Synthetic and Systems Biotechnology 9 (2024) 391–405

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Original Research Article

Discovering peptides and computational investigations of a multiepitope vaccine target *Mycobacterium tuberculosis*

Truc Ly Nguyen [a], Heebal Kim [a,b,c,*]

[a] Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, 08826, Republic of Korea
[b] Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, 08826, Republic of Korea
[c] eGnome, Inc., Seoul, 05836, Republic of Korea

ARTICLE INFO

Keywords:
Mycobacterium tuberculosis
Tuberculosis
Multiepitope vaccine
Docking molecular
Molecular dynamics simulation
Immune simulation

ABSTRACT

Mycobacterium tuberculosis (MTB) is the causative agent of tuberculosis (TB), a prevalent airborne infectious disease. Despite the availability of the Bacille Calmette-Guerin vaccine, its global efficacy remains modest, and tuberculosis persists as a significant global public health threat. Addressing this challenge and advancing towards the End MTB Strategy, we developed a multiepitope vaccine (MEV) based on immunoinformatics and computational approaches. Immunoinformatics screening of MBT protein identified immune-dominant epitopes based on Major Histocompatibility Complex (MHC) allele binding, immunogenicity, antigenicity, allergenicity, toxicity, and cytokine inducibility. Selected epitopes were integrated into an MEV construct with adjuvant and linkers, forming a fully immunogenic vaccine candidate. Comprehensive analyses encompassed the evaluation of immunological and physicochemical properties, determination of tertiary structure, molecular docking with Toll-Like Receptors (TLR), molecular dynamics (MD) simulations for all atoms, and immune simulations. Our MEV comprises 534 amino acids, featuring 6 cytotoxic T lymphocyte, 8 helper T lymphocyte, and 7 linear B lymphocyte epitopes, demonstrating high antigenicity and stability. Notably, molecular docking studies and triplicate MD simulations revealed enhanced interactions and stability of MEV with the TLR4 complex compared to TLR2. In addition, the immune simulation indicated the capacity to effectively induce elevated levels of antibodies and cytokines, emphasizing the vaccine's robust immunogenic response. This study presents a promising MEV against TB, exhibiting favorable immunological and physicochemical attributes. The findings provide theoretical support for TB vaccine development. Our study aligns with the global initiative of the End MTB Strategy, emphasizing its potential impact on addressing persistent challenges in TB control.

1. Introduction

Tuberculosis (TB), a bacterial disease that primarily affects the lungs, is preventable and treatable, but 10 million people still catch it annually, and 1.6 million people died from TB in 2021, almost entirely in low and middle-income countries [1]. TB has long been the world's deadliest infectious disease treatment, although it has suffered a setback and has been disrupted due to the COVID-19 pandemic [2]. Like the Bacillus Calmette-Guerin (BCG) vaccine, tools to fight TB are imperfect. However, there is "hopeful" innovation in vaccines like DNA vaccines, live attenuated and killed whole-cell vaccines (WCVs), multiepitope vaccines (MEVs), etc. Among them, the therapeutic DNA vaccine is a promising vaccine, which is a promising strategy against tuberculosis. However, challenges with the delivery and expression of the DNA vaccine and potential issues with inducing an adequate immune response in all individuals may limit its efficacy [3]. Live attenuated and killed whole-cell vaccines (WCVs) also offer promising vaccination strategies against tuberculosis. However, their efficacy may be compromised in immunocompromised individuals, and there is a risk of virulence reversion in live attenuated vaccines, leading to the potential for disease transmission [4]. MEVs are a type of vaccine that can be composed of cytotoxic T lymphocyte (CTL), helper T lymphocyte (HTL), and linear B lymphocyte (LBL) epitopes in a series or overlapping epitope peptides [5]. They are designed to induce multi antigenic immunity against significant complex pathogens with different strain variants [6]. MEVs can be used to prevent and treat tumors or viral infections [5,7,8]. Besides,

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author: Department of Agricultural Biotechnology, Seoul National University, Seoul, 08826, Republic of Korea.
E-mail address: heebal@snu.ac.kr (H. Kim).

https://doi.org/10.1016/j.synbio.2024.03.010
Received 31 January 2024; Received in revised form 9 March 2024; Accepted 12 March 2024
Available online 21 March 2024

the advantage of epitope vaccines over traditional subunit vaccines lies in the ability to combine immunodominant human HTL, CTL, and LBL epitopes from different antigens, thereby enhancing immunogenicity and reducing adverse effects. Currently, many MEVs based on epitope designs from multiple antigens against tuberculosis are being researched [9–19]. Among them, Jiang et al. selected 17 latent tuberculosis infection and regions of difference (LTBI-RD) antigens (Rv1511, Rv1736c, Rv1737c, Rv1980c, Rv1981c, Rv2031c, Rv2626c, Rv2653c, Rv2656c, Rv2659c, Rv2660c, Rv3425, Rv3429, Rv3872, Rv3873, Rv3878, and Rv3879) to identify immunodominant epitopes [12]. Similarly, Bellini et al. designed and characterized a multistage peptide-based vaccine from 15 protein antigens associated with various activities of the MTB life cycle, including Rv3908c, Rv1886c, Rv1384, Rv1436, Rv3874, Rv0288, Rv0867c, Rv1174c, Rv1334, Rv0475, Rv0440, Rv0125, Rv1733c, Rv1039c, and Rv1039c [11]. In addition, a novel peptide-based vaccine was designed based on HTL, CTL, and B cell epitopes predicted from 17 protective antigens of MTB. These 17 candidate antigens were Ag85A, Ag85B, ESAT6, EspA, Mpt63, MTB32A, PPE18, RpfB, TB10.4, CFP10, MPT51, MPT64, MTB8.4, PPE44, PPE68, RpfA, and RpfB [13]. Notably, Bibi et al. showed that MEV might activate humoral and cellular immune responses and may be a possible tuberculosis vaccine candidate [18]. In that work, a novel MEV designed by targeting Rv2608, Rv2684, Rv3804c (Ag85A), and Rv0125 (MTB 32A) which has been predicted to have different B-cell and T-cell epitopes. Another promising candidate is an MEV against MTB exploiting secreted exosome proteins (Evs) [19]. However, the potential for immune evasion by MTB through antigenic variation may limit the vaccine's effectiveness, and there is a need for ongoing monitoring and updating of vaccine components to address this challenge.

Beyond these considerations, H37Rv is the most widely used MTB strain, and its protein Rv0256c, also known as PPE2 (Proline-Proline-Glutamate 2), has an essential role in immune activation and infection of the host [20]. Rv0256c has been found to translocate to the nucleus of host cells and bind to the promoter region of inducible nitric oxide synthase (iNOS), suppressing iNOS gene transcription [21,22], ultimately protecting the mycobacterium from nitric oxide (NO) mediated killing. Additionally, Rv0256c has been found to inhibit myeloid hematopoiesis and reactive oxygen species (ROS) production [23,24]. In a previous study, Rv0256c induces a strong B cell response in tuberculosis patients [25]. These findings suggest that Rv0256c is a crucial protein contributing to MTB survival and pathogenesis. Therefore, in the present study, we aim to select the Rv0256c protein as the target sequence to design a MEV candidate. Through immunoinformatics techniques, we predicted CTL, HTL, and LBL epitopes. These epitopes were shown to be highly antigenic, nontoxic, and nonallergic. The potential for these epitopes to cause autoimmunity was also examined. In addition, the Toll-Like Receptor 4 (TLR4) agonist (RpfE) peptide was added to the vaccine design as an adjuvant to boost its immunogenicity. We further evaluated the vaccine construct's population coverage, antigenicity, allergenicity, toxicity, and physicochemical features. Afterwards, the tertiary structures of the vaccine construct were predicted, refined, and validated. The resultant tertiary structure was then docked with immune receptors TLR2 and TLR4. Furthermore, the stability of interactions was verified using molecular dynamics (MD) simulations for all atoms of the docking complexes in triplicate. Finally, to assess the immunogenicity and immunological response of the MEV, in silico immune simulations were carried out. This study provides valuable insights for MTB vaccine development and contributes towards the End MTB Strategy.

2. Materials and methods

The systematic workflow used in this study is depicted in Fig. 1.

2.1. Retrieval sequence, screening antigenicity, and allergenicity of target protein

The FASTA sequence of Rv0256c was obtained from the UniProt database (https://www.uniprot.org/) with accession number P9WI47. To screen for antigenicity, we employed VaxiJen v2.0 to predict the antigenicity of the Rv0256c protein, with a threshold value of 0.4 set up (http://www.ddg-pharmfac.net/VaxiJen/VaxiJen/VaxiJen.html) [26]. This server is focused on auto cross-covariance (ACC) transformation and alignment-independent prediction that maintains predictive accuracy of 70–89%. Protein Rv0256c showed antigenicity above the threshold value and was selected for further analysis [26]. For predicting allergenicity, the Rv0256c protein sequence was expanded for further analysis based on AllergenFP v.1.0, a bioinformatics tool for allergenicity prediction (https://ddg-pharmfac.net/AllergenFP/) [27]. The result with a non-allergen property was selected for further analysis. To enhance humoral and cell-mediated immunity, both B-cell and T-cell antigens were predicted.

2.2. CTL epitopes prediction and assessment

The Immune Epitope Database and Analysis Resources (IEDB) MHC I server was used to predict CTL epitopes (http://tools.iedb.org/mhci/) [28]. This server predicts $CD8^+$ T cell epitopes based on proteasomal C-terminal cleavage, MHC-I binding, and TAP transport efficiency. The 9-mer and 10-mer epitopes were predicted using the ANN 4.0 algorithm weight matrix, artificial neural networks, and IC_{50} value. We used the entire human HLA allele reference set. Finally, the predicted peptides were sorted as per the predicted IC_{50}. As in the previous study, only epitopes with an IC50 over 500 were chosen [29].

2.3. HTL epitopes prediction and assessment

The MHC II server of IEDB was used to predict HTL epitopes (http://tools.iedb.org/mhcii/) [28]. NN-align 2.3 (Net MHC II 2.3) algorithm was used to predict the epitopes. The entire HLA human reference set was used. A 15mer epitope length was defined. Finally, the results were then arranged as per their modified ranks. Additionally, the ability to secrete IFN-γ, IL-4 and IL-10 of these chosen epitopes was predicted using the servers IFNepitope (http://crdd.osdd.net/raghava/ifnepitope/) [30], IL4pred (https://webs.iiitd.edu.in/raghava/il4pred/) [31], and IL-10 pred (http://crdd.osdd.net/raghava/IL-10pred/) [32], respectively. All chosen epitopes showed the ability to secrete these cytokines.

2.4. B-cell epitopes prediction and assessment

An online server ABCpred was used to predict linear B-cell epitopes (http://crdd.osdd.net/raghava/abcpred/) [33]. The ABCpred uses an artificial neural network to predict linear B-cell epitopes in an antigen sequence. A 0.51 threshold was applied individually to each selected protein. Epitopes were chosen to have a 16mer length.

2.5. Antigenicity, allergenicity, and toxicity prediction of chosen epitopes

All chosen epitopes were tested to assess their antigenicity, allergenicity, and toxicity. For antigenicity prediction, CTL epitopes, HTL epitopes, and conformational B cell epitopes amino acid sequences were submitted to the VaxiJen v2.0 server [26]. Both a threshold of 0.4 and bacteria were explicitly mentioned. Highly antigenic epitopes were selected for further analysis. Consequently, the allergenicity of the B cell, CTL, and HTL epitopes was predicted using the freely available allergenicity prediction tool AllerTOP v.2.0 (https://www.ddg-pharmfac.net/AllerTOP/index.html). AllerTOP v.2.0 predicts the allergens based on machine learning methods like auto and cross-covariance transformation, k nearest neighbors, and amino acid E-descriptors [34]. All

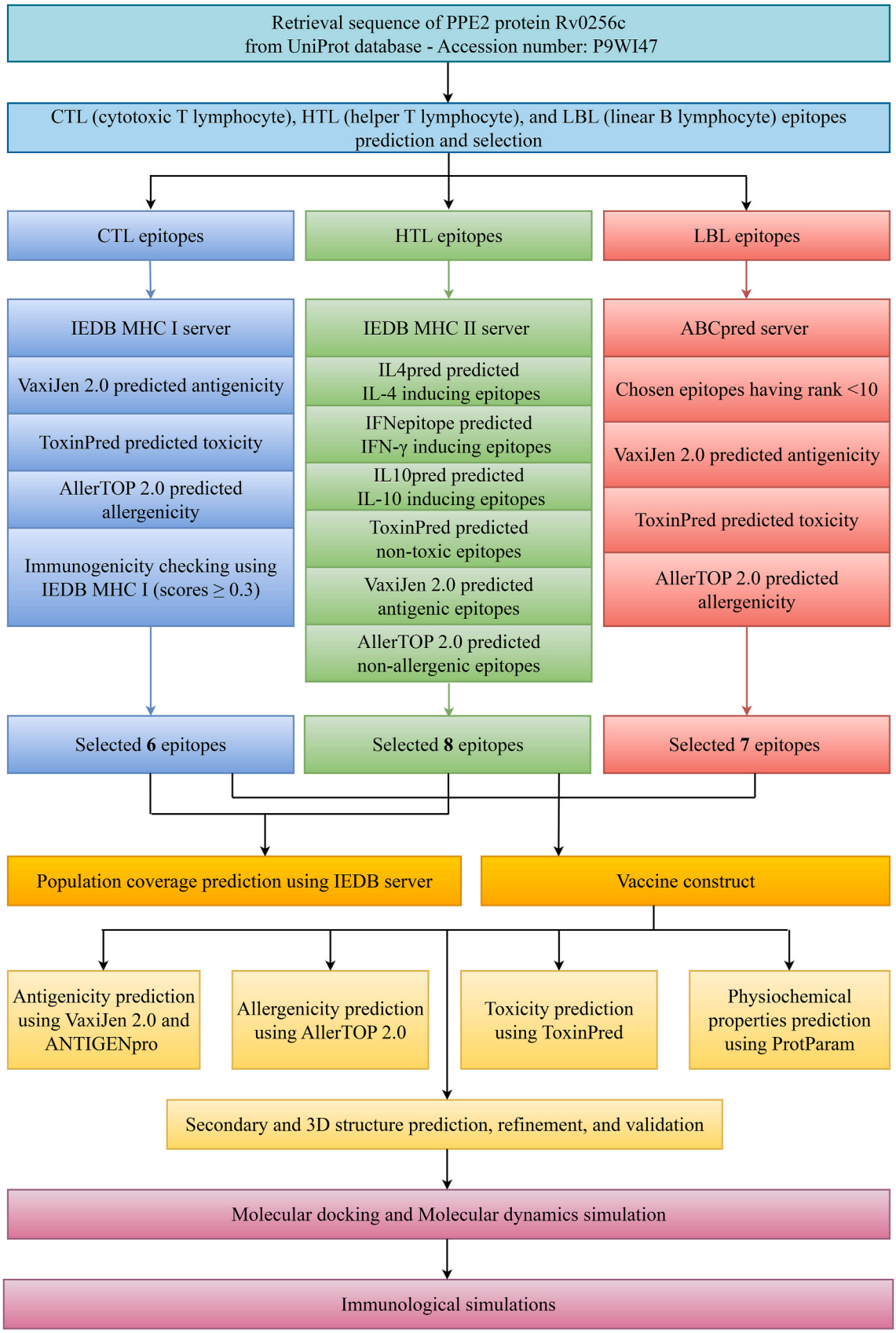

Fig. 1. The systematic workflow of this study.

settings were left at their default values. Lastly, the toxicity of the B cell, CTL, and HTL epitopes was predicted using the freely server ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) [35]. Only the epitopes identified as antigenic, non-allergenic, and non-toxic were retained for further research.

2.6. Population coverage of the epitopes

The diversity in MHC allele distribution reflects the world's geographical and cultural diversity. As a result, vaccination coverage is determined by the MHC alleles that its epitopes recognize. By using the population coverage tool in the IEDB database (http://tools.iedb.org/population/) [36], the epitopes and their corresponding MHC class I and class II alleles as input, we calculated the combined coverage of our T lymphocyte epitopes. These alleles genotypic frequencies employed in the IEDB database were obtained from the Allele Frequency Net Database (AFND) (http://www.allelefrequencies.net/) [37]. Currently, AFND offers allele frequencies for 115 countries and 21 different ethnicities grouped into 16 different geographical areas.

2.7. Multiepitope vaccine design and its properties evaluation

Highly antigenic, nonallergenic, and nontoxic epitopes were selected to design the vaccine. These best epitopes were linked through EAAAK, GPGPG, KK, and AAY linkers to construct a potential MEV. These linkers were incorporated into separate domains to let them act separately and enhance the vaccine's immunogenicity [38] as they are cleavable, flexible, and rigid [39]. Apart from these epitopes and linkers, the TLR4 agonist (RpfE) peptide was added to the vaccine design as an adjuvant for increasing the immune response [40].

The ability to elicit humoral-cellular immune responses and the knowledge of a particular antigen associated with an immune response are referred to as immunogenicity and antigenicity, respectively. Therefore, the antigenicity and immunogenicity of a candidate vaccine are vital [41]. ToxinPred and AllerTOP 2.0 were used to predict the toxicity and allergenicity of the vaccine construct. Allergenicity was checked to ensure that the vaccine did not exhibit any reactions (allergic) once injected into the body. The vaccine candidate should have a high level of antigenicity because this attribute defines an antigen's capacity to trigger an immune response and the development of memory cells. Antigenicity prediction was made using VaxiJen v2.0 and ANTIGENpro server [42]. Both methods are alignment-free. VaxiJen v2.0 functions by utilizing several physicochemical properties of the protein, while ANTIGENpro is a machine learning algorithm-based microarray analysis data-based server. Using the open web server ProtParam, many physiochemical features, including amino acid composition, Aliphatic Index (AI), molecular weight, Instability Index (II), Grand Average of Hydropathicity (GRAVY), and theoretical isoelectric point (pI) were evaluated [43].

2.8. Structure prediction, refinement, and validation of the vaccine construct

The vaccine's secondary structure motifs calculated with PDBsum [44] were computed using v.3.0 of Gail Hutchinson's PROMOTIF program [45]. PDBsum is a web server (http://www.ebi.ac.uk/pdbsum) that offers structural details about the entries in the Protein Data Bank (PDB). Protein secondary structure, interactions between proteins and ligands and DNA, PROCHECK structural quality evaluations, and numerous more analyses are among the mostly image-based analyses. The PROMOTIF program examines a protein coordinate file and gives information regarding the structural motifs present in the protein. Currently, the program evaluates the following structural features: secondary structure, β- and γ-turns, helical geometry and interactions, β-strands and β-sheet topology, β-hairpins, etc. In order to illustrate each type of motif in the protein, PROMOTIF generates postscript files along with a summary page.

For predicting the three-dimensional (3D) structure of the vaccine, we used several servers such as SWISS-MODEL (https://swissmodel.expasy.org/) - a fully automated protein structure homology-modeling server [46], Iterative Threading ASSEmbly Refinement (I-TASSER) (https://zhanggroup.org/I-TASSER/) [47], or a deep learning approach AlphaFold2 via ColabFold v1.5.5 (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb) [48,49]. Afterwards, the GalaxyRefine module of the GalaxyWEB server (http://galaxy.seoklab.org/) was used to refine the vaccine's 3D structures [50]. Through the process of 3D structural refinement, the vaccine's near-accurate native structure was preserved while local mistakes were corrected and the accuracy of initially anticipated structures was improved. Subsequently, ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) [51] and PROCHECK v.3.5 (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) [52] were two of the many publicly available tools utilized for 3D structural validation. ProSA-web is a web server that evaluates the overall quality and local model quality of 3D models based on the Z-score value [53,54]. Meanwhile, the PROCHECK program analyzes the Ramachandran plot for the 3D structure of the vaccine construct to assess residual coverage in favored, allowed, and disallowed regions. A good quality structure would be expected to have over 90% of residues in the most favored regions.

2.9. Molecular docking and molecular dynamics simulations studies

The structure of TLR2 and TLR4 were downloaded from the Protein Data Bank (PDB) with ID 2Z7X and 3FXI, respectively. The 3D structure of MEV and immune receptors were docked using the ClusPro server (https://cluspro.bu.edu) - a widely used protein-protein docking server that predicts the 3D structures of protein complexes [55]. The following three steps are used by ClusPro to examine the molecular docking of vaccine with TLRs: (1) rigid body docking by sampling billions of conformations, (2) grouping of the 1000 lowest energy structures generated to identify the largest clusters based on root-mean-standard deviation (RMSD), (3) energy minimization for steric clash removal. Accordingly, the server provided 30 model complexes, out of which the model having the lowest binding energy (kcal/mol) was selected for dynamics. Subsequently, PDBsum was used for analysis and to find interacting residues between the vaccine and TLR2, 4 (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html) [44].

MD simulations were conducted using the GROMACS program on a Linux operating system to assess the stability of the complexes [56]. The CHARMM27 force field and spce water were employed to generate topology files, resulting in a system with 24,783 atoms from 1603 residues for the MEV-TLR2 complex and 26,709 atoms from 1736 residues for the MEV-TLR4 complex. Each complex was placed in a cubic box (12x12x12 for MEV-TLR2 and 11x11x11 for MEV-TLR4) to maintain integrity with 297,338 and 227,396 solvent molecules, respectively. To neutralize the charge, 27 Cl^- ions were added to the MEV-TLR2 complex, and 5 Na^+ ions were added to the MEV-TLR4 complex. The energy minimization utilized the steepest descent algorithm with 50,000 steps, stopping when the maximum force was $<$1000.0 kJ/mol/nm. Position restraints were applied during equilibration, including NVT equilibration at 300 K with 50,000 steps (100 ps) and NPT equilibration at 1 bar reference pressure with an additional 50,000 steps (100 ps). Production simulations for all-atom systems (916,797 atoms in MEV-TLR2 and 708,897 atoms in MEV-TLR4) were carried out using the NPT ensemble for 50,000,000 steps (100 ns). After completing the 100 ns MD simulation, analyses were performed, including calculating the root mean square deviation (RMSD) of backbone residues, root mean square fluctuation (RMSF) of C-alpha, radius of gyration (Rg), and solvent accessible surface area (SASA). Each complex was simulated in triplicate to ensure result accuracy, robustness, and dependability.

2.10. Immunological responses induced by the vaccine construct

C-IMMSIMversion 10.1 is an immune simulation tool that can assess the vaccine's immunological response (https://kraken.iac.rm.cnr.it/C-IMMSIM/) [57]. It predicts a position-specific scoring matrix used to understand immune response magnitude, which shows the result of vaccine dosage concerning different time intervals. This server describes a mammalian immune system's humoral and cellular response against vaccine construct. To quantify the impact of antigens and foreign particles on immune activity, an agent-based method based on the position-specific scoring matrix and machine learning techniques was applied. Except for the time steps, which were set at 1, 84, and 168, the simulation was run using the default parameters. Injection occurs at time = 0 in time step 1, and each time step lasts for 8 h. The entire simulation consisted of 1050 steps. Since most commercial vaccines prescribe a four-week delay between doses, three injections were anticipated to be needed at four-week intervals [58].

3. Results

3.1. Retrieval sequence, screening antigenicity, and allergenicity of target protein

After retrieving the FASTA sequence of the Rv0256c protein from the UniProt database, immunoinformatics analysis was used to predict the antigenicity and allergenicity of the protein. The Rv0256c protein was predicted as a probable antigen (0.4302) based on the VaxiJen v2.0 server with a threshold of 0.4. AllergenFP v.1.0 results indicated that the protein had non-allergenic properties as it held the highest Tanimoto similarity index of 0.94. Thus, we confirmed that the Rv0256c protein sequence could be considered for CTL, HTL, and B-cell epitope prediction.

3.2. CTL, HTL, and B-cell epitopes prediction and assessment

The Rv0256c protein sequence predicted CTL epitopes using the MHC-I tool from the IEDB server. From the IEDB server, the completed human HLA allele reference set was selected for epitope prediction. Epitopes with IC50 over 500 were chosen for further study. Further on, epitopes were checked for toxicity, antigenicity, and allergenicity. Only non-toxic, antigenic, and non-allergenic epitopes were chosen (Supplementary Data Sheet 1). An immunogenicity check was done using the IEDB server, and immunogenic epitopes with scores ≥0.3 were selected. Finally, six epitopes were included in the vaccine construct (Table 1).

The IEDB MHC II server was used for HTL epitope prediction. Epitopes were checked for toxicity, antigenicity, and allergenicity. Only non-toxic, antigenic, and non-allergenic epitopes were chosen. Epitopes were checked and selected based on their ability to induce Il-4, Il-10, and IFN-γ (Supplementary Data Sheet 2). Finally, we included eight possible epitopes in the vaccine that induced the abovementioned cytokines (Table 2).

B-cell epitopes having a rank <10 predicted using the ABCpred webserver were selected for further studies. Furthermore, we selected only those epitopes that were non-toxic, non-allergenic, and antigenic to further incorporate into the vaccine design (Supplementary Data Sheet 3). We used the web servers ToxinPred, AllerTOP, and VaxiJen, in that order. Finally, seven epitopes were selected for the vaccine construct (Table 3).

3.3. Population coverage analysis

The population coverage of the 14 T-lymphocyte epitopes (combined CTL and HTL) employed in this designed vaccine was evaluated using IEDB population coverage analysis. The IEDB database assessed the distribution of their 55 corresponding MHC alleles in 16 geographical areas and 101 countries. The region-wise coverage of alleles is represented in Fig. 2. Notably, our vaccine demonstrates a global coverage rate of 99.74%.

3.4. Multiepitope vaccine design and its properties evaluation

For the MEV construction, highly antigenic, non-allergenic, and non-toxic epitopes were selected. According to the results in Tables 1–3, six CTL epitopes, eight HTL epitopes, and seven B-cell epitopes were selected. The selected epitopes were linked with amino acid linkers like EAAAK, GPGPG, KK, and AAY. Furthermore, to increase the immune response, the TLR4 agonist (RpfE) peptide was added as an adjuvant to the N-terminal of the vaccine since, in their study, Lee et al. discovered the new finding that MTB directly binds TLR4 and initiates TLR4 signaling, which in turn causes DCs to produce IL-1 beta and express co-stimulatory and MHC antigen presentation molecules [40]. As a result, the amino acid sequence of the constructed vaccine is mentioned in Fig. 3.

The antigenicity, allergenicity, toxicity, and physicochemical analyses have been listed in Table 4. The vaccine was predicted to be antigenic by VaxiJen and ANTIGENpro with scores of 0.9363 and 0.9399, respectively. The vaccine was expected to be non-allergenic by AllerTOP. It was found to be non-toxic by ToxinPred. The ProtParam server determined the vaccine construct's molecular weight to be 53.80 kDa, and its 10.10 pI suggested that it had basic properties. Of those residues, there were 534 amino acids; 64 were positively charged, and 33 were negatively charged. The II was calculated to be 28.63 in terms of instability, indicating that the construct is stable following expression (a value above 40 predicts that the protein may be unstable). According to the AI calculation, the construct is thermostable, with a value of 62.77. The Grand Average of Hydropathicity (GRAVY), which indicates how hydrophilic a substance is, was estimated to be negative (−0.354). Based on these results, this MEV construct can be predicted as a potential vaccine candidate.

3.5. Structure prediction, refinement, and validation of the vaccine construct

The vaccine's secondary structure motifs are shown in Fig. 4 which were computed using v.3.0 of Gail Hutchinson's PROMOTIF program. Particularly, among 534 residues, there are 0.7% β-strand, 25.5% α-helix, 1.1% 3_{10}-helix, 10.5% β-turn, 1.2% γ-turn, 0.7% β-hairpins, and 60.3% others.

Table 1
CTL epitopes for vaccine construction.

Peptide	Length	Antigenicity scores[a]	Toxicity	Allergenicity	Immunogenicity scores[b]
LMATNFFGIN	10	0.9704	Non-Toxin	Non-Allergen	0.37602
FSGFDPWLPS	10	0.9231	Non-Toxin	Non-Allergen	0.33227
PANIAFALGY	10	0.6439	Non-Toxin	Non-Allergen	0.33119
VIQPFINWL	9	0.5510	Non-Toxin	Non-Allergen	0.31448
SPANIAFALG	10	0.5157	Non-Toxin	Non-Allergen	0.31063
GNPATIAFT	9	1.0466	Non-Toxin	Non-Allergen	0.30027

[a] Scores ≥ 0.4.
[b] Score ≥ 0.3.

Table 2
HTL epitopes for vaccine construction.

Peptide	Il-4 inducer	IFN-γ scores	Il-10 inducer	Toxicity	Antigenicity scores[a]	Allergenicity
AQARKAVGTGVRKKT	Yes	0.08	Yes	Non-Toxin	1.1876	Non-Allergen
ARKAVGTGVRKKTPE	Yes	0.02	Yes	Non-Toxin	1.2335	Non-Allergen
LNSAAQARKAVGTGV	Yes	0.41	Yes	Non-Toxin	0.8428	Non-Allergen
QAMFSGFDPWLPSLG	Yes	0.02	Yes	Non-Toxin	0.6122	Non-Allergen
QARKAVGTGVRKKTP	Yes	0.07	Yes	Non-Toxin	1.1250	Non-Allergen
SAAQARKAVGTGVRK	Yes	0.39	Yes	Non-Toxin	1.1115	Non-Allergen
VGDLNSAAQARKAVG	Yes	0.16	Yes	Non-Toxin	0.5454	Non-Allergen
AAQARKAVGTGVRKK	Yes	0.16	Yes	Non-Toxin	1.0585	Non-Allergen

[a] Scores ≥ 0.4.

Table 3
LBL epitopes for vaccine construction.

Peptide	Antigenicity scores[a]	Toxicity	Allergenicity
TGVRKKTPEPDSAEAP	0.8432	Non-Toxin	Non-Allergen
PVAAIAPSIPTPTPTP	0.9063	Non-Toxin	Non-Allergen
TGSPQGAGTLGFAGTT	0.8953	Non-Toxin	Non-Allergen
RGYEYLDLDPETGHDP	0.9490	Non-Toxin	Non-Allergen
AQARKAVGTGVRKKTP	1.1161	Non-Toxin	Non-Allergen
APQIVKANAPTAASDE	0.7838	Non-Toxin	Non-Allergen
AWLVQASANSAAMATR	0.6719	Non-Toxin	Non-Allergen

[a] Scores ≥ 0.4.

Subsequently, the 3D structure of vaccine construct was predicted utilizing SWISS-MODEL, I-TASSER, and AlphaFold2. Firstly, when employing SWISS-MODEL, the value of Global Model Quality Estimate (GMQE) and QMEANDisCo Global of the two models were too low, only 0.25 and 0.09 for model 1 and model 2, respectively (Fig. S1). As we know, GMQE and QMEANDisCo Global scores give an overall model quality measurement between 0 and 1, with higher numbers indicating higher expected quality. Therefore, we moved to the I-TASSER server to predict and acquire the top five final 3D models. Structure information, including C-score, estimated TM-score, estimated root-mean-square deviation (RMSD), number of structural decoys, and cluster density, of these 5 models is shown in Table S1. C-score is a confidence score for estimating the quality of predicted models. In general, the C-score is typically in the range of [−5, 2], where the higher value signifies the higher confidence of the model. TM-score and RMSD are usually used to measure the accuracy of structure modeling when the native structure is known. Based on that, we chose model 1 (Fig. S2A) with the C-score of −2.08, TM score of 0.47 ± 0.15, and estimated RMSD was 12.5 ± 4.3 Å. However, analyzing the Ramachandran plot of this model disclosed only 55.0% of residues in the most favored regions (Fig. S2B). Hence, we conducted structural refinement using the GalaxyRefine server and got five refined models with the structure information displayed in Table S2. Refer to Table S2, model 2 demonstrated the highest global distance test-high accuracy (GDT-HA) score of 0.9148 (the higher value, the more accurate), and the lowest root mean square deviation (RMSD) score of 0.524 (lower value indicating greater stability) (Fig. S2C). Nonetheless, when we applied the PROCHECK tool to validate the stereochemical quality of this model structure, there were only 76.0% of residues in the most favored regions (Fig. S2D). A good quality model would be expected to have over 90% of residues in the most favored regions. For that reason, we could not use the prediction result obtained from I-TASSER. Consequently, we employed the deep learning approach of AlphaFold2 and predicted the top five 3D structure models of the vaccine via ColabFold v2.3.2 based on the local distance difference test (pLDDT) ranking (Fig. S3). Among them, the 1st rank_model 2 with the best estimated reliability, was selected as the predicted structure (Fig. S3A). Nevertheless, this prediction yielded a pLDDT score below 50 (Fig. S3B) and the Ramachandran plot analysis revealed only 44.2% of residues in most favored regions (Fig. S3C), we performed further structural refinement employing the GalaxyRefine server which generated five

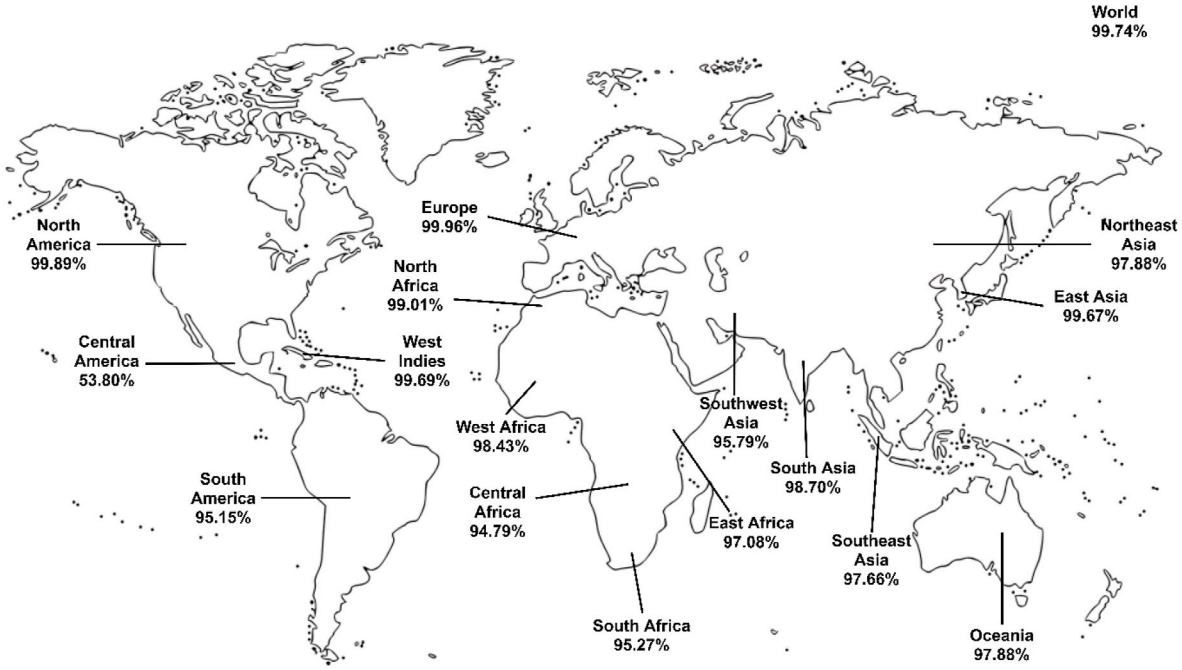

Fig. 2. Population coverage for T-lymphocytes.

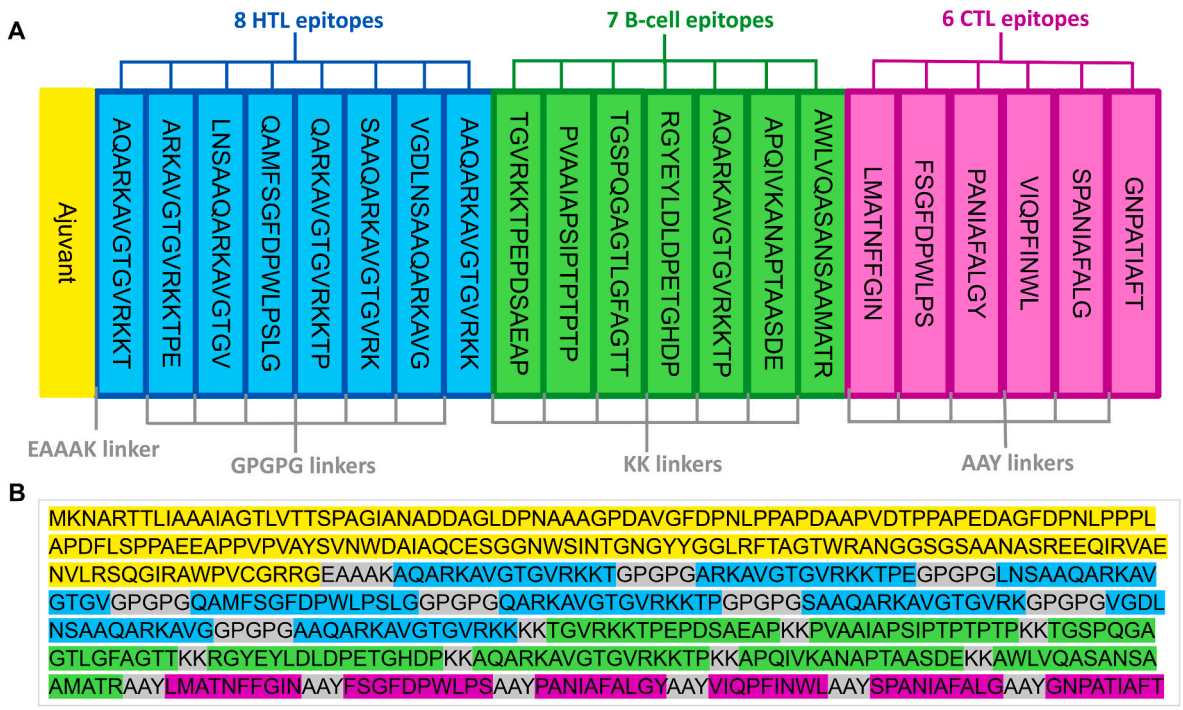

Fig. 3. Structural details of MEV. Schematic representation (A) and sequence (B) of the final vaccine construct with linkers, adjuvant, and epitopes sequentially and appropriately.

Table 4
Evaluation of antigenicity, toxicity, allergenicity, toxicity, and physicochemical properties of the vaccine construct.

Features	Assessment	Remark
Number of amino acids	534	–
Molecular weight	53797.90	Average
Total number of atoms	7580	Average
Theoretical pI	10.10	Basic nature
Total number of negatively charged residues (Asp + Glu)	33	–
Total number of positively charged residues (Arg + Lys)	64	–
Aliphatic index (AI)	62.77	Thermostable
Instability Index (II)	28.63	Stable
Estimated half-life (mammalian reticulocytes, in vitro)	30 h	Satisfactory
Estimated half-life (yeast cells, in vivo)	>20 h	Satisfactory
Estimated half-life (Escherichia coli, in vivo)	>10 h	Satisfactory
Grand Average of hydropathicity (GRAVY)	−0.351	Hydrophilic
Antigenicity	0.9363 (VaxiJen v2.0, threshold 0.4)	Antigenic
	0.9399 (ANTIGENPro)	Antigenic
Allergenicity	Non-allergen (AllerTOP v2.0)	Non-allergen
Toxicity	ToxicPred	Non-toxic

refined models and their properties (Table 5).

As the result shows in Table 5, model 1 held the highest GDT-HA score of 0.8095, the lowest RMSD score of 0.862, and the lowest MolProbity score of 1.211 (lower MolProbity value indicates better model quality), and its 3D structure was represented in cartoon in Fig. 5A. Subsequently, to validate the quality of this 3D structure after the refining process, ProSA-web and PROCHECK tool were utilized. As shown in Fig. 5B, the Z-score was determined to be −5.98. Notably, after refinement, the Ramachandran plot showed 92.8%, 5.0%, and 2.2% of residues were present in the favored, allowed, and disallowed regions, respectively (Fig. 5C). Additionally, corresponding quality scores assessed through QMEAN4 are presented in Fig. 5D with a value of −5.58. Overview, model 1 (Fig. 5A) was the best compared to other models and was selected as the vaccine candidate for further study, including molecular docking and simulations.

3.6. Molecular docking of vaccine construct with immune receptors

The molecular docking was performed for two complex systems, MEV-TLR2 and MEV-TLR4, using the ClusPro server, which generated the top 30 models for each system. Among these models, the model with the lowest negative docking score was selected as the best docked complex. Specifically, the models with −1376.3 kcal/mol (MEV-TLR2) and −1545.0 kcal/mol (MEV-TLR4) were selected for further analysis (Table 6).

A 3D representation of the surface and cartoon of the docking complex MEV-TLR2 was presented in Fig. 6A and B, and MEV-TLR4 was shown in Fig. 7A and B. In addition, interacting residues between MEV-TLR2 (Fig. 6C–F) and MEV-TLR4 (Fig. 7C–F) were visualized using PDBsum. Our results showed that 36 hydrogen bonds and 10 salt bridges were formed between the residues of MEV and two chains of TLR2 (Fig. 6E and F). Similarly, 68 hydrogen bonds and 13 salt bridges were formed between the residues of the vaccine and two chains of TLR4 (Fig. 7E and F). Based on these findings, MEV had excellent performance in strongly binding to TLR2 and 4 to produce a strong immune response.

3.7. Molecular dynamic simulations of vaccine with immune receptors

In order to obtain the stability and dynamic behavior of the interactions between MEV with TLR2 and TLR4 receptors during 100 ns simulation, statistical parameters such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), the radius of gyration (Rg), and solvent accessible surface area (SASA) were examined in triplicate (Fig. 8). After 50 ns, the RMSD values for both remained stable, indicating that the complexes maintained a comparatively stable structure, according to Fig. 8A and B, the complexes consistently retained moderate structural stability. During the simulation, the

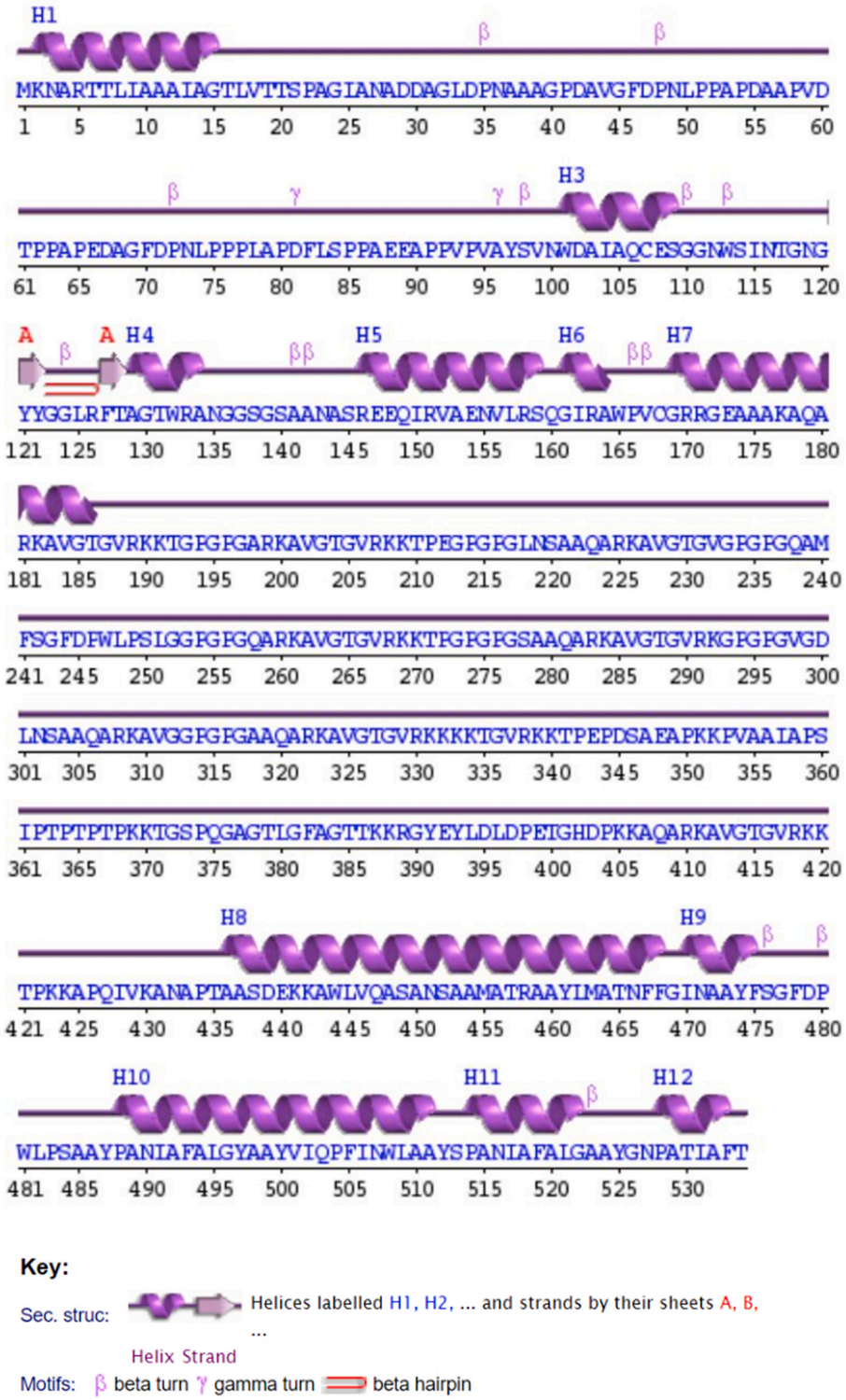

Fig. 4. Secondary structure prediction of the vaccine construct representation in schematic "wiring diagram" including strands (pink arrows), helices (purple springs), and other motifs in red (e.g., β-hairpins, γ-turns, etc).

average RMSD values of the MEV-TLR2 and MEV-TLR4 complexes were 1.767 ± 0.127 and 1.634 ± 0.151 nm, respectively.

Referring to the trajectories of Fig. 8C and D, the MEV regions displayed the most significant fluctuations and an increasing trend in RMSF values, suggesting that the MEV regions were more dynamic or flexible than the TLRs regions. Besides, the Rg offers valuable information regarding the tendency of complex structures to expand during MDs. For the MEV-TLR2 complex (Fig. 8E), the Rg values slightly increased from the start of the simulations until 10 ns, indicating compaction or tightening of the complex. The Rg values gradually decreased, then stable after 50 ns until they reached 100 ns. For the MEV-TLR4 complex (Fig. 8F), the Rg gradually reduced from the beginning of the simulation up to 40 ns, then stabilized until it reached 100 ns. Throughout the simulation, the complexes maintained an overall relatively compact

Table 5
Structure information obtained from GalaxyWEB.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	3.593	27.4	7.7	45.3
MODEL 1	0.8095	0.862	1.211	1.9	0.3	96.2
MODEL 2	0.8071	0.885	1.295	2.3	0.0	95.9
MODEL 3	0.8062	0.882	1.454	2.8	0.0	94.4
MODEL 4	0.8038	0.883	1.376	2.2	0.3	94.4
MODEL 5	0.7949	0.911	1.214	1.4	0.3	95.1

structure, as indicated by the average Rg value of 4.889 ± 0.187 and 4.663 ± 0.044 nm for MEV-TLR2 and MEV-TLR4, respectively.

In addition, Fig. 8G and H shows the average SASA values of MEV-TLR2 and MEV-TLR4 complex were 889.117 ± 21.779 and 915.881 ± 7.935 nm^2, respectively. The SASA values gradually decreased throughout the simulation, indicating that the complexes got more compact or less exposed to solvent. Significant structural rearrangement may have occurred during the beginning of simulation until 40 ns (for MEV-TLR2) and 30 ns (MEV-TLR4) period, as evidenced by the significant and quick decrease in SASA during this time frame. After that, both were stable until the end of the simulation period.

3.8. Immune responses induced by the vaccine

The immunological simulation data was produced by the C-IMMSIMserver following three successive injections of the vaccine candidate. Fig. 9A shows that a vaccine's initial exposure produces a relatively low immunoglobulin response, whereas a subsequent exposure produces an increased immunoglobulin response. Additionally, Fig. 9A shows that immunoglobulins such as IgM and IgG are more abundant than other immunoglobulins such as IgM, IgG1, and IgG1 + IgG2. Details of the cytokine levels, especially IFN-g, increased significantly and were above 400,000 ng/ml, which are visualized in Fig. 9B, while Fig. 9C indicates responses of B lymphocytes. Our simulation study also identified helper T-cell and cytotoxic T-cell responses (Fig. 9D and E), confirming the realistic character of the server-predicted immune response because helper T-cell activity is crucial for activating B-

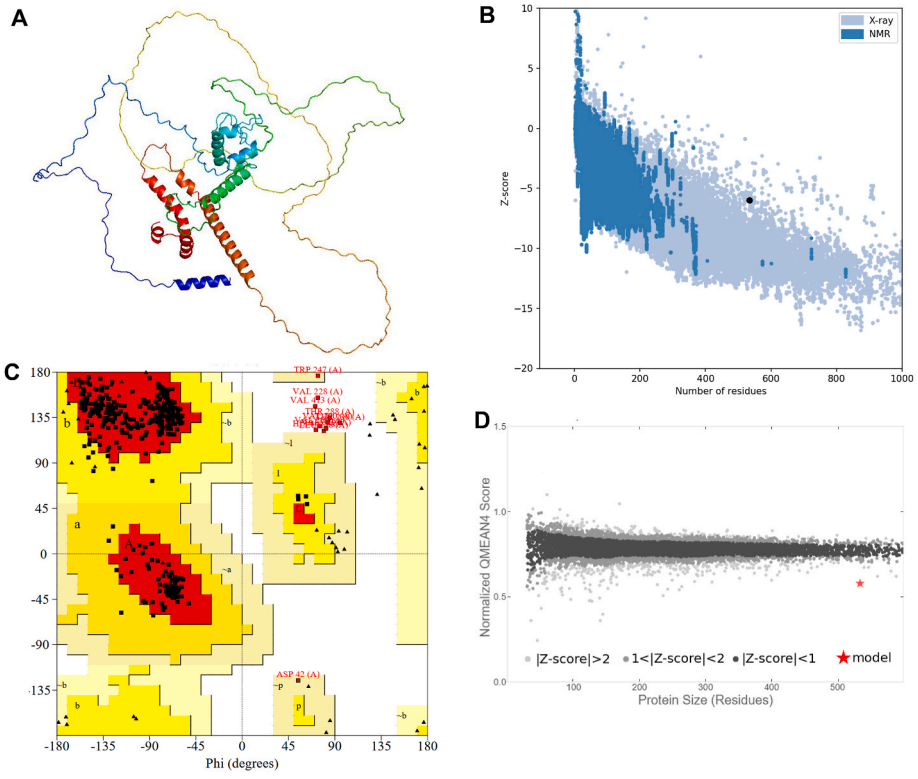

Fig. 5. Prediction, refinement, and validation of the tertiary structure of the vaccine. (A) The 3D structure representation in the cartoon by PyMOL. (B) The Z-score was obtained from ProSA-Web. (C) Ramachandran plot gained from PROCHECK. (D) Normalized QMEAN score composed of four statistical potential terms (QMEAN4) of the vaccine.

Table 6
Molecular docking of the vaccine with TLR2 and TLR4.

Target (PDB ID)	Center (kcal/mol)	Lowest energy (kcal/mol)		No. of interface residues	Interface area (Å^2)	No. of hydrogen bonds	No. of salt bridges
TLR2 (2Z7X)	−1079.1	−1376.3	Chain A	TLR2A: 8 Vaccine: 8	TLR2A: 414 Vaccine: 406	13	6
			Chain B	TLR2B: 46 Vaccine: 35	TLR2B: 1716 Vaccine: 1954	23	4
TLR4 (3FXI)	−1441.9	−1545.0	Chain A	TLR4A: 48 Vaccine: 46	TLR4A: 2298 Vaccine: 2278	44	8
			Chain B	TLR4B: 39 Vaccine: 30	TLR4B: 1486 Vaccine: 1641	24	5

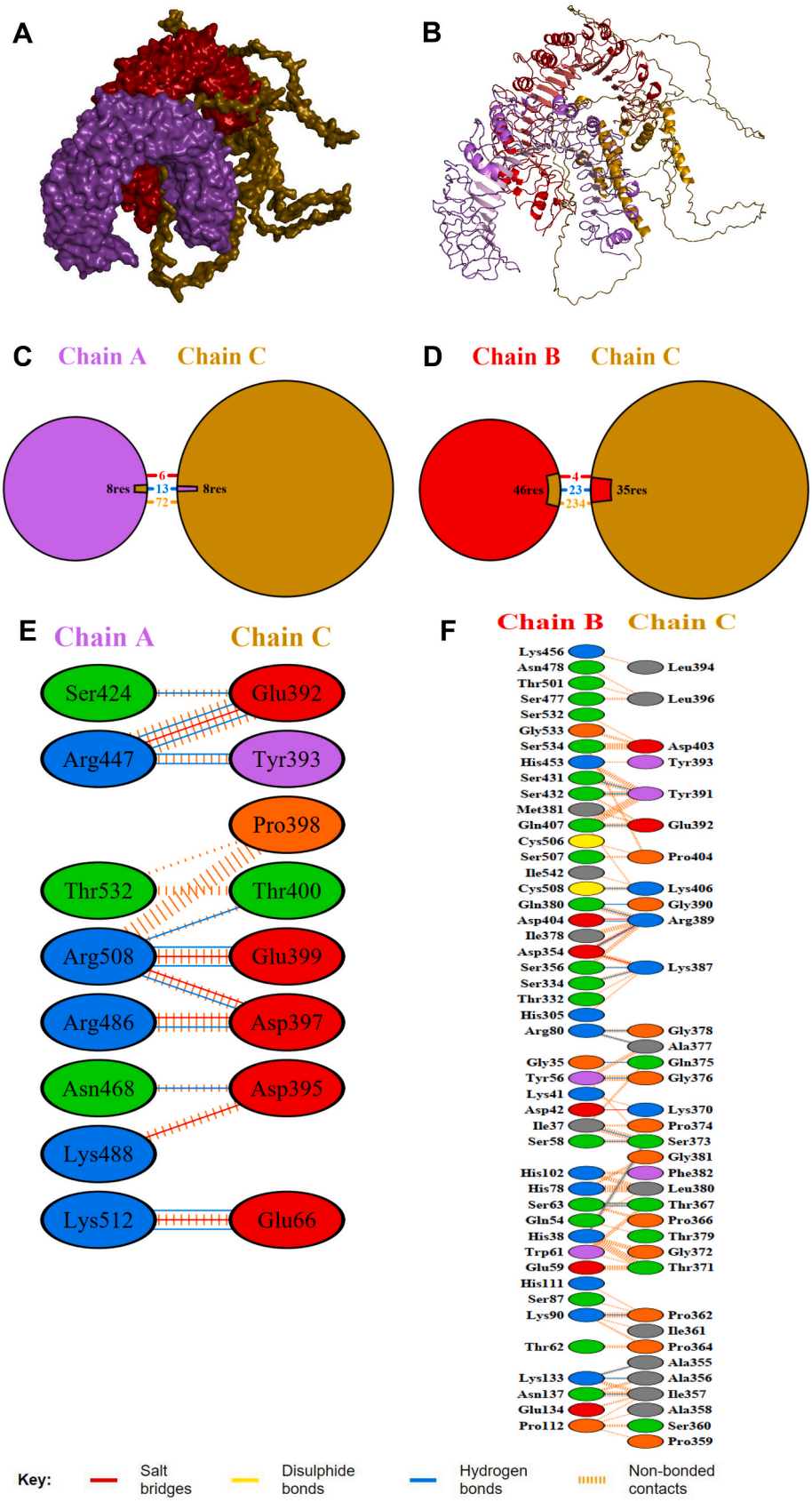

Fig. 6. Molecular docking between TLR2 chains A (purple) and B (red) with the vaccine construct (chain C – dark yellow). (A–B) Three-dimensional representation of the docking complex in surface and cartoon, respectively. (C–D) Schematic diagram of interactions between TLR2 and the vaccine. (E–F) Residue interactions between TLR2 and the vaccine construct. Salt-bridges (red lines), hydrogen bonds (blue lines), and non-bonded contacts (orange dashed line) between residues on either side of the vaccine-receptor interface.

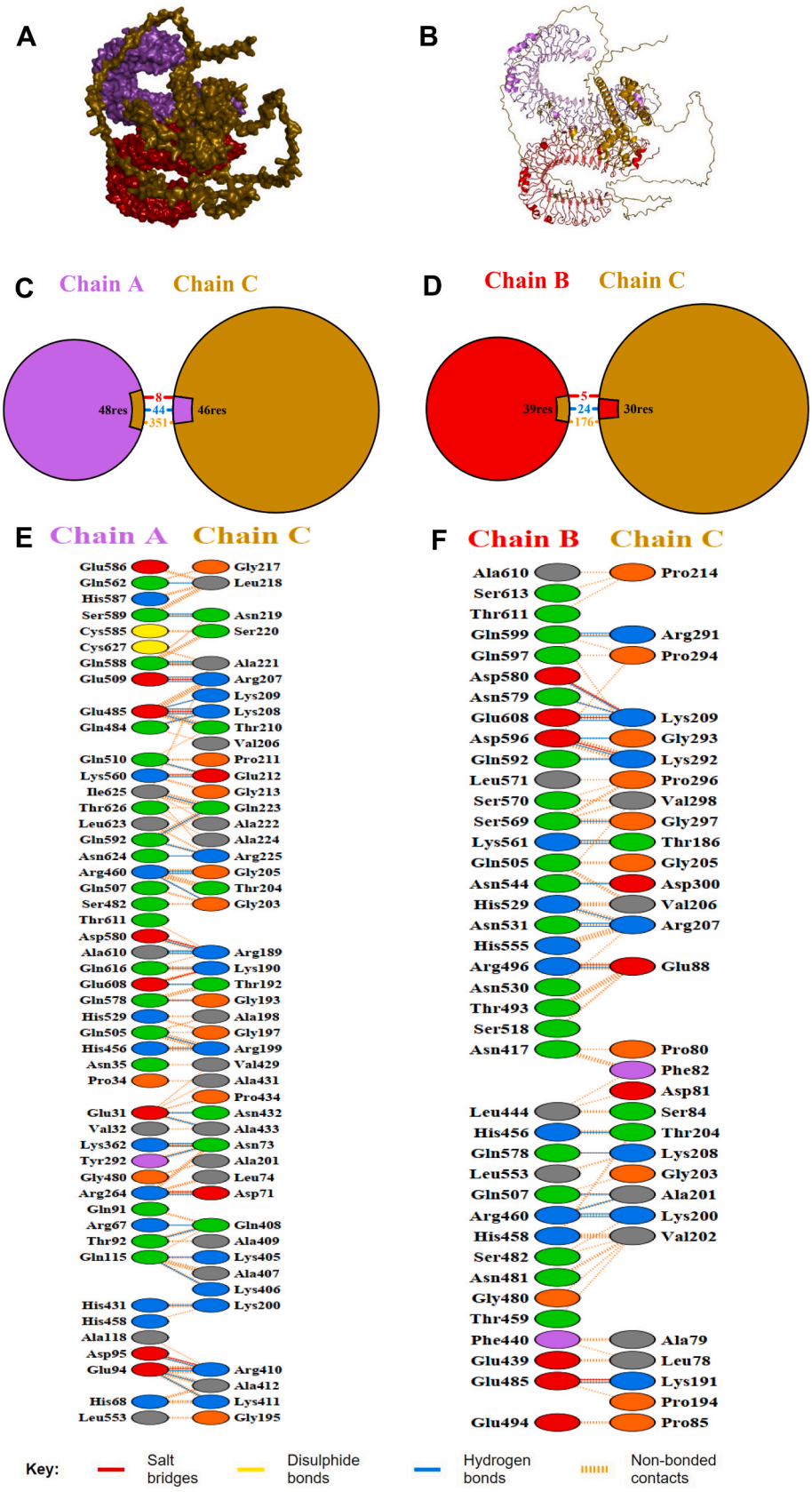

Fig. 7. Molecular docking between TLR4 chains A (purple) and B (red) with the vaccine construct (chain C – dark yellow). (A–B) Three-dimensional representation of the docking complex in surface and cartoon, respectively. (C–D) Schematic diagram of interactions between TLR4 and the vaccine. (E–F) Residue interactions between TLR4 and the vaccine construct. Salt-bridges (red lines), hydrogen bonds (blue lines), and non-bonded contacts (orange dashed line) between residues on either side of the vaccine-receptor interface.

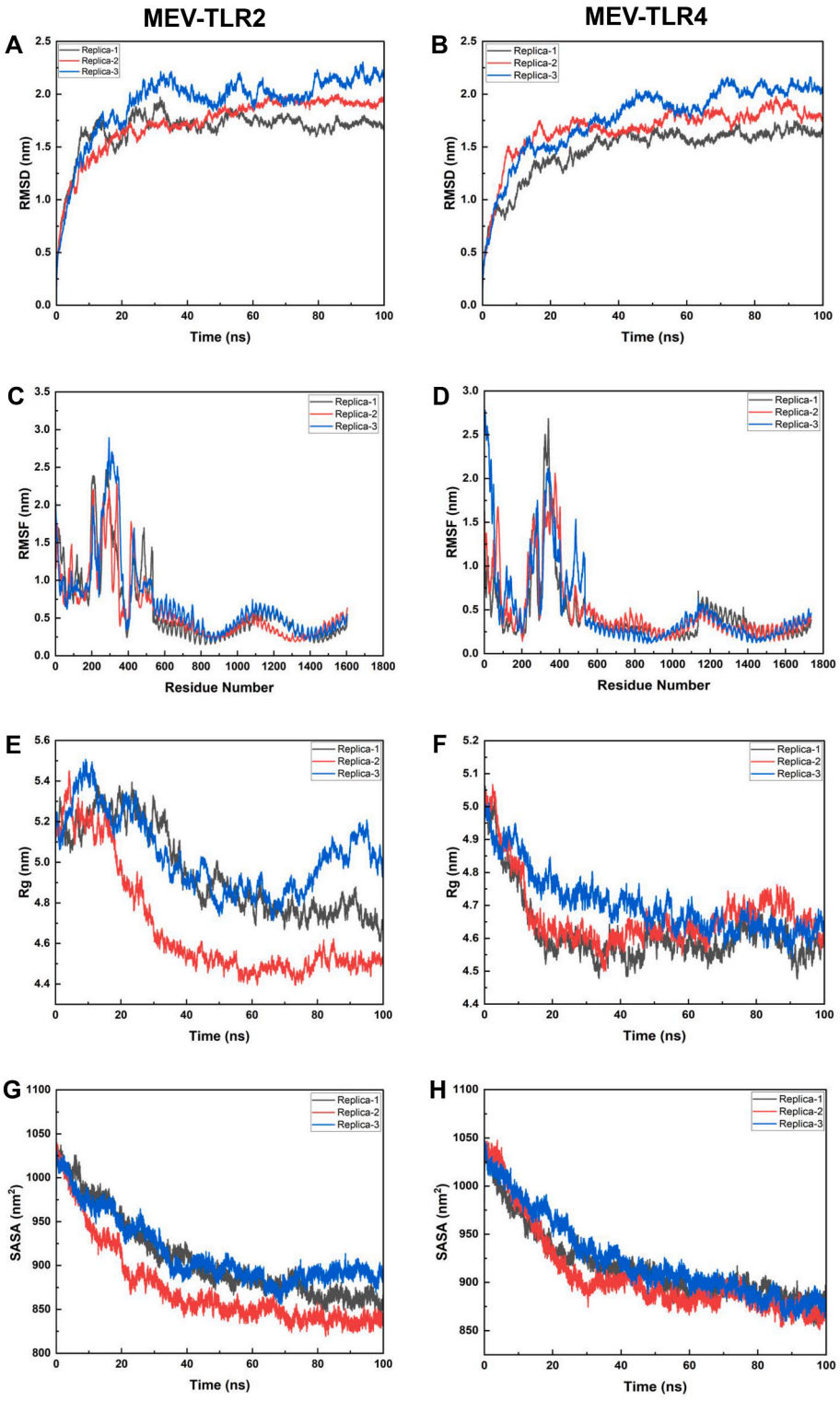

Fig. 8. MD simulations of the docking complexes between MEV and TLR2/4. (A–B) Root mean square deviation plot for MEV-TLR2 and MEV-TLR4 complex, respectively. (C–D) Root mean square fluctuation of MEV-TLR2 and MEV-TLR4 complex, respectively. (E–F) Radius of gyration analysis between MEV-TLR2 and MEV-TLR4 complex, respectively. (G–H) Solvent-accessible surface area of MEV-TLR2 and MEV-TLR4 complex during MD simulations, respectively.

cells. During the self-memorization process following pathogen exposure, memory cells play a crucial role in preventing and regulating viral infection and reinfection. The successful injection of the vaccine candidate resulted in an increase in the regulatory components of the immune system, including DC cells, macrophages, NK cells, interleukins, and cytokines (Fig. 9C–H). These results imply that the MEV is a highly effective next-generation vaccine based on peptides that can stimulate a robust immune response against MTB infection.

4. Discussion

Despite advances in vaccine technology, there are still no vaccines

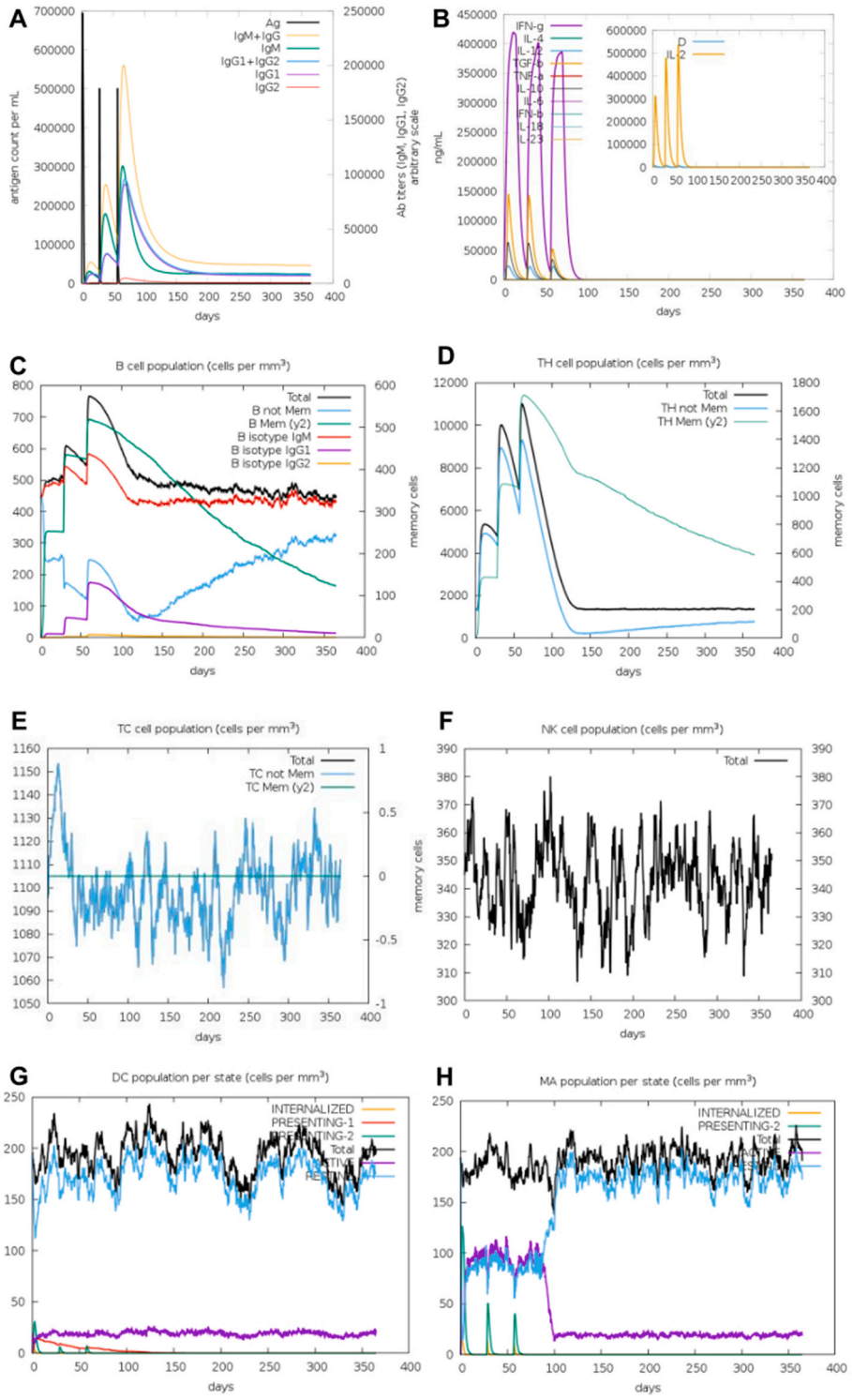

Fig. 9. The innate and adaptive immune responses induced by the MEV in the C-IMMSIMserver. (A) Immunoglobulin responses upon exposure to the vaccine. (B) Concentration of cytokines and interleukins. (C) B cell population. (D) T helper cell population. (E) T cytotoxic cell population. (F) Behavior of the population of Natural Killer cells. (G) Behavior of the population of Dendritic cells. (H) The population of macrophages after vaccination.

against some infectious diseases, including tuberculosis. The infectious pathogens underlying these diseases evade and alter host immune responses, making vaccine development difficult. In this study, we aim to design a MEV candidate against tuberculosis to contribute towards the End MTB Strategy. Several vaccines for MT have been developed to provide possible candidates for novel vaccine designs [15,18,19,59]. While a few new proteins with antigenic qualities were chosen for the current study to identify epitopes and develop the vaccine, most known antigenic proteins or proteins found in exosome vesicles were used in these previous studies to uncover antigenic epitopes. Besides, although many vaccine candidates are available, each vaccine employs different algorithms and features. The in vitro and in vivo vaccination production process is far more complex, expensive, and time-consuming than the MEV. A range of laboratory medical studies are also required for the

final epitope selection. A computational technique that saves time and predicts a peptide or epitope sequence that might be used to make a lab-based MEV is called in silico methodology.

The present study focuses on a PPE family protein, Rv0256c (PPE2), which induced a strong B cell response in tuberculosis patients [25]. We carried out a computer analysis using a range of immunoinformatics techniques to find 21 potent epitopes that could be helpful in the fight against tuberculosis. Therefore, using this method, we may reduce the cost and length of wet lab investigations. Finally, we speculate that the designed vaccine is extracellular, highly immunogenic, antigenic, nontoxic, and nonallergic; as such, it could be a promising MEV candidate for MTB based on computational analysis. Additional wet-lab validation is required to confirm the 21 epitopes' effectiveness as MEV in this study.

Firstly, we retrieved the protein sequence and evaluated antigenicity and allergenicity. Then, the immunoinformatic techniques were used to screen and construct potential epitopes from the sequence of Rv256c protein for B and T cells. The antigenicity, allergenicity, and several physiochemical properties of the developed multi-epitope vaccination were then evaluated. The vaccine construct contains 534 amino acids, comprising six cytotoxic T lymphocyte, eight helper T lymphocyte, and seven linear B lymphocyte epitopes, along with adjuvants and linkers. The antigenicity score of MEV predicted by VaxiJen 2.0 and ANTIGENpro was 0.9363 and 0.9399, respectively. The MEV was predicted as stable and thermostable with an instability index and aliphatic index of 28.64 and 62.77, respectively. Afterwards, tertiary structure prediction, refinement, and validation were conducted. Ramachandran plot analysis reveals that 97.8% of the amino acid residues were in the most favored and allowed regions.

Subsequently, molecular docking and MD simulations were utilized in evaluating the complex stability. There were a total of 46 interaction sites in MEV-TLR2 and 81 interaction sites in MEV-TLR4 (Figs. 6 and 7). These results showed that the interactions between MEV and TLR2, 4 were strong, and the docking effect was good, especially, the MEV-TLR 4 complex. Triplicate MD simulations were calculated to confirm the poses found by docking results. The MD calculation lets us establish if MD finds the most populated cluster from docking. The static view provided by docking should be verified by using MD. In this study, a MD simulation could be helpful to confirm if the primary contacts found will be maintained during the MD to present more reliable results. In general, combining two in silico techniques (docking, MD) could improve the reliability of the results. We need to ensure that all the system's chemical and physical properties have reached an equilibrium where their averages no longer change as a function of time. A simple way to test this is by measuring the RMSD of the backbone concerning the start (Fig. 8A and B). For the RMSD, the average is taken over the particles, giving time-specific values; for RMSF, the latter is averaged over time, giving each particle (residue) value. RMSF is a simple tool to measure the rigidity of the polypeptide chain. It calculates the deviations of the C-alpha atom's coordinates from their average position. The flexibility pattern reflects the location of secondary structure elements in the protein structure (Fig. 8C and D). Besides, we studied the compactness of the receptors TLR2 and TLR4 interaction with the vaccine using Rg. The receptor remains compact, and no unusual folding or unfolding was observed throughout the 100 ns. The vaccine with TLR4 is more tightly packed than TLR2 (Fig. 8E and F). Furthermore, we calculated the total solvent-accessible surface area to understand the system's SASA of the binding region. We observed that initially, the area of structures contact by the solvent molecules was higher, and it decreased gradually with time and showed stable values, showing significant interaction can be assumed between the vaccine and TLR2, 4. In particular, the complex between the vaccine and TLR4 showed more stability than TLR2 (Fig. 8G and H). To sum up, our molecular docking studies and triplicate MD simulations revealed superior interactions and stability of the vaccine with the TLR4 complex compared to the TLR2 complex. This aligns with the outcomes of a previous investigation [60]. Alderwick et al. in that study, demonstrated that human immune cells that were infected with MTB express more TLR4, which played a role in the interaction with MTB and activated TLR4 related signaling, which in turn enhanced Th2 signaling and led to the development of tuberculosis disease.

Finally, we performed in silico immune simulation to characterize the immunogenicity and immune response of the vaccine. Fig. 9 shows the outcomes of the immunological simulation after the MEV was administered. After the host immune system was exposed to MEV several times, there was a discernible rise in secondary and tertiary antibody levels, which were higher than primary antibody detection levels. As a result, there was a quick drop in antigen concentration, which suggests quick clearance (Fig. 9A). Interestingly, T-cells were found to have increased initially but then somewhat decreased. Moreover, cytokine levels, especially IFN-g, increased significantly and were above 400,000 ng/ml (Fig. 9B). Fig. 9C–H shows the distribution of immune cell populations in different states. These immune cell populations showed a notable overall increase, exemplified by the maturation of memory cells. These results demonstrate that the vaccine designs can provide immunity against the MTB and strongly suggest the formation of immunological memory.

5. Conclusions

Despite the availability of the BCG vaccine, tuberculosis persists as a significant global public health threat, which underscores the critical need for a prophylactic and immunotherapeutic vaccine against MTB. The application of an in silico approach represents the initial crucial step in the development of such a vaccine. In this study, we achieved success by designing an MEV against MTB, utilizing immunoinformatics and computational methodologies. The MEV, composed of highly antigenic and immunogenic epitopes, demonstrated safety, stability, and robust interactions with receptors TLR2 and TLR4. Notably, our molecular docking studies and triplicate MD simulations revealed enhanced interactions and stability of the vaccine with the TLR4 complex compared to TLR2. The executed immunological simulation further supported our hypothesis by demonstrating a robust immune response induced by the vaccination. This investigation not only establishes a scientific foundation for understanding the mechanisms of MEV against TB but also provides valuable theoretical support for the ongoing development of TB vaccines. Additionally, our contributions align with the global initiative of the End MTB Strategy.

Funding

This work was supported by the BK21 FOUR Program of the Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea.

Data availability

All data generated or analyzed during this study are included in this article.

CRediT authorship contribution statement

Truc Ly Nguyen: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, preparation, Writing – review & editing. **Heebal Kim:** Conceptualization, Resources, Writing – original draft, preparation, Writing – review & editing, Supervision, Both authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are thankful to the laboratory of Bioinformatics and Population genetics, Seoul National University, for providing the computing infrastructure to implement and execute the research work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2024.03.010.

References

[1] Chakaya J, et al. The WHO Global Tuberculosis 2021 Report - not so good news and turning the tide back to End TB. Int J Infect Dis : IJID : official publication of the International Society for Infectious Diseases 2022;124(Suppl 1):S26–9.
[2] Zimmer AJ, et al. Tuberculosis in times of COVID-19. J Epidemiol Community 2022;76(3):310–6.
[3] Li JM, Zhu DY. Therapeutic DNA vaccines against tuberculosis: a promising but arduous task. Chin Med J 2006;119(13):1103–7.
[4] Scriba TJ, et al. Vaccination against tuberculosis with whole-cell mycobacterial vaccines. J Infect Dis 2016;214(5):659–64.
[5] Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cell Mol Immunol 2018;15(2):182–4.
[6] Khairkhah N, et al. Immunological investigation of a multiepitope peptide vaccine candidate based on main proteins of SARS-CoV-2 pathogen. PLoS One 2022;17(6): e0268251.
[7] Mao Y, et al. Designing a multi-epitope vaccine against Peptostreptococcus anaerobius based on an immunoinformatics approach. Synthetic and Systems Biotechnology 2023;8(4):757–70.
[8] Nguyen TL, Kim H. Designing a multiepitope vaccine against eastern equine encephalitis virus: immunoinformatics and computational approaches. ACS Omega 2024;9(1):1092–105.
[9] Gong W, et al. Peptides-based vaccine MP3RT induced protective immunity against Mycobacterium tuberculosis infection in a humanized mouse model. Front Immunol 2021;12:666290.
[10] Nayak SS, Sethi G, Ramadas K. Design of multi-epitope based vaccine against Mycobacterium tuberculosis: a subtractive proteomics and reverse vaccinology based immunoinformatics approach. J Biomol Struct Dyn 2023;41(23):14116–34.
[11] Bellini C, et al. Design and characterization of a multistage peptide-based vaccine platform to target Mycobacterium tuberculosis infection. Bioconjugate Chem 2023; 34(10):1738–53.
[12] Jiang F, et al. PP19128R, a multiepitope vaccine designed to prevent latent tuberculosis infection, induced immune responses in silico and in vitro assays. Vaccines 2023;11(4):856.
[13] Cheng P, et al. Bioinformatics analysis and consistency verification of a novel tuberculosis vaccine candidate HP13138PB. Front Immunol 2023;14.
[14] Andongma BT, et al. In silico design of a promiscuous chimeric multi-epitope vaccine against Mycobacterium tuberculosis. Comput Struct Biotechnol J 2023;21: 991–1004.
[15] Ruaro-Moreno M, et al. Design of a multi-epitope vaccine against tuberculosis from Mycobacterium tuberculosis PE_PGRS49 and PE_PGRS56 proteins by reverse vaccinology. Microorganisms 2023;11(7):1647.
[16] Cheng P, Wang L, Gong W. In silico analysis of peptide-based biomarkers for the diagnosis and prevention of latent tuberculosis infection. Front Microbiol 2022;13.
[17] Khan MT, et al. Multi-epitope vaccine against drug-resistant strains of Mycobacterium tuberculosis: a proteome-wide subtraction and immunoinformatics approach. Genomics & Informatics 2023;21(3):e42.
[18] Bibi S, et al. In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. Sci Rep 2021;11(1).
[19] Sharma R, et al. An immunoinformatics approach to design a multi-epitope vaccine against Mycobacterium tuberculosis exploiting secreted exosome proteins. Sci Rep 2021;11(1).
[20] Pal R, Ghosh S, Mukhopadhyay S. Moonlighting by PPE2 protein: focus on mycobacterial virulence. J Immunol 2021;207(10):2393–7.
[21] Bhat KH, et al. The PPE2 protein of Mycobacterium tuberculosis translocates to host nucleus and inhibits nitric oxide production. Sci Rep 2017;7. 39706-39706.
[22] Bhat KH, et al. PPE2 protein of*Mycobacterium tuberculosis*may inhibit nitric oxide in activated macrophages. Ann N Y Acad Sci 2013;1283(1):97–101.
[23] Srivastava S, et al. Mycobacterium tuberculosis PPE2 protein Interacts with p67phox and inhibits reactive oxygen species production. J Immunol 2019;203(5): 1218–29.
[24] Pal R, Mukhopadhyay S. PPE2 protein of Mycobacterium tuberculosis affects myeloid hematopoiesis in mice. Immunobiology 2021;226(1):152051.
[25] Abraham PR, et al. Mycobacterium tuberculosis PPE protein Rv0256c induces strong B cell response in tuberculosis patients. Infect Genet Evol 2014;22:244–9.
[26] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinf 2007;8. 4-4.
[27] Dimitrov I, et al. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 2014;30(6):846–51.
[28] Fleri W, et al. The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. Front Immunol 2017;8. 278-278.
[29] Al Tbeishat H. Novel in Silico mRNA vaccine design exploiting proteins of M. tuberculosis that modulates host immune responses by inducing epigenetic modifications. Sci Rep 2022;12(1).
[30] Dhanda SK, Vir P, Raghava GPS. Designing of interferon-gamma inducing MHC class-II binders. Biol Direct 2013;8. 30-30.
[31] Dhanda SK, et al. Prediction of IL4 inducing peptides. Clin Dev Immunol 2013; 2013. 263952-263952.
[32] Nagpal G, et al. Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. Sci Rep 2017;7. 42851-42851.
[33] Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Struct, Funct, Bioinf 2006;65(1):40–8.
[34] Dimitrov I, et al. AllerTOP v.2—a server for in silico prediction of allergens. J Mol Model 2014;20(6).
[35] Gupta S, et al. In silico approach for predicting toxicity of peptides and proteins. PLoS One 2013;8(9):e73957.
[36] Bui H-H, et al. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinf 2006;7(1):153.
[37] Gonzalez-Galarza F, Faviel, et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. Nucleic Acids Res 2019;48(D1):D783–8.
[38] Hajighahramani N, et al. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus. Infect Genet Evol 2017;48:83–94.
[39] Chawla M, et al. Immunoinformatics-aided rational design of a multi-epitope vaccine targeting feline infectious peritonitis virus. Front Vet Sci 2023;10: 1280273.
[40] Lee SJ, et al. A potential protein adjuvant derived from Mycobacterium tuberculosis Rv0652 enhances dendritic cells-based tumor immunotherapy. PLoS One 2014;9(8):e104351.
[41] Ilinskaya AN, Dobrovolskaia MA. Understanding the immunogenicity and antigenicity of nanomaterials: past, present and future. Toxicol Appl Pharmacol 2016;299:70–7.
[42] Magnan CN, et al. High-throughput prediction of protein antigenicity using protein microarray data. Bioinformatics 2010;26(23):2936–43.
[43] Gasteiger E, et al. Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Humana Press; 2005. p. 571–607.
[44] Laskowski RA, et al. PDBsum: structural summaries of PDB entries. Protein Sci 2018;27(1):129–34.
[45] Hutchinson EG, Thornton JM. PROMOTIF—a program to identify and analyze structural motifs in proteins. Protein Sci 1996;5(2):212–20.
[46] Waterhouse A, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46(W1):W296–303.
[47] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinf 2008;9. 40-40.
[48] Mirdita M, et al. ColabFold - making protein folding accessible to all. Cold Spring Harbor Laboratory; 2021.
[49] Jumper J, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9.
[50] Ko J, et al. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res 2012;40(W1):W294–7.
[51] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35(Web Server issue):W407–10.
[52] Laskowski RA, et al. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26(2):283–91.
[53] Sharma AD, et al. T cell epitope based vaccine design while targeting outer capsid proteins of rotavirus strains infecting neonates: an immunoinformatics approach. J Biomol Struct Dyn 2023:1–19.
[54] Akhtar N, et al. Secreted aspartyl proteinases targeted multi-epitope vaccine design for Candida dubliniensis using immunoinformatics. Vaccines 2023;11(2):364.
[55] Kozakov D, et al. The ClusPro web server for protein–protein docking. Nat Protoc 2017;12(2):255–78.
[56] Abraham MJ, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1–2: 19–25.
[57] Rapin N, Lund O, Castiglione F. C-Immsim 10.1 server. PLoS Pathog 2012;8.
[58] Dey J, et al. Designing of multi-epitope peptide vaccine against Acinetobacter baumannii through combined immunoinformatics and protein interaction–based approaches. Research: Immunologic; 2023.
[59] Albutti A. An integrated computational framework to design a multi-epitopes vaccine against Mycobacterium tuberculosis. Sci Rep 2021;11(1).
[60] Alderwick LJ, et al. The mycobacterial cell wall—peptidoglycan and arabinogalactan. Cold Spring Harbor Perspectives in Medicine 2015;5(8):a021113.