novel recombinant BCG strain overexpressing antigens Ag85A and Ag85B. Clinical & developmental immunology. 2012;2012:563838.

15. Li W, Deng G, Li M, Zeng J, Zhao L, Liu X, et al. A recombinant adenovirus expressing CFP10, ESAT6, Ag85A and Ag85B of Mycobacterium tuberculosis elicits strong antigen-specific immune responses in mice. Molecular immunology. 2014;62(1):86-95.

16. Gupta SK, Wilson T, Maclean PH, Rehm BHA, Heiser A, Buddle BM, et al. Mycobacterium avium subsp. paratuberculosis antigens induce cellular immune responses in cattle without causing reactivity to tuberculin in the tuberculosis skin test. Frontiers in immunology. 2022;13:1087015.

17. Sivakumaran D, Blatner G, Bakken R, Hokey D, Ritz C, Jenum S, et al. A 2-Dose AERAS-402 Regimen Boosts CD8(+) Polyfunctionality in HIV-Negative, BCG-Vaccinated Recipients. Frontiers in immunology. 2021;12:673532.

18. Nemes E, Hesseling AC, Tameris M, Mauff K, Downing K, Mulenga H, et al. Safety and Immunogenicity of Newborn MVA85A Vaccination and Selective, Delayed Bacille Calmette-Guerin for Infants of Human Immunodeficiency Virus-Infected Mothers: A Phase 2 Randomized, Controlled Trial. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2018;66(4):554-63.

19. Smaill F, Xing Z. Human type 5 adenovirus-based tuberculosis vaccine: is the respiratory route of delivery the future? Expert review of vaccines. 2014;13(8):927-30.

20. Wajja A, Nassanga B, Natukunda A, Serubanja J, Tumusiime J, Akurut H, et al. Safety and immunogenicity of ChAdOx1 85A prime followed by MVA85A boost compared with BCG revaccination among Ugandan adolescents who received BCG at birth: a randomised, open-label trial. The Lancet Infectious diseases. 2023.

21. Shurygina AP, Zabolotnykh N, Vinogradova T, Khairullin B, Kassenov M, Nurpeisova A, et al. Preclinical Evaluation of TB/FLU-04L-An Intranasal Influenza Vector-Based Boost Vaccine against Tuberculosis. International journal of molecular sciences. 2023;24(8).

22. Tkachuk AP, Bykonia EN, Popova LI, Kleymenov DA, Semashko MA, Chulanov VP, et al. Safety and Immunogenicity of the GamTBvac, the Recombinant Subunit Tuberculosis Vaccine Candidate: A Phase II, Multi-Center, Double-Blind, Randomized, Placebo-Controlled Study. Vaccines. 2020;8(4).

23. Hussein J, Zewdie M, Yamuah L, Bedru A, Abebe M, Dagnew AF, et al. A phase I, open-label trial on the safety and immunogenicity of the adjuvanted tuberculosis subunit vaccine H1/IC31® in people living in a TB-endemic area. Trials. 2018;19(1):24.

24. Bekker LG, Dintwe O, Fiore-Gartland A, Middelkoop K, Hutter J, Williams A, et al. A phase 1b randomized study of the safety and immunological responses to vaccination with H4:IC31, H56:IC31, and BCG revaccination in Mycobacterium tuberculosis-uninfected adolescents in Cape Town, South Africa. EClinicalMedicine. 2020;21:100313.

25. Guo X, Lu J, Li J, Du W, Shen X, Su C, et al. The Subunit AEC/BC02 Vaccine Combined with Antibiotics Provides Protection in Mycobacterium tuberculosis-Infected Guinea Pigs. Vaccines. 2022;10(12).

26. Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, et al. Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional

genomics of microbial pathogens. Molecular microbiology. 1999;33(6):1103-17.

27. Downing KJ, Betts JC, Young DI, McAdam RA, Kelly F, Young M, et al. Global expression profiling of strains harbouring null mutations reveals that the five rpf-like genes of Mycobacterium tuberculosis show functional redundancy. Tuberculosis (Edinburgh, Scotland). 2004;84(3-4):167-79.

28. Mattow J, Jungblut PR, Schaible UE, Mollenkopf HJ, Lamer S, Zimny-Arndt U, et al. Identification of proteins from Mycobacterium tuberculosis missing in attenuated Mycobacterium bovis BCG strains. Electrophoresis. 2001;22(14):2936-46.

29. Schuck SD, Mueller H, Kunitz F, Neher A, Hoffmann H, Franken KL, et al. Identification of T-cell antigens specific for latent mycobacterium tuberculosis infection. PloS one. 2009;4(5):e5590.

30. Reece ST, Nasser-Eddine A, Dietrich J, Stein M, Zedler U, Schommer-Leitner S, et al. Improved long-term protection against Mycobacterium tuberculosis Beijing/W in mice after intra-dermal inoculation of recombinant BCG expressing latency associated antigens. Vaccine. 2011;29(47):8740-4.

31. Liang Y, Zhao Y, Bai X, Xiao L, Yang Y, Zhang J, et al. Immunotherapeutic effects of Mycobacterium tuberculosis rv3407 DNA vaccine in mice. Autoimmunity. 2018;51(8):417-22.

32. Liang Y, Li X, Yang Y, Xiao L, Liang Y, Mi J, et al. Preventive effects of Mycobacterium tuberculosis DNA vaccines on the mouse model with latent tuberculosis infection. Frontiers in immunology. 2023;14:1110843.

33. Coppola M, van den Eeden SJ, Wilson L, Franken KL, Ottenhoff TH, Geluk A. Synthetic Long Peptide Derived from Mycobacterium tuberculosis Latency Antigen Rv1733c Protects against Tuberculosis. Clinical and vaccine immunology : CVI. 2015;22(9):1060-9.

34. Zhang W, Jiang H, Bai YL, Kang J, Xu ZK, Wang LM. Construction and immunogenicity of the DNA vaccine of Mycobacterium Tuberculosis dormancy antigen rv1733c. Scandinavian journal of immunology. 2014;79(5):292-8.

35. De Groot AS, Moise L, Terry F, Gutierrez AH, Hindocha P, Richard G, et al. Better Epitope Discovery, Precision Immune Engineering, and Accelerated Vaccine Design Using Immunoinformatics Tools. Frontiers in immunology. 2020;11:442.

36. Fleri W, Paul S, Dhanda SK, Mahajan S, Xu X, Peters B, et al. The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design. Frontiers in immunology. 2017;8:278.

37. Oli AN, Obialor WO, Ifeanyichukwu MO, Odimegwu DC, Okoyeh JN, Emechebe GO, et al. Immunoinformatics and Vaccine Development: An Overview. ImmunoTargets and therapy. 2020;9:13-30.

38. Urrutia-Baca VH, Gomez-Flores R, De La Garza-Ramos MA, Tamez-Guerra P, Lucio-Sauceda DG, Rodríguez-Padilla MC. Immunoinformatics Approach to Design a Novel Epitope-Based Oral Vaccine Against Helicobacter pylori. Journal of computational biology : a journal of computational molecular cell biology. 2019;26(10):1177-90.

39. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2–a server for in silico prediction of allergens. Journal of molecular modeling. 2014;20(6):2278.

40. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature methods. 2011;8(10):785-6.

41. Shen HB, Chou KC. A top-down approach to enhance the power of predicting human protein subcellular localization: Hum-mPLoc 2.0. Analytical biochemistry. 2009;394(2):269-74.

42. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC bioinformatics. 2007;8:4.

43. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP. In silico approach for predicting toxicity of peptides and proteins. PloS one. 2013;8(9):e73957.

44. Ye J, McGinnis S, Madden TL. BLAST: improvements for better sequence analysis. Nucleic acids research. 2006;34(Web Server issue):W6-9.

45. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Computer applications in the biosciences : CABIOS. 1995;11(6):681-4.

46. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.

47. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research. 2007;35(Web Server issue):W407-10.

48. Vakser IA. Long-distance potentials: an approach to the multiple-minima problem in ligand-receptor interaction. Protein engineering. 1996;9(1):37-41.

49. Protein Data Bank: the single global archive for 3D macromolecular structure data. Nucleic acids research. 2019;47(D1):D520-d8.

50. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. Journal of molecular biology. 2007;372(3):774-97.

51. Ponomarenko J, Bui HH, 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.

52. Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC bioinformatics. 2010;11:568.

53. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. Biology direct. 2013;8:30.

54. Dhanda SK, Gupta S, Vir P, Raghava GP. Prediction of IL4 inducing peptides. Clinical & developmental immunology. 2013;2013:263952.

55. Dhall A, Patiyal S, Sharma N, Usmani SS, Raghava GPS. Computer-aided prediction and design of IL-6 inducing peptides: IL-6 plays a crucial role in COVID-19. Briefings in bioinformatics. 2021;22(2):936-45.

56. Nagpal G, Usmani SS, Dhanda SK, Kaur H, Singh S, Sharma M, et al. Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. Scientific reports. 2017;7:42851.

57. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics (Oxford, England). 2016;32(4):511-7.

58. Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. PLoS computational biology. 2013;9(10):e1003266.

59. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC bioinformatics. 2006;7:153.

60. Alland C, Moreews F, Boens D, Carpentier M, Chiusa S, Lonquety M, et al. RPBS: a web resource for structural bioinformatics. Nucleic acids research. 2005;33(Web Server issue):W44-9.

61. Rapin N, Lund O, Bernaschi M, Castiglione F. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PloS one. 2010;5(4):e9862.

62. Michael R. Green JS. Molecular Cloning: A Laboratory Manual Science Press Co., Ltd; 2017.

63. Michael R. Green JS. Molecular Cloning: A Laboratory Manual (Fourth Edition): Three-Volume Set. 2012.

64. Moodley A, Fatoba A, Okpeku M, Emmanuel Chiliza T, Blessing Cedric Simelane M, Pooe OJ. Reverse vaccinology approach to design a multi-epitope vaccine construct based on the Mycobacterium tuberculosis biomarker PE_PGRS17. Immunologic research. 2022;70(4):501-17.

65. Scheiblhofer S, Laimer J, Machado Y, Weiss R, Thalhamer J. Influence of protein fold stability on immunogenicity and its implications for vaccine design. Expert review of vaccines. 2017;16(5):479-89.

66. Saylor K, Gillam F, Lohneis T, Zhang C. Designs of Antigen Structure and Composition for Improved Protein-Based Vaccine Efficacy. Frontiers in immunology. 2020;11:283.

67. Stewart P, Patel S, Comer A, Muneer S, Nawaz U, Quann V, et al. Role of B Cells in Mycobacterium Tuberculosis Infection. Vaccines. 2023;11(5).

68. Accapezzato D, Visco V, Francavilla V, Molette C, Donato T, Paroli M, et al. Chloroquine enhances human CD8+ T cell responses against soluble antigens in vivo. The Journal of experimental medicine. 2005;202(6):817-28.

69. Delamarre L, Pack M, Chang H, Mellman I, Trombetta ES. Differential lysosomal proteolysis in antigen-presenting cells determines antigen fate. Science (New York, NY). 2005;307(5715):1630-4.

70. Delamarre L, Couture R, Mellman I, Trombetta ES. Enhancing immunogenicity by limiting susceptibility to lysosomal proteolysis. The Journal of experimental medicine. 2006;203(9):2049-55.

71. Rodriguez MS, Desterro JM, Lain S, Lane DP, Hay RT. Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. Molecular and cellular biology. 2000;20(22):8458-67.

72. Khamsri B, Fujita M, Kamada K, Piroozmand A, Yamashita T, Uchiyama T, et al. Effects of lysine to arginine mutations in HIV-1 Vif on its expression and viral infectivity. International journal of molecular medicine. 2006;18(4):679-83.

73. Timms RT, Zhang Z, Rhee DY, Harper JW, Koren I, Elledge SJ. A glycine-specific N-degron pathway mediates the quality control of protein N-myristoylation. Science (New York, NY). 2019;365(6448).

74. Faridgohar M, Nikoueinejad H. New findings of Toll-like receptors involved in Mycobacterium tuberculosis infection. Pathogens and global health. 2017;111(5):256-64.

75. Ciesielska A, Matyjek M, Kwiatkowska K. TLR4 and CD14 trafficking and its influence on LPS-induced pro-inflammatory signaling. Cellular and molecular life sciences : CMLS. 2021;78(4):1233-61.

76. Geng J, Shi Y, Zhang J, Yang B, Wang P, Yuan W, et al. TLR4 signalling via Piezo1 engages and enhances the macrophage mediated host response during bacterial infection. Nature communications. 2021;12(1):3519.

77. Saelee C, Hanthamrongwit J, Soe PT, Khaenam P, Inthasin N, Ekpo P, et al. Toll-like receptor-mediated innate immune responses by recognition of the recombinant dormancy-associated Mycobacterium tuberculosis proteins Rv2659c and Rv1738. PloS one. 2022;17(9):e0273517.

78. Pahari S, Negi S, Aqdas M, Arnett E, Schlesinger LS, Agrewala JN. Induction of autophagy through CLEC4E in combination with TLR4: an innovative strategy to restrict the survival of Mycobacterium tuberculosis. Autophagy. 2020;16(6):1021-43.

79. Jo EK, Park JK, Dockrell HM. Dynamics of cytokine generation in patients with active pulmonary tuberculosis. Current opinion in infectious diseases. 2003;16(3):205-10.

80. Lienhardt C, Azzurri A, Amedei A, Fielding K, Sillah J, Sow OY, et al. Active tuberculosis in Africa is associated with reduced Th1 and increased Th2 activity in vivo. European journal of immunology. 2002;32(6):1605-13.

81. Scriba TJ, Netea MG, Ginsberg AM. Key recent advances in TB vaccine development and understanding of protective immune responses against Mycobacterium tuberculosis. Seminars in immunology. 2020;50:101431.

82. Saqib M, Khatri R, Singh B, Gupta A, Kumar A, Bhaskar S. Mycobacterium indicus pranii as a booster vaccine enhances BCG induced immunity and confers higher protection in animal models of tuberculosis. Tuberculosis (Edinburgh, Scotland). 2016;101:164-73.

83. Biselli R, Mariotti S, Sargentini V, Sauzullo I, Lastilla M, Mengoni F, et al. Detection of interleukin-2 in addition to interferon-gamma discriminates active tuberculosis patients, latently infected individuals, and controls. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2010;16(8):1282-4.

84. Penn-Nicholson A, Tameris M, Smit E, Day TA, Musvosvi M, Jayashankar L, et al. Safety and immunogenicity of the novel tuberculosis vaccine ID93 + GLA-SE in BCG-vaccinated healthy adults in South Africa: a randomised, double-blind, placebo-controlled phase 1 trial. The Lancet Respiratory medicine. 2018;6(4):287-98.

85. Ogongo P, Tezera LB, Ardain A, Nhamoyebonde S, Ramsuran D, Singh A, et al. Tissue-resident-like CD4+ T cells secreting IL-17 control Mycobacterium tuberculosis in the human lung. The Journal of clinical investigation. 2021;131(10).

86. Wik JA, Skalhegg BS. T Cell Metabolism in Infection. Frontiers in immunology. 2022;13:840610.

87. Serbina NV, Lazarevic V, Flynn JL. CD4(+) T cells are required for the development of cytotoxic CD8(+) T cells during Mycobacterium tuberculosis infection. Journal of immunology (Baltimore, Md : 1950). 2001;167(12):6991-7000.

88. Okada M, Kita Y, Nakajima T, Hashimoto S, Nakatani H, Nishimatsu S, et al. The study of novel DNA vaccines against tuberculosis: induction of pathogen-specific CTL in the mouse and monkey models of tuberculosis. Human vaccines & immunotherapeutics. 2013;9(3):515-25.

89. Lew MH, Norazmi MN, Nordin F, Tye GJ. A novel peptide vaccination augments cytotoxic CD8(+) T-cell responses against Mycobacterium tuberculosis HspX antigen. Immunobiology. 2022;227(3):152201.

90. Smith SM, Brookes R, Klein MR, Malin AS, Lukey PT, King AS, et al. Human CD8+ CTL specific for the mycobacterial major secreted antigen 85A. Journal of immunology (Baltimore, Md : 1950). 2000;165(12):7088-95.

91. Launois P, DeLeys R, Niang MN, Drowart A, Andrien M, Dierckx P, et al. T-cell-epitope mapping of the major secreted mycobacterial antigen Ag85A in tuberculosis and leprosy. Infection and immunity. 1994;62(9):3679-87.

Table 1. The servers and databases of bioinformatics analysis used in this study

Servers/databases	Website
ProtParam	https://web.expasy.org/protparam/
VaxiJen	http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html
AllerTop	http://www.ddg-pharmfac.net/AllerTOP
ToxinPred	https://webs.iiitd.edu.in/raghava/toxinpred/multi_submit.php
Blastp	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins
SignaIP 4.1	https://services.healthtech.dtu.dk/service.php?SignalP-4.1
TMHMM 2.0	https://services.healthtech.dtu.dk/service.php?TMHMM-2.0
Cell-PLoc 2.0	http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/
NetPhos	https://services.healthtech.dtu.dk/services/NetPhos-3.1
NetNGlyc	https://services.healthtech.dtu.dk/services/NetNGlyc-1.0

Servers/databases	Website
YinOYang	https://services.healthtech.dtu.dk/services/YinOYang-1.2
SOPMA	https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.htm
bkuyun	https://www.bkunyun.com
Prosa	https://prosa.services.came.sbg.ac.at/prosa.php
GRAMM	https://gramm.compbio.ku.edu/gramm
PDB	https://www.rcsb.org/
PDBePISA	https://www.ebi.ac.uk/pdbe/pisa/
ElliPro	http://tools.iedb.org/ellipro/
MHC-I Binding Predictions	http://tools.iedb.org/mhci/
Class I Immunogenicity	http://tools.iedb.org/immunogenicity/
MHC-II Binding Predictions	http://tools.iedb.org/mhcii/
CD4 T cell immunogenicity prediction	http://tools.iedb.org/CD4episcore/
IFNepitope	http://crdd.osdd.net/raghava/ifnepitope/predict.php
IL4pred	https://webs.iiitd.edu.in/raghava/il4pred/predict.php
IL6pred	https://webs.iiitd.edu.in/raghava/il6pred/predict3.php
IL-10pred	https://webs.iiitd.edu.in/raghava/il10pred/predict3.php
Population Coverage	http://tools.iedb.org/population/
PEP-FOLD 3.5 server	https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3
C-ImmSim	http://150.146.2.1/C-IMMSIM/index.php

Table 2. Physicochemical and biological properties of the protein encoded by W541 DNA vaccine

Properties	Results predicted
Total number of amino acids	704
Molecular weight	74644.63
Formula	C3329H5035N923O993S24
Theoretical pI	6.29
Total number of negatively charged residues (Asp+Glu)	52
Total number of positively charged residues (Arg+Lys)	48
Total number of atoms	10304
Number of phosphorylation sites	57
Number of glycosylation sites	7
Estimated half-life	1.1 hours (mammalian reticulocytes, in vitro). 3 min (yeast, in vivo). 2 min (Escherichia coli, in vivo).
Instability index (II)	45.37(unstable)
Aliphatic index (AI)	68.85
Grand average of hydropathicity (GRAVY)	-0.294
Antigenicity	0.6527(ANTIGEN)
Allergenicity	Non-allergenicity
Toxicity	Non-toxin
Signal peptide	No
TMHMM result	Outside
Subcellular localization (Cell-PLoc, Gram-positive bacterial)	Extracell
Subcellular localization (Cell-PLoc, Hum-mPLoc)	Cytoplasm, lysosome

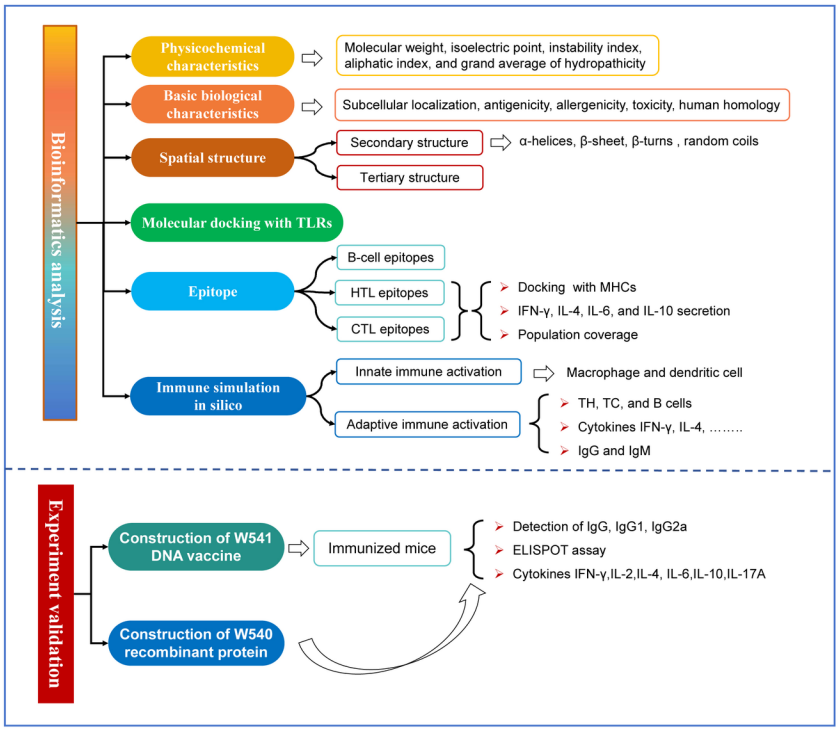
Table 3A. Docking results of selected HTL epitopes with corresponding MHC alleles

Antigens	Peptides	Alleles
Ag85A	EYLQVPSPSMGRDIK	HLA-DRB1*04:01,HLA-DRB1*04:01,HLA-DRB1*01:01,HLA-DRB1*01:01,HLA-DR
	YLQVPSPSMGRDIKV	HLA-DRB1*04:01,HLA-DRB1*04:01,HLA-DRB1*01:01,HLA-DRB1*01:01,HLA-DR
	QVPSPSMGRDIKVQF	HLA-DRB1*03:01,HLA-DRB1*03:01,HLA-DRB4*01:01,HLA-DRB4*01:01
	VPSPSMGRDIKVQFQ	HLA-DRB1*03:01,HLA-DRB1*03:01,HLA-DRB4*01:01,HLA-DRB4*01:01
	PSPSMGRDIKVQFQS	HLA-DRB1*03:01,HLA-DRB1*03:01,HLA-DRB4*01:01,HLA-DRB4*01:01,HLA-DQ
	DIKVQFQSGGANSPA	HLA-DRB1*04:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB4*01:01,HLA-DRB1*
	TGSAVVGLSMAASSA	HLA-DQA1*01:02/DQB1*06:02,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*01:01
	VGLSMAASSALTLAI	HLA-DRB1*09:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*07:01,HLA-DQA1*
	WVYCGNGKPSDLGGN	HLA-DQA1*05:01/DQB1*03:01
Ag85B	PGLVGLAGGAATAGA	HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*01:01,HLA-DRB1*15:01,HLA-DQA1*
	LRAQDDYNGWDINTP	HLA-DQA1*01:01/DQB1*05:01
	KPTGSAAIGLSMAGS	HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*01:01
	PTGSAAIGLSMAGSS	HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*09:01
	TGSAAIGLSMAGSSA	HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*01:01
	GSAAIGLSMAGSSAM	HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*09:01,HLA-DQA1*01:02/DQB1*06:02
	SAAIGLSMAGSSAMI	HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*09:01,HLA-DQA1*01:02/DQB1*06:02
	SMAGSSAMILAAYHP	HLA-DQA1*01:02/DQB1*06:02,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*15:01
	DPSQGMGPSLIGLAM	HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*09:01,HLA-DRB1*01:01

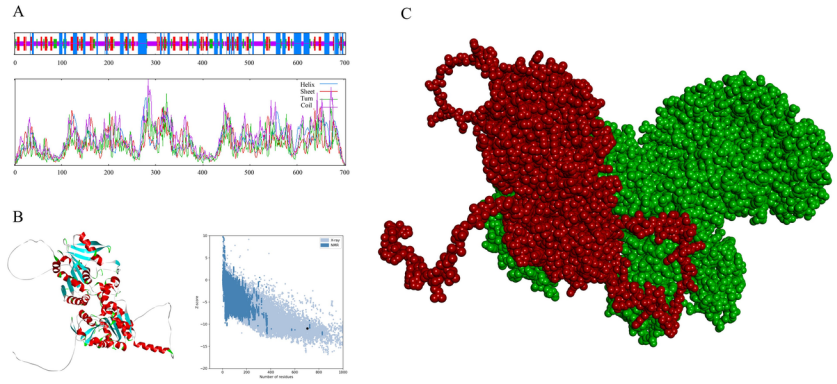
Table 3B. Docking results of selected CTL epitopes with corresponding MHC alleles

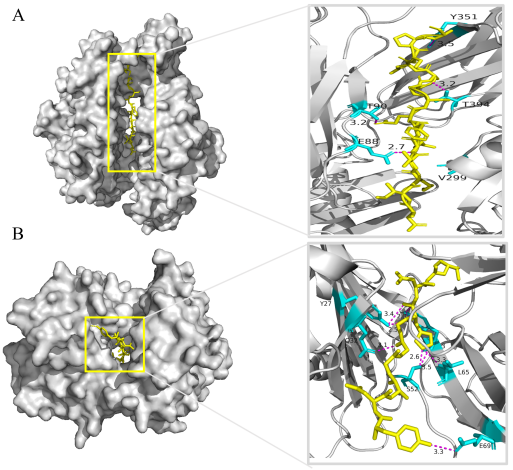
Antigens	Peptides	Alleles	Start	End	Length	Antigeni
Ag85A	FSGWDINTPA	HLA-A*02:06,HLA-A*02:01,HLA-A*68:02,HLA-A*02:03	48	57	10	0.9887
	SGWDINTPA	HLA-A*02:06	49	57	9	1.4382
	SGWDINTPAF	HLA-A*23:01,HLA-B*15:01	49	58	10	1.5629
	NTPAFEWYD	HLA-A*68:02	54	62	9	0.927
	ALTLAIYHP	HLA-A*02:03	132	140	9	1.0436
	FQDAYNAGGG	HLA-A*02:06,HLA-A*02:03,HLA-A*68:02	240	249	10	1.4901
	YNAGGGHNGV	HLA-A*68:02,HLA-A*02:06,HLA-A*02:03	244	253	10	2.4651
	NAGGGHNGV	HLA-A*68:02	245	253	9	3.0386
	WEYWGAQLN	HLA-B*40:01,HLA-B*40:01	264	272	9	0.904
	RALGATPNTG	HLA-A*02:06	280	289	10	0.8506
Ag85B	GTAAAVVLP	HLA-A*68:02	309	317	9	0.9815
	VGLAGGAATA	HLA-A*02:03	320	329	10	1.0136
	GLAGGAATA	HLA-A*02:03,HLA-A*02:01,HLA-A*02:06	321	329	9	1.3338
	AGGAATAGAF	HLA-B*15:01	323	332	10	1.1862
	GGAATAGAF	HLA-B*15:01	324	332	9	1.0961
	GSAAIGLSM	HLA-B*58:01,HLA-B*15:01	450	458	9	1.2027
	RLWVYCGNGT	HLA-A*02:03	537	546	10	0.8136
	TPNELGGAN	HLA-B*35:01	546	554	9	1.3143
	TPNELGGANI	HLA-B*07:02	546	555	10	1.2042
	ELGGANIPA	HLA-A*68:02	549	557	9	0.8819
	AYNAAGGHNA	HLA-A*24:02	574	583	10	1.5728
	YNAAGGHNAV	HLA-A*68:02,HLA-A*02:03,HLA-A*02:06	575	584	10	1.6111
	NAAGGHNAV	HLA-A*68:02,HLA-B*35:01,HLA-A*02:06,HLA-A*02:03	576	584	9	1.9957

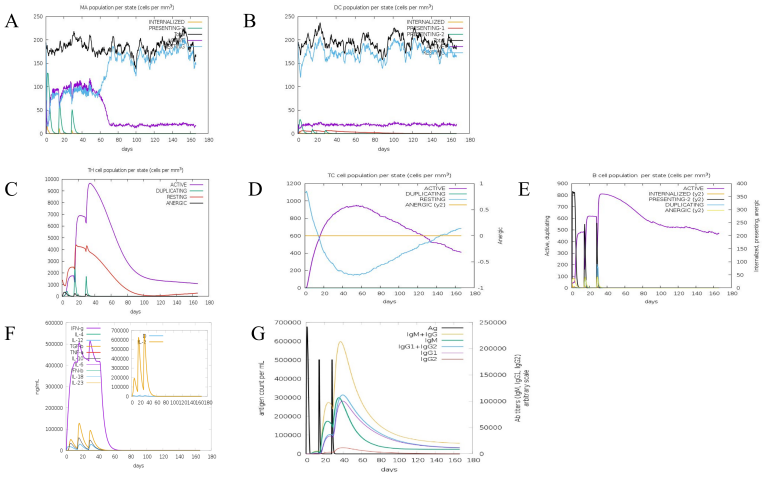
Antigens	Peptides	Alleles	Start	End	Length	Antigeni
	NAAGGHNAVF	HLA-B*15:01,HLA-B*35:01	576	585	10	1.4758
	AAGGHNAVF	HLA-B*15:01,HLA-B*35:01	577	585	9	1.1841
	AVFNFPPNG	HLA-A*30:01	583	591	9	0.9399
Rv1733c	FAAAAGTAVQ	HLA-B*15:01	679	688	10	0.8588

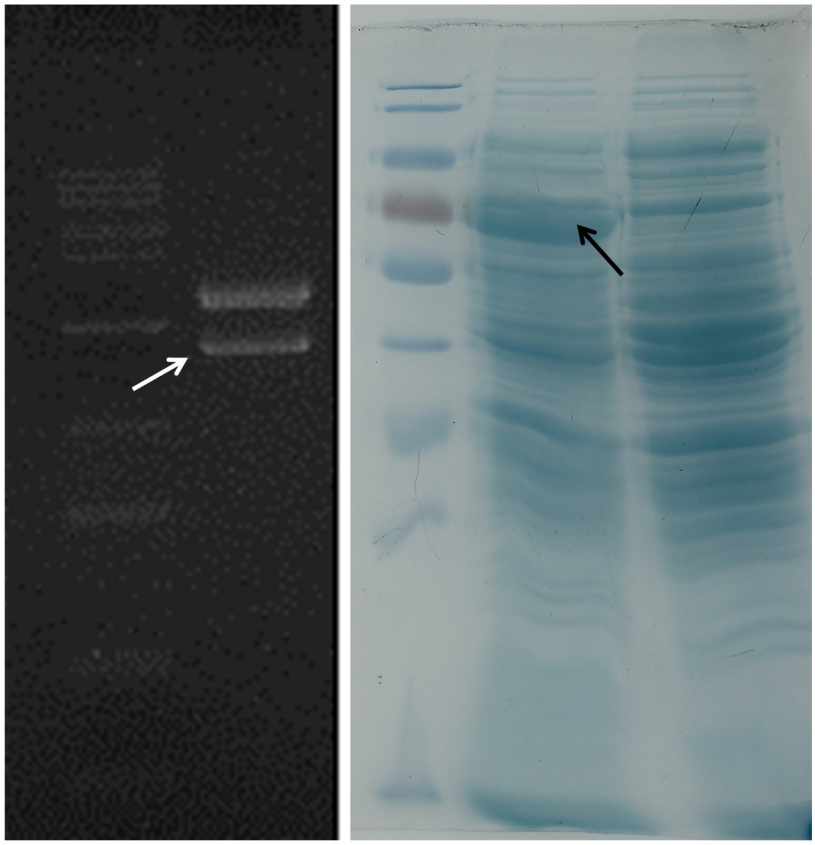


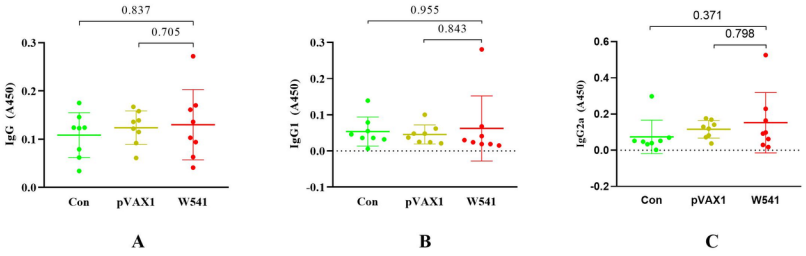
FSRPGLPVEYLQVPSPSMGRDIKVQFQSGGANSPALYLLDGLRAQDDFSGWDINTPAFEWYDQSGLS
VVMPVGGQSSFYSDWYQPACGKAGCQTYKWETFLTSELPGWLQANRHVKPTGSAVVGLSMAASSA
LTLAIYHPQQFVYAGAMSGLLDPSQAMGPTLIGLAMGDAGGYKASDMWGPKEDPAWQRNDPLLNV
GKLIANNTRVWVYCGNGKPSDLGGNNLPAKFLEGFVRTSNIKFQDAYNAGGGHNGVFDFPDSGTHS
WEYWGAQLNAMKPDLQRALGATPNTGPAPQGAGSGGGSGGRRLMIGTAAAVVLPGLVGLAGGAA
TAGAFSRPGLPVEYLQVPSPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYY
QSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAAIGLSMA
GSSAMILAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDPT
QQIPKLVANNTRLWVYCGNGTPNELGGANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNFPPNGT
HSWEYWGAQLNAMKGDLQGPGPGLRQHASRYLARVEAGGPGPGSGVLIPARRPQNLLDVTAEPAR
GRKRTLSDVLNEMRGPGPGIPFAAAAGTAVQDSRSHVYAHQAQTRHP

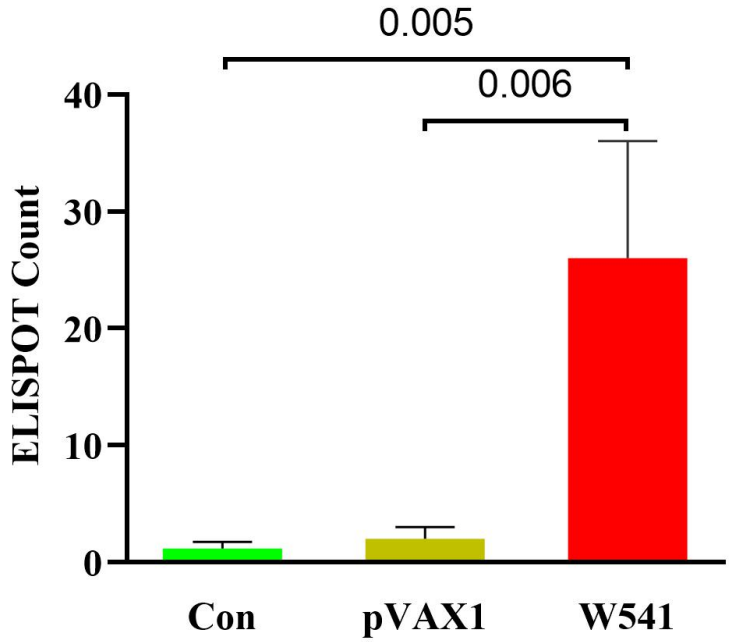


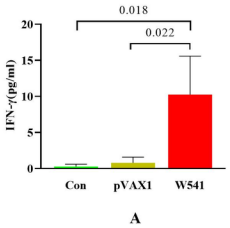




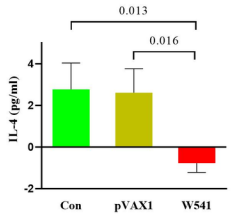




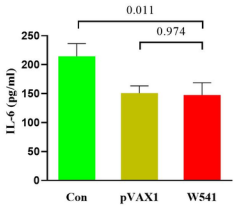




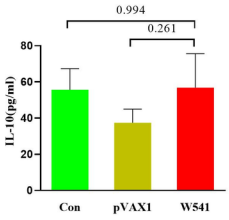
A



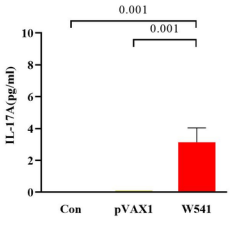
B



C



D



E