

A Multi-Epitope Vaccine Candidate with Predicted Broad Genotype Coverage Against *Toxoplasma gondii* Designed by Reverse Vaccinology and Immunoinformatics Strategies

Mostafa Zaher

mustafaraafat41@gmail.com

Animal Health Research Institute <https://orcid.org/0000-0002-8267-2702>

Dalia M. El-Husseini

Animal Health Research Institute

Research Article

Keywords: Multi-epitope vaccine, Immunoinformatics, *Toxoplasma gondii*, Reverse vaccinology, Zoonotic parasitic disease

Posted Date: October 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11527358/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Background

Toxoplasma gondii (*T. gondii*) is a globally prevalent zoonotic pathogen with major veterinary importance, causing significant livestock reproductive losses and facilitating zoonotic transmission to humans. Despite extensive efforts, there is no available efficient vaccine except for Toxovax, which is only approved for sheep and suffers from some limitations. The parasite's complex multistage life cycle and strain diversity present major obstacles for vaccine development. Traditional vaccine strategies, including killed, attenuated, and subunit vaccines, provide only limited protection, in addition to posing potential translational drawbacks. Multi-epitope vaccine (MEV) approaches have emerged as a promising alternative that combines conserved immunodominant epitopes, enabling a robust immune response. Our study aimed to computationally design and evaluate a novel MEV candidate capable of eliciting predicted broad genotypic coverage against *T. gondii*. It thus, serves as a One Health candidate intended for veterinary disease control, which consequently serves the public health by reducing foodborne zoonotic transmission risk and translational potential for cross-species experimental validation.

Methods

Eighteen *T. gondii* proteins expressed in different life stages of multiple genotypes were screened for antigenicity, allergenicity, and epitope conservancy. Selected proteins were analyzed to predict MHC-I, MHC-II, and B-cell epitopes using stringent thresholds. Followed by evaluating the candidate epitopes for their cytokine induction potential and theoretical population coverage. Selected epitopes were then assembled into constructs and assessed for physicochemical properties, secondary and tertiary structure stability, immune response simulation, and molecular docking to TLR2 and TLR4. The candidate construct was subjected to complex stability analysis through molecular dynamics (MD) simulation. Codon optimization and in silico cloning were then performed for potential *E. Coli* expression.

Results

Forty-three conserved and non-allergenic epitopes were selected and constituted different constructs. Out of which, MEV-2A demonstrated high predicted antigenicity, 100 % hypothetical worldwide population coverage, Th1-skewed cytokine induction, high stability, and computational binding affinity towards TLR2 and TLR4, suggesting that MEV-2A is a promising candidate for further validation.

Conclusion

MEV-2A is a computationally designed and optimized multi-epitope vaccine targeting multi-stages and multi-genotypes of *T. gondii*, aimed to serve as a One Health candidate for livestock disease control and human zoonotic risk reduction, which warrants further experimental *in vitro* and *in vivo* validations.

Clinical trial number: not applicable.

Introduction

Toxoplasmosis, a disease caused by the obligate intracellular protozoan parasite *Toxoplasma gondii* (*T. gondii*), is a widespread zoonotic infection that affects approximately 30% to 33% of the global human population, as well as numerous warm-blooded animals. It is considered one of the most critical, life-threatening diseases with global medical, veterinary, and economic significance [1, 2]. In livestock, *T. gondii* infection causes reproductive loss, food safety concerns, and reduces productivity, reinforcing its importance within a One Health framework [3]. The infection rates are notably higher in developing and underdeveloped countries than in developed ones because prevalence is affected by climate, stray cat populations, and socioeconomic conditions [4]. *T. gondii* exhibits significant genetic diversity, leading to the presence of diverse strains and isolates. They are structured into global major clonal lineages (Type I, II, III, and atypical recombinants) based on specific molecular markers and virulence profiles [5]. Modern genomic surveillance reveals a highly diverse *T. gondii* macro-population structured into at least six major clades and 15 to 16 haplogroups, in addition to extensive array of atypical, non-archetypal recombinant genotypes [6]. To achieve effective control, a successful vaccine must encompass antigens conserved across these diverse genotypes to ensure broad protective coverage in both human and animal hosts.

From a clinical vaccinology perspective, the complex transitions between these distinct life-cycle stages dictate separate biological endpoints. While ideal vaccine candidate would aim for absolute infection prevention (sterile immunity), practical intervention strategies against complex protozoans frequently focus on separate, overlapping targets[7]. The life cycle of *T. gondii* involves feline definitive hosts, where sexual reproduction occurs, causing the shedding of numerous environmentally resistant oocysts. Upon sporulation, these infect intermediate hosts, triggering tachyzoites' replication and dissemination, followed by the conversion to slow-replicating bradyzoites that form resilient tissue cysts in neural and muscle tissues, causing lifelong chronic infection. The transmission between intermediate hosts may occur horizontally *via* food consumption or vertically *via* transplacental [8].

The parasite's complex life cycle and multiple variants impart challenges for disease control and vaccine design. Therapeutics, including pyrimethamine and sulfadiazine, suffer from limited efficacy against latent bradyzoite cysts, high cost, and have various side effects [9]. A One Health approach to develop an effective vaccine to address toxoplasmosis has become a compelling and feasible opportunity, especially with the growing understanding of parasite transmission pathways, epidemiology, and primary risk populations bridging human medicine, veterinary health, and environmental safety [10].

Despite decades of research, no effective vaccine against toxoplasmosis exists except for Toxovax, which was formulated from a live attenuated S48 *T. gondii* strain and is only approved exclusively for sheep to reduce reproductive losses [11]. In addition, it risks reversion to virulence and has a short shelf life [12]. Other vaccine strategies were explored, including excretory–secretory antigens, subunit, DNA, multi-antigenic, mRNA, carbohydrate, and exome-based vaccines. However, these have frequently failed to confer broad and long-lasting immunity [13]. A primary limitation is their reliance on a single strain or

specific antigenic target, which may not represent the current circulating strains of the same or other genotypes, failing to elicit a proper immune response. These drawbacks underscore the crucial need to develop vaccines that can elicit immune responses against different *T. gondii* genotypes, combined with a compelling delivery system to improve efficacy [13].

Reverse vaccinology and immunoinformatics enabled rational identification and evaluation of immunodominant B- and T-cell epitopes to induce robust humoral and cellular immunity. This reduces time, effort, and cost invested in experimental epitope screening, consequently accelerating vaccine development research. Many studies have designed multi-epitope vaccines (MEVs) for various pathogens, such as Foot-and-mouth disease virus [14], Dengue virus [15], Leishmania Parasite [16], and *Acinetobacter baumannii* [17]. For complex intracellular parasites like *T. gondii*, this approach is valuable for prioritizing the CD8 + cytotoxic response crucial for controlling the intracellular parasite. Previous studies on *T. gondii* investigated immunization with MEV showed prolonged survival time [18]. SAG-related sequence (SRS), bradyzoite antigen 1 (BAG1), rhomboid proteins (ROMs), dense granule proteins (GRAs), microneme proteins (MICs), profilin, heat-shock proteins (HSPs), calcium-dependent protein kinases (CDPKs), and rho-try proteins (ROPs) have been recognized as the most promising vaccine targets [19–21]. Identification of immunodominant epitopes from these proteins and designing MEV with the highest-scoring epitopes would greatly enhance immunization. Similarly, considering epitopes from antigens expressed in different life-cycle stages would improve the vaccine's protective effect [18]. However, the variation in epitope sequences within the same region may affect the efficacy of the vaccine; thus, the parasites' genotypes and subtypes should be considered from the very beginning of the epitope prediction process.

To systematically bridge the demands of veterinary health and human public safety, the engineered construct in this study is conceptualized as a universal One Health candidate. Recognizing that absolute sterile immunity remains an elusive benchmark for multi-stage apicomplexan parasites, the primary design objectives of this vaccine candidate are focused on disease control and transmission reduction. Specifically, our reverse vaccinology framework combines tachyzoite-, bradyzoite-, and oocyst-specific antigens to theoretically prime an immune profile optimized to restrict acute tachyzoite dissemination, thereby mitigating livestock reproductive losses and congenital transmission, while simultaneously limiting the chronic tissue cyst burden in muscular tissues, which represent the primary vector for foodborne zoonotic transmission to humans and definitive host [22]. Herein, we employed a framework to design a novel MEV encompassing conserved antigenic epitopes from eighteen *T. gondii* proteins capable of eliciting immunity against different *T. gondii* strains (Fig. 1). Candidate proteins were selected based on validated literature reports and records in the ToxoDB database demonstrating expression during key life-cycle transitions. Our rational design integrates reverse vaccinology, molecular docking, molecular dynamics, and immunoinformatic tools to design and predict the efficacy of the candidate vaccine, laying the groundwork for subsequent *in vitro* and *in vivo* evaluation.

Methods

Protein retrieval and selection

Sequences of 18 *T. gondii* proteins were retrieved from ToxoDB (<https://toxodb.org/toxo/app>) [23]. Sequences represented multiple genotypes of *T. gondii*, including ARI, CAST, COUG, GT1, MAS, ME49, RH, RUB, FOU, TgCATBr9, P89, VAND, TgCatPRC2, GAB2-2007-GAL-DOM2 and VEG, which were chosen to adequately represent global genetic diversity, including clonal lineages and atypical regional genotypes **supplementary-1 Table S1.2a**. Thus, ensuring that the epitope prediction would be more universal and less prone to genotype-specific immune response. The selected proteins are: BAG1, GRA6, GRA7, MIC1, MIC3, MIC13, Profilin, ROM4, ROM5, ROP2, ROP18, ROP38, and some members of the SRS protein family (SRS29B, SRS34A, SRS44, SRS49, SRS57, and SRS36D) **supplementary-1 Table S1.2b**. The selection of these target antigens was guided by the presence of published studies proving either alone or in combination could elicit effective immune response and provide protection after *in vivo* challenge, confirmed expression in key life cycles to target different phases of infection, and the presence of the antigens' sequences across all available reference strains to ensure broad genotype coverage. Proteins accession numbers and strains are presented in **supplementary-1 Table S1.3**.

Multiple sequence alignment and conservancy analysis

Conserved regions across different genotypes were identified by aligning protein sequences using the ClustalW algorithm. The resulting sequences were then aligned against the human and murine genomes using BLASTp. The E-value of 0.05 was applied as a cut-off value to exclude high and even weakly matched sequences with host proteins to minimize the risk of cross-reactivity and potential autoimmune response. Therefore, only sequences showing E-values > 0.05 were selected, indicating negligible host homology.

Antigenicity and allergenicity prediction

Antigenicity of the selected proteins was predicted using VaxiJen V2.0 (<https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) and ANTIGENpro (<https://scratch.proteomics.ics.uci.edu/>). Proteins with antigenicity scores ≥ 0.5 were classified as probable antigens. Performing antigenicity screening on proteins already reported as antigens served as an essential computational validation step to quantify and verify their immunogenic scores prior to epitope mapping. In addition, proteins were evaluated for their allergenicity using the AllerTOP V2.0 prediction tool (https://www.ddg-pharmfac.net/allertop_test/). Proteins predicted to be allergenic or non-antigenic were excluded from subsequent analyses.

Epitope prediction

Linear B-cell epitopes were predicted from protein sequences using ABCpred web tool (<http://crdd.osdd.net/raghava/abcpred/>) with 16-mer peptide length and a threshold score of ≥ 0.5 . This webtool prediction model was reported to have an accuracy of 65.93% using cross-validation folds[24] Unlike linear B-cell

epitopes, conformational B-cell epitopes were not considered because the intended vaccine construct is composed of sequentially linked epitopes by specific linkers; therefore, the 3D conformation of the source antigen is not preserved in the final construct [25].

Major Histocompatibility Complex I (MHC-I) binding epitopes were predicted and identified using NetMHCpan 4.2 (<https://services.healthtech.dtu.dk/services/NetMHCpan-4.2/>), which uses artificial neural networks [26]. Protein sequences were screened for 9-mer epitope lengths against all human leukocyte alleles (HLA) supertype representatives (HLA-A01:01,HLA-A02:01,HLA-A03:01,HLA-A24:02,HLA-A26:01,HLA-B07:02,HLA-B08:01,HLA-B27:05,HLA-B39:01,HLA-B40:01,HLA-B58:01,HLA-B15:01) and mouse H-2 alleles. The strong binders showing a percentile rank of ≤ 0.5 and conserved across the *T. gondii* genotypes were selected.

MHC-II epitopes were predicted using the NetMHCIIpan 4.3 server (<https://services.healthtech.dtu.dk/services/NetMHCIIpan-4.3/>) [27] to identify 15-mer core sequences. Proteins were aligned against three human MHC class II isotypes: HLA-DR, HLA-DQ, and HLA-DP, in addition to the mouse H-2 alleles. Epitopes were selected if their rank was ≤ 1 , were conserved across ≥ 10 MHC-II alleles, and covered both species.

The standard epitope lengths were selected to ensure optimal processing and presentation requirement that matches the mammalian immune system. To handle overlapping or duplicate epitopes identified during the initial screening, overlapping segments within the same protein sequence were prioritized first based on their conservation scores, then score and finally their source antigen relevance to prevent structural redundancy in the final constructs.

The final selected T-cell epitopes were evaluated for their ability to induce IFN- γ , IL-10, and IL-4 using IFNepitope (<http://crdd.osdd.net/raghava/ifnepitope/>), IL-10pred (<https://webs.iitd.edu.in/raghava/il10pred/predict3.php>), and IL4pred (<http://crdd.osdd.net/raghava/il4pred/>), respectively. Epitopes were classified as positive inducers based on the default parameters of the used tools. Specifically, IFNepitope (> 0), IL4pred (> 0.2), and IL10pred (> 0.2) classification thresholds are tool-defined (binary positive/negative predictions).

Population coverage analysis

Population coverage of the selected T-cell epitopes was assessed using the IEDB population coverage tool (<https://tools.iedb.org/population/>) against the global population. Individual epitopes were mapped alongside their predicted strong HLA Class I and Class II binders (Supplementary File 1, Tables S1.4) to evaluate coverage distribution[28]. Coverage was calculated using the class I and II combined mode for the global population to predict the average number of epitope-allele combinations per individual (hit coverage), and the projected percentage of individuals in a population covered with at least one binding allele in the construct (population coverage).

Vaccine construct design

Selected epitopes were arranged in three main MEV constructs. The MEV-1 construct was designed with B-cell epitopes positioned at the N-terminal preceding T-cell epitopes. In the MEV-2 construct, B-cell epitopes were placed after T-cell epitopes at the protein C-terminus. The MEV-3 construct only contained T-cell epitopes, excluding B-cell epitopes. This systematic variation in epitope arrangement was implemented to evaluate how the physical positioning of highly flexible linear B-cell regions affects structural folding, thermodynamic processing, and solvent exposure of antibody-binding domains. Also, to investigate whether the B-cell region would collapse and become buried or remain on the exposed surface accessible to B-cell receptors when in proximity to the rigid linker and adjuvant.

Each construct has three variants, one without an adjuvant (original) and one with cholera Toxin B subunit as an adjuvant (A), and one with 50s ribosomal protein L7/L12 as an adjuvant (B), yielding a total of nine constructs. These two adjuvants were selected specifically for toxoplasmosis. Because the parasite is primarily acquired through ingesting tissue cysts or oocysts, generating robust mucosal immunity is a must. Thus, they were employed to address this requirement due to their documented capacity to trigger strong mucosal and systemic cellular responses.

For benchmarking, two previously published and validated multi-epitope vaccine constructs for *T. gondii* were used as positive control (C1, C2), and a construct composed of non-HLA-binding T-cell epitopes was used as negative control (C3).

Epitopes were connected using suitable protein linkers. For all constructs with adjuvants, EAAAK was used as a rigid protein linker between the adjuvant and the MEV. The epitopes were connected by protein linkers; KK linkers for B-cell epitopes, AAY linkers for MHC-I epitopes, and GP GPG Linkers for MHC-II epitopes.

Physicochemical and structural characterization

Constructs' physicochemical properties, including molecular weight, stability index, aliphatic index, theoretical isoelectric point (pI), and Grand Average of Hydropathicity (GRAVY) were computed using ProtParam (<https://web.expasy.org/protparam/>) [29].

Constructs' antigenicity was predicted using VaxiJen v2.0, ANTIGENpro, and Predicted Antigenic Peptides web server (<http://imed.med.ucm.es/Tools/antigenic.pl>).

Constructs' solubility was predicted using protein-sol (<https://protein-sol.manchester.ac.uk/>), Recombinant protein solubility predictor (RPSP) (<https://biotech.ou.edu/>), CamSol (<https://www.vendruscolo.ch.cam.ac.uk/camsolmethod.html>), and SolPro (<https://scratch.proteomics.ics.uci.edu/>).

Possible secondary structures attained by the constructs, including coils, α -helix, and β -sheets were predicted, and their percentages were calculated using PSI-PRED (<https://bioinf.cs.ucl.ac.uk/psipred/>) [30].

Immune response simulation

C-ImmSim (<https://kraken.iac.rm.cnr.it/C-IMMSIM/index.php>) was used to evaluate the immunogenicity of designed constructs. The C-ImmSim is an agent-based model simulator that predicts the immune system response against certain induction [31]. Simulations were performed using the HLA alleles that showed high prevalence during epitope prediction. The simulations were run using random seed 12,345, simulation volume value 10, simulation steps value 100. Each construct was introduced in three doses (steps 1, 48, and 84), with 10,000 units per injection. The output of the simulation included antigen clearance kinetics, B/T cell population, antibody titres (Immunoglobulin M (IgM), Immunoglobulin G (IgG)), and cytokine profile (Interferon gamma (IFN- γ), Interleukin (IL)-2, IL-4, IL-10, and IL-12).

Tertiary structure prediction

The best-performing vaccine constructs by previous analyses were selected for tertiary structure prediction. They were modelled using two different algorithms: trRosetta (<https://yanglab.qd.sdu.edu.cn/trRosetta/>) [32] and AlphaFold2 ColabFold V1.5.5 (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) [33]. The predicted raw 3D structures were refined using GalaxyRefine (<https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>) to correct residue positions and angles. The optimal model was selected based on the stereochemical quality metrics, indicating stable structure formation. The metrics involved Ramachandran plot analysis and ERRAT quality score for both original and refined models using SAVES V6.1 webserver (<https://saves.mbi.ucla.edu/>). Predicted structures were also compared based on their MolProbity, clash score, and poor rotamer generated during refinement. Furthermore, the final 3D structural coordinates were analysed visually to pinpoint the spatial orientation of the N- and C-terminally placed B-cell epitopes to verify their physical exposure on the outer solvent-accessible surface of the protein.

Molecular docking analysis

Antigen-presenting cells' immune receptors, Toll-like receptor 2 (TLR2; PDB ID: 2Z7X) and TLR4 (PDB ID: 4G8A), were used for docking analysis with refined MEV structures. The TLR crystal structures were retrieved from the Protein Data Bank, visualized, and prepared by BIOVIA Discovery Studio 2024 (<https://www.3ds.com/products/biovia/discovery-studio>) through removing water molecules and heteroatoms, and adding polar hydrogens.

Molecular docking analysis was performed to evaluate the structural interaction between the final chimeric MEV constructs and the innate immune receptors (TLR2 and TLR4). Individual epitopes or native source proteins was not assessed due to the lack of the final MEV's conformational constraints within those molecules. Instead, the complete tertiary structures of the MEV were utilized to ensure that structural characteristics, surface charges distributions, and steric constraints introduced by the joining linkers are taken into the account during receptor-MEV pairing.

Protein-protein docking was performed using ClusPro 2.0 (<https://cluspro.org/help.php>) [34], generating multiple cluster models per complex via rigid-body docking followed by energy minimization. The Best models were selected by evaluating docked complexes based on cluster size and relative docking energy scores. For downstream analysis, these energy scores were treated as relative computational indices rather than absolute physical free energies. To characterize the interface properties, the hydrogen bond count was calculated, alongside the predicted binding free energy (ΔG) and predicted dissociation constant (Kd) using PRODIGY server (<https://rascar.science.uu.nl/prodigy/>) [35]. The total buried interface area was quantified using PDBEPIISA (<https://www.ebi.ac.uk/pdbe/pisa/>) [36].

Construct selection procedure

The optimal multi-epitope vaccine prototype selection procedure involved a two-stage sequential gating framework to the initial candidates:

- **Stage 1: Sequence-level checklist filtering:** Candidate within each construct family (MEV-1, MEV-2, MEV-3) were compared based on pass/fail screening matrix to eliminate structural or expression defects. To pass, variants required an instability index < 40.0, positive multi-algorithm solubility parameters (SolPro > 0.90, CamSol > 0.0, and elevated RPSP/Protein-sol scores), antigenicity thresholds (VaxiJen > 0.5, ANTIGENpro > 0.90), and optimal host cell engagement (C-ImmSim). Candidates show promising higher values within each family were selected for stage 2.
- **Stage 2: Post-Docking Weighted scoring matrix:** The selected core candidates were subjected to 3D prediction, refinement and molecular docking against TLR2 and TLR4. The final construct was selected based on multi-criteria weighted scoring formula:

$$\varvec{F} = \varvec{F} \varvec{w}_F + \varvec{C} \varvec{w}_C + \varvec{S} \varvec{w}_S + \varvec{B} \varvec{w}_B + \varvec{K} \varvec{w}_K = \sum_i \varvec{w}_i \varvec{S}_i \varvec{w}_i$$

Weights were distributed across the six docking matrices: ClusPro docking energy (30%), Members of the cluster (20%), predicted ΔG (20%), Number of bonds (10%), total buried interface area (10%), and predicted dissociation constant Kd (10%). The construct yielded the maximum cumulative score across each receptor was chosen as the definitive construct.

B-cell surface topology and relative solvent accessibility (RSA) quantification

The risk of spatial occlusion, internal masking or structural collapse of integrated B-cell epitopes within the tertiary structure of multi-epitope construct were assessed using precise surface visibility analysis using BIOVIA discovery studio 2024.

Prior area calculation, hydrogen atoms were appended and any residual heteroatom or water molecules were removed to clear molecular boundaries. Solvent -accessible surface was computationally generated for the whole construct using a spherical solvent probe radius of 1.4 Å to map water molecule contacts. The c-terminal B-cell epitopes from the leading construct was isolated using interactive sequence display window. The total absolute solvent accessible surface area (SASA) for the c-terminal block was calculated under the selection properties palette.

The absolute SASA footprint ($SASA_{\text{measured}}$) for each individual B-cell epitope were extracted. Because absolute area scales with the varying physical size of unique side-chain chemistries, individual values were normalized into relative solvent accessibility (RSA) percentages to determine the true exposure status. The RSA for each individual epitopes was calculated using this formula:

$$RSA (\%) = \left(\frac{SASA_{\text{measured}}}{\sum_{i=1}^n SASA_{\text{max},i}} \right) \times 100$$

Where $SASA_{\text{max},i}$ represents the standard empirical maximum allowable solvent-accessible surface area for an isolated amino acid residue i embedded within extended Gly-x-Gly tripeptide context as defined by Tien et al (2013).

Residue segments maintaining calculated average RSA $\geq 20\%$ -25% were classified as structurally exposed, whereas average values exceeding a 50% threshold were defined as highly exposed loop architecture.

Molecular dynamics simulation

The protein topologies were prepared in GROMACS 2025.2 [37, 38] using the CHARMM36 family of protein force fields (CHARMM36/CHARMM36m) as distributed for GROMACS [39]. Each protein complex was solvated in a cubic box with the TIP3P water model and neutralized by adding Cl^- counterions. Energy minimization employed the steepest-descent algorithm for up to 50,000 steps, terminating at a maximum force below $200 \text{ kJ mol}^{-1} \text{ nm}^{-1}$, with the Verlet cutoff scheme, PME electrostatics, and 1.2 nm real-space cutoffs for Coulomb and van der Waals interactions.

Equilibration proceeded in two phases. First, NVT (100 ps) at 300 K using the stochastic velocity-rescaling (V-rescale) thermostat. Second, NPT (100 ps) at 300 K and 1 bar using the Parrinello-Rahman barostat. Long-range electrostatics were treated with the particle-mesh Ewald (PME) method, and long-range dispersion corrections to energy and pressure were applied.

Production Molecular dynamics (MD) simulations were performed for 100 ns in the NPT ensemble with a 2 fs timestep at 300 K and 1 bar. Short-range cutoffs were 1.2 nm for both Coulomb and van der Waals interactions (Verlet scheme). Hydrogen bond lengths were constrained using Linear Constraint Solver (LINCS) algorithm. Trajectories were post-processed to remove periodic boundary conditions before analysis.

Following MD, stability and compactness were assessed using standard GROMACS tools: Root Mean Square Deviation (RMSD) of backbone atoms per TLR and vaccine chains (gmxf), residue-wise Root Mean Square Fluctuation (RMSF) of Ca atoms (gmxf), and radius of gyration (Rg) (gmxf). Solvent exposure was quantified as Solvent Accessible Surface Area (SASA) (gmxf). Intermolecular hydrogen bonds between the vaccine construct and TLR chains were computed with chain-specific index files using gmxf hbond. All analyses used the PBC-corrected production trajectories.

Three independent replicates were run for each complex. Each replicate was started from a separate set of initial velocities taken from a Maxwell-Boltzmann distribution at the NVT step, so the three runs are statistically independent. The equilibrium values were averaged over the second half of each trajectory, from 50 to 100 ns.

Stability and interface metrics (RMSD, radius of gyration, SASA, hydrogen bond counts) were extracted from the equilibrated 50–100 ns window of each three-replicate simulation and reported as mean $\pm$ standard deviation. Coefficient of variation ($CV = SD/\text{mean} \times 100\%$) was calculated to assess across-replica reproducibility, with $CV < 30\%$ considered acceptable for protein complex simulations. Differences between MEV-2A-TLR2 and MEV-2A-TLR4 complexes were compared using Welch's t-test (two-tailed, $\alpha = 0.05$), and Cohen's d effect sizes were calculated to report practical significance alongside p-values. All statistical analyses were performed in R version 4.2.0; formal statistical tests were not applied to time-course data to avoid violating the independence assumption.

Residue-level contacts at the receptor-vaccine interface were measured with MDAAnalysis. In each frame, a residue on one chain was counted as in contact with the other chain if any of its heavy atoms lay within $4.5 \text{ \AA}$ of a heavy atom of the partner chain. The contact occupancy of a residue is the fraction of frames in which it was in contact, and these contacts were followed over time and compared across the three replicates.

VD-MM/GBSA

The docked complexes of the MEV construct and immune receptors were further evaluated using the HawkDock V2 server (<http://cadd.zju.edu.cn/hawkdock/>) to estimate binding energetics via the Variable Dielectric Molecular Mechanics/ Generalized Born Surface Area (VD-MM/GBSA) approach [40]. The VD-MM/GBSA approach applies variable, environment-dependent dielectric values based on atomic properties. Thus, prevents overestimation of electrostatic interactions, yielding more accurate predictions of the binding free energy unlike the conventional Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) approach, which only applies a uniform internal dielectric constant. The system calculated the total binding free energy for each receptor-MEV construct complex and performed per-residue energy decomposition to identify the amino acids that contributed most strongly to the interaction interface. This analysis provided a detailed assessment of the stability and strength of receptor engagement following the initial docking stage.

In Silico codon optimization and cloning

The selected MEV construct was reverse-translated to DNA using the Jcat web tool (<https://www.jcat.de/>) to optimize codon usage for *E. coli* BL21 (DE3). The nucleotide sequence was further optimized using VectorBuilder's codon-optimization tool (<https://en.vectorbuilder.com/tool/codon-optimization.html>) to match codon profile of the expression bacterial strain and removal of possible internal restriction cut sites for NdeI and XhoI. Optimized sequence was in silico cloned into the pET-28a(+) vector between NdeI and XhoI restriction sites. The construct included a His-Tag and Sumo-Tag at the N-terminal before the construct sequence, and are in frame with the MEV construct under T7 promoter control as indicated by SnapGene (<https://www.snapgene.com/>).

Data Availability and Traceability

All used and generated datasets, sequences and structural models during this study are organized across three supplementary materials and open access repository to ensure data transparency, data integrity and cross-platform reproducibility:

Supplementary Material 1 (Inputs)

contains all experimental setup, including computational server configurations, target protein accession numbers and raw population coverage allele frequency inputs. Each dataset is presented in single spreadsheet and listed on the main page of the workbook.

Supplementary Material 2 (curated outputs)

presents the final predicted epitopes (MHC-I, MHC-II, and linear B-cell epitopes), cytokine profile analysis output, construct physicochemical and immune simulation analysis, Tertiary structure and molecular docking analysis. Each dataset is presented in single spreadsheet and listed on the main page of the workbook.

Supplementary material 3 (Figures)

a document containing high-resolution figures for key analytical milestones, include secondary structure analysis, immune simulation analysis, molecular dynamic analysis.

Zenodo repository: permanently archives of all raw computational logs, sequence and unrefined structural pdb files and uncurated epitope prediction result data. The repository is organized into folders numbered chronologically according to the workflow sequence. All files are publicly accessible via the persistent link: <https://zenodo.org/records/19582147> and under DOI: <https://doi.org/10.5281/zenodo.19582147>

Results

Prediction and screening of epitopes

Eighteen *T. gondii* protein sequences were retrieved from ToxoDB and screened to identify potential immunogenic targets. Of these, thirteen proteins were predicted to be antigenic (antigenicity score > 0.5) and non-allergenic. These proteins were used for epitope prediction, while the remaining five proteins, including SRS29B, SRS34A, SRS44, MIC1, and ROP8, were excluded due to low antigenicity or allergenic potential (Table 1).

Table 1

T. gondii Proteins evaluation. The table shows the antigenicity and allergenicity prediction results of the tested proteins.

Protein	Primarily Existing Life Cycle Stage(s)	Antigenicity Score (Average)	Interpretation	Allergenicity
BAG1	Bradyzoite	0.571	ANTIGEN	NON-ALLERGEN
GRA6	Tachyzoite, Bradyzoite, Oocyst	0.887	ANTIGEN	NON-ALLERGEN
GRA7	Tachyzoite, Bradyzoite, Oocyst	0.538	ANTIGEN	NON-ALLERGEN
MIC1	Tachyzoite, Sporozoite	0.480	NON-ANTIGEN	NON-ALLERGEN
MIC3	Bradyzoite (Upregulated)	0.857	ANTIGEN	NON-ALLERGEN
profilin	All Phases (Constitutive)	0.545	ANTIGEN	NON-ALLERGEN
ROM4	Tachyzoite, Bradyzoite	0.666	ANTIGEN	NON-ALLERGEN
ROM5	Tachyzoite	0.688	ANTIGEN	NON-ALLERGEN
ROP18	Tachyzoite	0.449	NON-ANTIGEN	ALLERGEN
ROP2	Tachyzoite, Bradyzoite, Sporozoite	0.597	ANTIGEN	NON-ALLERGEN
ROP38	Tachyzoite, Bradyzoite	0.585	ANTIGEN	NON-ALLERGEN
SRS29B (SAG-1)	Tachyzoite (Exclusive)	0.787	ANTIGEN	ALLERGEN
SRS34A (SAG-2)	Tachyzoite	0.576	ANTIGEN	ALLERGEN
SRS44 (CST)	Tachyzoite	0.524	ANTIGEN	ALLERGEN
SRS57 (SAG-3)	Bradyzoite (Exclusive)	0.644	ANTIGEN	NON-ALLERGEN
SRS36D (SAG5B)	Bradyzoite-specific	0.778	ANTIGEN	NON-ALLERGEN
SRS49D (SAG2C)	Tachyzoite, Bradyzoite	0.658	ANTIGEN	NON-ALLERGEN
MIC13	All Phases (Constitutive)	0.556	ANTIGEN	NON-ALLERGEN

The selected antigenic proteins were used to predict linear B-cell epitopes, MHC class I and II epitopes. A threshold score of ≥ 0.5 and percentile rank cut-offs of ≤ 0.5 and ≤ 1.0 were applied for selecting strong binding epitopes for B-cell, MHC-I, and MHC-II, respectively. Additionally, epitopes that showed

conservancy in all tested genotypes and interacted with both murine and human alleles were selected. Predicted B-cell epitopes were derived from the members of the SRS family and MIC3, representing surface-associated antigens. Predicted MHC-I epitopes were derived from GRA6, ROM4, ROM5, and ROP38, representing secretory and invasion proteins that are typically exposed to the host immune system. Predicted MHC-II epitopes were derived equally from all proteins.

Applying these filtrations **supplementary-1 Table S1.5**, a total of 43 unique epitopes were selected: 25 MHC-I, 10 MHC-II, and 8 B-cell epitopes (**Supplementary-2 Table S2.1-2.3**). MIC3 contributed to the highest epitope density (7 epitopes) and was represented across three cell types (3 B-cell epitopes, 3 MHC-I epitopes, 1 MHC-II epitopes).

Cytokine induction potential

The capability of cytokine-induction by the selected T-cell epitopes was evaluated through their ability to stimulate T-helper (Th)1/Th2 immune response. Of the 35 selected epitopes, 11 (31.4%) induced IFN- γ , 2 (5.7%) induced IL-10, and 14 (40%) induced IL-4.

Notably, the GRA7 epitope (KRSFKDLLKK) was the highlight of the prediction test as it scored positive for all three cytokines. Additionally, MIC13-EESTHHITW, MIC3-HPRYHAHTV, ROP2-ALLEGFLRY, PROFILIN-QQDWKDAAVFRALDG, and SRS57-RSKITYFGTLTQKAP epitopes were predicted to activate both IFN- γ and IL-4 (**Supplementary-2 Table S2.4**).

Population coverage analysis

The selected epitopes achieved a theoretical global population coverage of 100% for both classes combined (MHC-I and MHC-II), with a predicted average hit of 21.2 across worldwide populations (**Supplementary-2 Table S2.5**). While the high predicted score suggests hypothetical broad potential allele recognition across populations, the actual protective immunity remains subject to proper delivery, antigen recognition, and processing.

Multi-epitope vaccine construct design and evaluation

The nine designed constructs candidates, namely MEV-1, MEV-1A, MEV-1B, MEV-2, MEV-2A, MEV-2B, MEV-3, MEV-3A, and MEV-3B, which vary in epitope positioning and adjuvant inclusion (Fig. 2), were subjected to evaluation regarding antigenicity, physicochemical properties, and immunogenicity. The construct lengths ranged from 514 to 793 residues, reaching molecular weights ranging from 53.8 to 69.5 kDa.

All constructs were predicted to exhibit strong antigenicity with a vaxiJen score of 0.57 ± 0.03 and ANTIGENpro score averaging 0.94. Moreover, all constructs that included either adjuvant A or B showed higher solubility compared to the same construct without adjuvant, except for the MEV-3, which exhibited higher solubility without adjuvant when predicted via RPSP. In contrast, when the solubility prediction was performed with SolPro, all constructs exhibited high solubility, except for the MEV-3 construct without an adjuvant (**Supplementary-2 Table S2.6**).

Constructs' secondary structure prediction and analysis via PSI-PRED revealed a balance between α -helices and β -strands in MEV-1 and MEV-2 constructs. In MEV-3 constructs, α -helices were more represented (average of 42.1%). Generally, the average compositions of all constructs ranged from 33.8 to 43.6% for α -helices, 11.2 to 15.1% for β -strands, and 43.4 to 52.9% for coils (**Supplementary-2 Table S2.6 and Supplementary-3 Fig. S1**).

The immune response potential of each construct was modelled and assessed through the C-ImmSim server to rank the candidates. Each construct was administered in a simulated three-dose regimen of 10,000 units/dose with simulation steps 1, 48, and 84. Simulation outcomes indicated a noticeable predicted humoral immune response characterized by a dose-dependent increase in simulated immunoglobulin titres after administration. Modelled IgG1 and IgG2 levels were significantly increased, especially after secondary and tertiary doses, suggesting computational immunological memory formation. Cytokine profiling revealed a strong Th1-skewed immune response bias, suggested by robust simulated induction of IFN- γ and IL-2, along with moderate IL-10 and IL-12 levels. This observation was prominent in MEV-2 and MEV-3 original constructs and their MEV-A variants. These simulated cytokine and antibody trajectories are model-based predictions under optimized parameters; thus, they do not necessarily reflect the real physiological immune responses measured in *in vivo* settings.

Based on the comparative checklist strategy detailed in the methodology, the MEV-A variants (MEV-1A, MEV-2A, and MEV-3A) were selected for further analysis as they were only candidates within each family to successfully pass most of the foundational stability, predicted solubility and antigenicity, and immune simulation characteristics (**Supplementary-2 Table S2.7 and Supplementary-3 Fig. S2**).

Tertiary structure prediction

Tertiary structures of the selected constructs were generated using two de novo prediction algorithms, trRosetta and AlphaFold. AlphaFold predicted structure models exhibited a more globular and compact form with per-residue confidence scores (pLDDT) greater than 80% in the N-terminal region of each protein, while the remaining regions exhibited lower than 40%, indicating high flexibility in the multi-epitope regions. On the other hand, trRosetta predicted models showed more relaxed structures (Fig. 3).

All produced model structures were refined using the GalaxyRefine server, followed by structure and stereochemical quality assessment using the SAVES web server. The trRosetta produced models initially demonstrated high ERRAT quality scores, which were slightly improved after refinement, with an average of 96.9. On the other hand, the AlphaFold initially demonstrated lower ERRAT quality scores (averaged 54.08), which were significantly increased to an average of 91.9 after refinement. The same pattern was also observed with the Ramachandran plot analysis, with the refined MEV-1A models exhibiting the highest percentages in the favored area.

Additional parameters were also used in the evaluation of produced tertiary structures, including clash score, which resulted in an average of 2.7, and a poor rotamer analysis, which resulted in an average of 0.2 across all refined models, confirming stereochemical reliability (**supplementary-2 Table S2.8**).

Molecular docking analysis

The refined tertiary structure models from both trRossette and AlphaFold were docked against the immune cell receptors, TLR2 and TLR4, using the ClusPro 2.0 web server to assess and analyze the interaction between each MEV construct and both TLRs. Docked complexes were ranked by cluster size, docking energy, intermolecular bonds, predicted binding free energy, predicted dissociation constant, and interface area.

For TLR4, all tested MEV constructs formed stable computational complexes, with relative docking energies ranging from – 1481 to -1666, interface areas ranging from 2116 to 3051 Å², and predicted binding energies (ΔG) ranging from – 20.7 to -26.5 kcal/mol. Among these, the AlphaFold-derived MEV-2A-TLR4 complex showed a balanced profile with a highly favorable binding energy of -1542.2, an interface area of 3051 Å², and a high number of formed bonds (82). It also exhibited a predicted binding energy (ΔG) of -26.5 kcal/mol and a predicted dissociation constant (Kd) of 2.1E-19 M as calculated by PRODIGY, suggesting a very strong interaction (Fig. 4, Table 2). MEV-1A and MEV-3A also exhibited strong binding tendencies to the TLR4, but with either smaller interface, fewer formed bonds or lower predicted ΔG compared to the MEV-2A.

Table 2
Docking analysis of the constructs with TLR4 and TLR2

Docking complex	Model	Energy	Members	No. of bonds	Within interface	Predicted Delta G (Kcal/mol)	Predicted K _d (M)	Cumulative Weighted score
MEV-1A_TLR4_trROSSETA	5	-1597.3	15	79	2863.8	-24.5	5E-18	781.37
MEV-1A_TLR4_Alpha	No interaction occurred							
MEV-2A_TLR4_Alpha	8	-1542.2	17	82	3051	-26.5	2.1E-19	784.66
MEV-2A_TLR4_trROSSETA	4	-1666.4	21	53	2246.5	-16.8	1.4E-12	737.43
MEV-3A_TLR4_Alpha	2	-1481.3	19	68	2124.9	-21.1	1.40E-15	671.7
MEV-3A_TLR4_trROSSETA	5	-1542.8	20	86	2116.6	-20.7	2.40E-15	691.24
MEV-1A_TLR2_Alpha	0	-1334.3	59	42	1492.4	-19.4	2.1E-14	569.41
MEV-1A_TLR2_trROSSETA	8	-1416.3	19	46	1571.3	-15.8	7E-12	593.58
MEV-2A_TLR2_Alpha	4	-1552.1	27	66	2466	-19	3.8E-14	728.03
MEV-2A_TLR2_trROSSETA	9	-1359.6	19	40	1474.5	-14.5	6.1E-11	566.03
MEV-3A_TLR2_Alpha	0	-1357.8	67	41	1370.4	-14.6	4.9E-11	564.8
MEV3A_TLR2_trROSSETA	0	-1128.6	58	25	1268.5	-12.5	1.5E-09	482.03

For TLR2, a similar pattern was observed, all tested MEV constructs formed stable complexes, with docking energies ranging from – 1128 to -1552, interface areas ranging from 1268 to 2466 Å², and predicted binding energies (ΔG) ranging from – 12.5 to -19.4 kcal/mol. The AlphaFold-derived MEV-2A-TLR2 complex exhibited the most favorable profile with a docking score of -1552.1, an interface area of 2466 Å², and a high number of formed bonds (66) (Fig. 5, Table 2).

The molecular docking results indicated that all the tested MEV constructs have structural capacity with both TLR2 and TLR4. Yet, the MEV-2A, specifically those derived from AlphaFold, exhibited favorable and consistent computational binding parameters with both receptor types. Furthermore, the PDBePISA analysis confirmed that these MEV-2A-TLR4 and MEV-2A-TLR2 complexes generated the largest interface areas, with scores of 3051 Å² and 2466 Å², respectively, indicating the formation of extensive networks of non-covalent interactions. Therefore, these complexes were selected for further evaluation through molecular dynamics simulations.

Following the integration of these dual-receptor docking metrics into post-docking weighted scoring matrix, MEV-2A emerged as the definitive vaccine construct prototype. By achieving the highest cumulative weighted score (Table 2), MEV-2A was uniquely selected over MEV-1A and MEV-3A for downstream structural validation via molecular dynamics simulations.

Surface exposure and solvent accessibility analysis of B-cell epitopes

Comprehensive surface topology and solvent accessibility profile was mapped to ensure that protein assembly did not introduce internal crowding of B-cell epitopes (Fig. 6). The SASA calculated for 142 amino acid c-terminal domain housing the eight linear B-cell epitopes and their (KK) linkers was 12,267.9 Å². when normalized against empirical maximum tripeptide reference values, the entire humoral displayed collective favorable RSA of 57.35%, indicating that the target antibody-binding region forms an expensive, solvent accessible loop network on the outer shell of the protein.

Residue-by-residue extraction demonstrated that all eight individual linear B-cell epitopes are highly exposed with independent SASA ranging from 1,179.96 Å² to 1595.84 Å² (**supplementary-2 Table S2.9**). This uniformly high exposure trend mathematically confirms that engineered rigid-to-flexible linker configuration effectively maintained segment spacing enable the entire B-cell domain to project directly into surrounding aqueous environment. This spatial orientation ensures that the critical epitopes maintained unhindered and fully available for host b-cell receptor (BCR) recognition.

Molecular Dynamics Simulations of MEV-2A–TLR2 and MEV-2A–TLR4 Complexes

Receptor and MEV-2A stability

Three independent 100-ns simulations were run for MEV-2A-TLR2 and MEV-2A-TLR4 complexes to assess the assembly stability and to check for potential receptors structural disruptions. In both systems, the receptor backbones reached a stable plateau within the first few nanoseconds maintaining low RMSD values between 0.2 to 0.4 nm for the rest of the run (Fig. 7A,C). Averaged over the equilibrated 50 to 100 ns window, the receptor backbone RMSD ranged from 0.28 to 0.43 nm (mean ± SD: 0.36 ± 0.07 nm) for the TLR2 receptor backbone and from 0.20 to 0.32 nm for TLR4 receptor backbone (Table 3). Both receptors demonstrated excellent across-replica reproducibility with coefficient of variation (CV) values of 21.5% for TLR2 and 21.9% for TLR4.

Table 3
Equilibrium structural and interfacial descriptors (mean ± SD over 50–100 ns) for the MEV-2A-TLR2 and MEV-2A-TLR4 complexes across three independent replicates.

Metric	Replicate 1	Replicate 2	Replicate 3	Mean ± SD
MEV-2A-TLR2 complex				
RMSD receptor (nm)	0.28 ± 0.04	0.38 ± 0.07	0.43 ± 0.07	0.36 ± 0.07
RMSD MEV-2A (nm)	3.15 ± 0.04	3.12 ± 0.08	2.81 ± 0.03	3.03 ± 0.17
Rg receptor (nm)	2.98 ± 0.03	3.01 ± 0.03	2.90 ± 0.03	2.96 ± 0.05
Rg MEV-2A (nm)	3.44 ± 0.03	3.44 ± 0.10	3.51 ± 0.08	3.46 ± 0.07
SASA receptor (nm ²)	268.23 ± 2.87	269.53 ± 2.78	268.58 ± 2.50	268.78 ± 2.63
SASA MEV-2A (nm ²)	529.65 ± 12.19	571.98 ± 17.00	580.33 ± 17.81	560.65 ± 24.12
Interface H-bonds	22.00 ± 3.01	68.35 ± 14.05	53.88 ± 28.32	48.1 ± 23.7
MEV-2A-TLR4 complex				
RMSD receptor (nm)	0.29 ± 0.04	0.20 ± 0.02	0.32 ± 0.02	0.27 ± 0.06
RMSD MEV-2A (nm)	3.08 ± 0.06	3.72 ± 0.10	2.80 ± 0.02	3.24 ± 0.48
Rg receptor (nm)	3.14 ± 0.04	3.21 ± 0.02	3.15 ± 0.01	3.17 ± 0.03
Rg MEV-2A (nm)	3.77 ± 0.05	3.46 ± 0.10	3.35 ± 0.03	3.53 ± 0.19
SASA receptor (nm ²)	268.68 ± 2.82	269.24 ± 2.31	272.50 ± 2.45	270.14 ± 2.64
SASA MEV-2A (nm ²)	604.96 ± 14.41	579.77 ± 20.61	573.41 ± 12.22	586.05 ± 18.06
Interface H-bonds	18.11 ± 3.75	10.81 ± 5.76	14.58 ± 4.67	14.5 ± 3.7

MEV-2A, a flexible multi-epitope construct, was more mobile at the start of each run, rising over the first 25 to 30 ns before levelling off (Fig. 7B,D). Its equilibrated backbone RMSD ranged from 2.81 to 3.15 nm in MEV-2A-TLR2 complexes (3.03 ± 0.17 nm) and from 2.80 to 3.72 nm in MEV-2A-TLR4 complexes (3.24 ± 0.48 nm) (Table 3), so despite its flexibility the antigen reached a stable conformation in every run.

Flexibility, compaction, and surface exposure of MEV-2A

The per-residue Ca RMSF profiles were consistent with RMSD behavior; The TLR2 and TLR4 residues fluctuated little, while MEV-2A was more flexible, mostly in its linkers, termini and epitope regions (**supplementary-3 Fig S3**). The radius of gyration (Rg) showed the same contrast between the steady receptor and the more adaptable antigen. This contrast was mirrored by Rg and SASA profiles. TLR2 Rg averaged 2.96 nm (range 2.90–3.01 nm, CV = 1.9%) with SASA values 265–275 nm², while TLR4 Rg averaged 3.18 nm (range 3.15–3.21 nm, CV = 1.1%) with SASA 280–295 nm². In sharp contrast, MEV-2A became more compact over time. The antigen's Rg falling from an extended starting value of roughly 4.5 to 5.0 nm to a plateau near 3.4 to 3.5 nm with TLR2 and 3.4 to 3.8 nm with TLR4 (supplementary-3 Fig S4; Table 3). Concurrently, its SASA values dropped from 870–890 nm² to stable equilibrated baselines (supplementary-3 Fig S5; Table 3), indicating that the vaccine construct folds into a highly compact, globular shape upon target receptor engagement.

Interface hydrogen bonds and contact mapping

The interface was held together by hydrogen bonds that formed early in each run and never fell to zero, indicating that MEV-2A stayed bound throughout (supplementary-3 Fig S6). The MEV-2A-TLR2 interface was significantly denser and steadily gained hydrogen bonds overtime, with equilibrated counts of 22.0 ± 3.0, 68.4 ± 14.1 and 53.9 ± 28.3 (48.1 ± 23.7 bond per equilibrated frame) across the three runs. Conversely, the MEV-2A-TLR4 interface proved thinner and steadier at 18.1 ± 3.8, 10.8 ± 5.8 and 14.6 ± 4.7 (14.5 ± 3.7 bond per frame) representing a 3.3-fold difference (Table 3). Despite the substantial

quantitative difference, this difference was not statistically significant (Welch's t-test: $t(4) = 2.42$, $p = 0.131$), likely due to high within-group variance in TLR2 runs (range 22.0-6.1 bond; CV = 49.4%) reflecting distinct binding modes sampled across replicates. However, the large effect size (Cohen's $d = 1.98$) indicates practical significance: TLR2 engagement favors a denser hydrogen-bonded interface compared to TLR4. Crucially, both complexes maintained continuous hydrogen bond formation throughout all simulations (minimum 10 bonds per frame in both systems), confirming that both receptor-MEV-2A interactions remain stable despite marked differences in contact density. Time-course analysis revealed distinct binding

kinetics: the TLR2 interface showed progressive hydrogen bond accumulation during the first 50 ns (slope approximately 0.8 bonds/ns), then stabilized at higher numbers, while TLR4 maintained relatively steady bonding from early in the run (supplementary-3 Fig S6), suggesting that TLR2 engagement involves a gradual, cumulative sampling of conformational states, whereas TLR4 rapidly samples a narrower configurational space.

Heavy-atom (4.5 Å cutoff) revealed that MEV-2A selectively presents distinct binding faces to interact with each receptor (supplementary-3 Fig S7). The TLR2 interface clustered within a, a hydrophobic and aromatic pocket on the receptor (residues 319 to 359, such as Ile319, Phe322, Tyr323, Phe325, Leu328, Tyr332, Thr335, Phe349, Val351, Pro352 and Leu355, together with Arg321, Arg167, His147, Glu190 and Glu215). On the vaccine side the residues facing TLR2 were even more consistent: the stretch from Val259 to Val271 maintained continuous contact across all replicates, driven by core interactions at Tyr264, Val265, Pro266 and Ser267. In contrast, TLR4 used a highly charged binding interface (residues 595 to 624, Lys595, Arg598, Glu603, Glu605, Arg606, Glu608, Asp614, Lys615 and Gln616). To engage this pocket, MEV-2A used a completely separate binding surface centered on residues 236 to 245 (Ser236, Val237, Ala238, Tyr240, His241, Pro242, Arg243, Tyr244 and His245) and 190 to 204 (Ala190, Tyr191, Arg193, Ser194, Lys201 and Tyr204). The time course of these contacts (supplementary-3 Fig S8) showed that these primary residues stayed engaged in all three runs for almost the whole 100 ns (the solid dark bands), so each complex is held by a defined, reproducible set of contacts, with MEV-2A presenting a different face of the antigen to each receptor.

In Silico Codon Optimization and Cloning

Codon usage of the MEV-2A construct was optimized for *E. coli* BL21 (DE3) as the selected host for expression via the VectorBuilder codon optimization tool.

The optimized gene yielded a 2,364 bp coding sequence with a codon adaptation index (CAI) of 0.93, and a GC content of 52.4%, consistent with efficient expression in *E. coli*.

The optimized sequence was verified to ensure it did not contain unwanted regulatory motifs or internal occurrences of the target restriction sites. The reverse translation of the optimized sequences confirmed full preservation of the original MEV-2A amino acid sequences.

This optimized sequence was in silico cloned into the pET-28a(+) vector between the NdeI and XhoI restriction sites with a preceding N-terminal Sumo and His-tags under the control of the T7 promoter using SnapGene (Fig. 8). The constructed cassette was then subjected to a simulated translation, resulting in the production of a recombinant protein with a predicted molecular weight of about 96.145 kDa, indicating that the MEV-2A is compatible with the standard *E. coli* expression systems.

Discussion

Our study applied an integrated immunoinformatics and structural vaccinology approach to design a multi-epitope vaccine targeting several *T. gondii* genotypes. The parasite's genotypic diversity and complex life cycle, including stage-specific antigen expression, intracellular persistence, and broad population distribution, necessitate a vaccine that integrates multiple genotypes and life cycle stages provoking both innate and adaptive immunity [41]. Previous work in *T. gondii* vaccine research has investigated antigens from key functional or structural protein families, including SRS, ROP, ROMs, MICs, and GRAs [42].

SRS proteins form the major coat of tachyzoites, mediating host-cell attachments and early immune recognition. SRS-based vaccines (DNA, protein, or Virus-Like particle (VLP)), particularly those utilizing SAG-1, SAG-2, SAG-3, SRS36D, and SRS49D, enhanced humoral and cellular immunity, with SRS36D and SRS49D reducing tissue cyst counts by over 70% [43, 44]. ROPs act as immune modulators and virulence factors injected during invasion; immunization with the ROP2 family provokes a Th1-skewed response characterized by high levels of IFN- γ , IL-2, and TNF- α [45]. ROMs, particularly ROM4 and ROM5, contribute to parasite motility and cell invasion, inducing humoral and cellular immunity accompanied by high levels of IFN- γ , IL-2, IL-12, and IL-23 [46, 47]. MIC adhesins (MIC3, MIC4, and MIC13) identified as a strong source of B- and T-cell epitopes that elicit mixed Th1/Th2 cytokines profiles, prolonged survival and lowered cyst burden in mice [48, 49]. GRAs modulate host immunity and shape the intracellular parasitophorous vacuole [50]. Lastly, Stage-specific antigens like the bradyzoite marker BAG1 offer modest protection against the chronic phases, which improves when combined with other antigens [51].

Successful experimental combinations that amplify specific IgG titers, IFN- γ release, and survival rates include ROP5–ROP7–SAG1 [52], ROP2-ROP5-ROP9-ROP16-ROP17-ROP18 [45], GRA5-SAG1-ROP2 and SAG1-SAG2-SAG3 [53]. Other studies used mixtures or fusions of GRAs or MICs, consistently showing enhanced immune response [54]. For example, MIC3-GRA7 [55] or GRA6-MIC13 [56, 57], all increased specific IgG titers, IFN- γ release, and enhance survival rates. Immunoinformatic studies have mirrored these combinations; examples include multi-stage vaccine targeting apicomplexan-specific kinases (CDPK1, 2, 3, and 5) early-stage coverage via SRS-derived proteins (TgVax452) [58], and tachyzoite-specific pathways combining ROP2, MIC3, and GRA7 or GRA6 and MIC3 [59]. Some studies also demonstrated that some proteins from parasite's internal cytoplasmic machinery and intracellular compartments exhibit immunogenic epitopes that can be successfully processed and presented [60, 61].

Because most previous studies focused on either the tachyzoites' surface or secretory antigens. Thus, our work extends these designs by combining a broader, multi-stage antigen set from different functional classes covering the key processes occurring during different parasitic life-cycle stages and circulating genotypes. Proteins associated with host-homology risk, such as HSPs and CDPKs [62], were excluded to minimise the risk of autoimmunity. To

completely eliminate the risk of vaccine-induced autoimmunity, an E-value threshold of 0.05 was implemented during host-homology filtering against human and mouse proteomes. In this context, a permissive cutoff acts as an exceptionally stringent safety barrier, maximizing alignment sensitivity to flag and exclude any parasite segment carrying even a minor risk of molecular mimicry or cross-reactivity.

Epitope mapping confirmed that MIC3 is a rich source of MHC-I, MHC-II, and B-cell epitopes, consistent with literature [63]. Epitopes from GRA6, ROM4, ROM5, and ROP38 contributed to provoking cellular immunity, whereas SRS proteins and MIC3 contributed to provoking humoral immunity. Cytokine profiling verified protective, Th1-biased baseline, identifying epitopes capable of inducing IFN- γ and IL-2, which are crucial for macrophage activation and tachyzoite clearance [50]. However, because an excessive Th1 response can act as an immunological double-edged sword, increasing the risk of localized tissue necrosis and inflammatory pathology, we explicitly included regulatory epitopes capable of inducing IL-10 and the Th2 driver IL-4 to ensure balanced cell-mediated control [64]. These selected epitopes demonstrated 100% theoretical worldwide population coverage, verifying broad applicability within computational human allele mapping models while retaining coverage for pre-clinical murine evaluations [16]. Exploring various epitope arrangements increased confidence in the selection of the final protein construct to proceed to downstream applications such as protein production and *in vivo* studies. This approach was adopted by previous immunoinformatic studies, such as developing a multi-epitope vaccine for malaria [65].

Adjuvant incorporation was a key determinant in optimizing vaccine performance. Adjuvant variants, particularly the “A” series (MEV-1A, MEV-2A, MEV-3A), successfully resolved the near-total baseline sequence insolubility of their parent variants, driving a huge increase in RPSP solubility probability (from 0.1% to 98.6%) and yielding positive CamSol scores. Immunoinformatic analysis showed that adjuvant variants especially MEV-2A, induced a strong humoral and cellular responses marked by increased IgG1/IgG2, IFN- γ and IL-2 levels, likely mediated by improved immune receptors engagement [66]. While cholera toxin subunit B effectively boosts immunogenicity in preclinical studies [67], and serves as an excellent benchmark for driving mucosal responses in veterinary models, human translation may require clinical-grade alternatives like MPLA or emulsion-based systems to mitigate systemic toxicity and high reactogenicity [68].

MEV recombinant proteins could benefit from modern adjuvants such as nanoparticles-based encapsulation (e.g., poly (lactic-co-glycolic acid) (PLGA) [69], chitosan[70] and lipid nanoparticles [71]) to shields the protein from enzymatic degradation and facilitate targeting delivery to immune cells, and enable less aggressive rout of administration such as oral and intranasal routs which matches the natural entry of *T. gondii*. Advanced nanocarriers like polyethyleneimine (PEI) trigger endosomal escape via the proton sponge effect, while calcium phosphate nanoparticles exhibit acid-dependent dissolution to maximize cytoplasmic bioavailability [72]. Alternative platforms beyond the protein-based systems include genetic delivery, which has become highly viable to evaluate the multi-epitope-based vaccines, such as viral vectors, DNA plasmids, and mRNA. These genetic platforms force endogenous antigen production and processing, which triggers stronger cytotoxic T-cell responses than protein platforms, besides ease of design and manufacture [73].

Structural modelling and receptor docking indicated favourable relative docking scores, large interface areas, and a persistent hydrogen-bond network. Crucially, these docking energies represent relative computational scoring values rather than true physical binding free energies and were utilized strictly for relative benchmarking between the constructs. To optimize selection, a post-docking weighted structural scoring matrix was implemented, evaluating candidates across six core metrics. MEV-2A emerged as the definitive vaccine prototype, outperformed its counterparts in dual-receptor binding thermodynamics with both TLR2 and TLR4. Our structural vaccinology pipeline prioritized docking and molecular dynamics of the fully integrated MEV over isolated epitopes. This approach is methodologically vital because multi-epitope constructs represent *de novo* designed proteins with novel conformational properties such as tertiary folding, spatial epitope exposure, and thermodynamic binding behaviour towards pattern recognition receptors (PRRs). Those properties are entirely different from their single original epitope/protein sequences. By treating the engineered construct as a single, macromolecular ligand, we were able to more accurately simulate how its structural flexibility promotes the engagement with TLR2 and TLR4 residues.

The subsequent 100-ns MD simulations confirmed that the MEV-2A-TLR2 and the MEV-2A-TLR4 complexes are dynamically stable. While the rigid receptor frameworks maintained exceptionally low RMSD plateaus (< 0.4 nm), the flexible chimeric antigen underwent rapid thermodynamic compaction, confirmed by the synchronized collapse of both its Rg and SASA. This structural flexibility allowed MEV-2A to engage with each receptor via distinct, non-overlapping spatial faces: binding TLR2 through extensive adaptable hydrophobic and aromatic networks, and TLR4 through charge-dense electrostatic patches, indicating a robust capacity of triggering simultaneous innate signalling cascades.

Finally, codon optimization and *in silico* cloning of MEV-2A into pET-28a (+) for expression in *E. coli* BL21 showed predicted compatibility with high CAI and favourable GC content, supporting experimental feasibility and facilitating protein production. While expanding the construct length was necessary to increase epitope density and maximize efficacy, it could consequently complicate protein production and purification. The total length of the construct could impose constraints, such as reduced translation efficiency, increased metabolic burden on host cells, and a high likelihood of accumulation in inclusion bodies due to solubility issues. To overcome the solubility issues and avoid inclusion bodies formation, SUMO Tag was incorporated into the design to increase solubility and enhance protein folding. Additionally, optimization protocols for protein production should be employed. Alternative methods, such as switching to eukaryotic systems, could benefit protein folding and high-yield production.

Despite these optimized parameters, this work shares a fundamental limitation with most immunoinformatic-driven vaccine designs, where the prediction tools for antigenicity, allergenicity, epitope prediction, population coverage, structure prediction, docking, immune profiling, and MD can have inherent biases and do not guarantee comparable effective protection when applied *in vivo*. In particular, favourable TLR2/TLR4 docking and stable MD trajectories demonstrate structural compatibility and a physically plausible binding mode, but they do not establish receptor agonism, downstream signalling activation, or productive antigen presentation; confirming these steps requires dedicated functional immunological assays. Furthermore, these simulated metrics represent idealized mathematical models of host-pathogen interactions and must not be overinterpreted as direct evidence of living biological safety or clinical efficacy. Therefore, an experimental vaccination challenge for both acute and chronic toxoplasmosis is indispensable to validate and prove the MEV-2A efficacy. Additionally, the formulation and delivery strategies for the constructed vaccine should be investigated to enhance receptor engagement, cellular, and humoral immune responses.

Conclusion

In this study, we employed a reverse vaccinology and immunoinformatics framework to design a robust vaccine for *T. gondii*. Our final designed and selected construct, MEV-2A, has deliberately combined 43 conserved T- and B-cell epitopes selected from Thirteen antigens representing multiple functional protein classes and life-cycle stages that are selected not only for antigenicity and conservancy across different genotypes but also for their documented robust *in vivo* protective effect. MEV-2A construct showed high predicted antigenicity, hypothetical broad global population coverage, favoured physicochemical properties, and a Th1- skewed immune profile. Furthermore, structural modelling, docking, and molecular dynamics simulations confirmed stability and well-maintained engagement with innate immune receptors TLR2 and TLR4.

Rather than relying on the high and often elusive benchmark of sterile infection prevention, this universal One Health construct is structurally engineered for disease control and transmission reduction. Its target application framework is optimized to theoretically trigger an immune response capable of suppressing acute tachyzoite dissemination while simultaneously reducing the chronic tissue cyst burden that serves as the major source of foodborne zoonotic transmission to humans. Thus, addressing a critical gap in toxoplasmosis control. These findings nominate the computationally predicted MEV-2A as a next-generation vaccine candidate against toxoplasmosis, awaiting *in vitro* and *in vivo* validation to confirm its efficacy and safety. Moreover, the presented framework provides a robust rational foundation and generalized template for multi-epitope and multistage vaccine design targeting parasites comprising different genotypes.

Abbreviations

BAG1

Bradyzoite antigen 1

CAI

Codon adaptation index

CDPKs

Calcium-dependent protein kinases

GRA

Dense granule proteins

GRAVY

Grand Average of Hydropathicity

HLA

Human leukocyte alleles

HSPs

Heat-shock proteins

IFN- γ

Interferon gamma

IgG

Immunoglobulin G

IgM

Immunoglobulin M

IL

Interleukin

Kd

Dissociation constant

MD

Molecular dynamics

MEV

Multiepitope vaccine

MIC

Microneme proteins

MHC

Major histocompatibility complex

pI

Isoelectric point

Rg

Radius of gyration

RMSD

Root mean square deviation

RMSF

Root mean square fluctuation

ROM

Rhomboid proteins

ROP

Rhoptry proteins

RPSP

Recombinant protein solubility predictor

SASA

Solvent accessible surface area

SRS

SAG-related sequence

T. gondii

Toxoplasma gondii

Th

T-helper

TLR

Toll-like receptor

VLP

Virus-like particle

ΔG

Binding free energy

Declarations

- **Ethics approval and consent to participate**

Not applicable

- **Consent for publication**

Not applicable

- **Data Availability**

All data generated or analyzed in the current study are available through this manuscript, supplementary materials, and through the Zenodo repository: <https://zenodo.org/records/19582147>

- **Competing interests**

The authors declare no competing interests.

- **Funding**

This work was supported by Science, Technology & Innovation Funding

Authority (STDF), Egypt, under the "One Health Grant" (project number 49100).

- **Authors' contributions**

MRZ and DME designed the study, analyzed the data, selected the promising epitopes, and wrote the manuscript. DME collected protein data, sequence alignment, conservancy analysis and performed antigenicity, allergenicity, and immunogenicity. MRZ performed epitope prediction for MHCI, MHCII, and B-cell, constituted the constructs, and executed their validations.

- **Acknowledgements**

We thank Dr. Mohamed Yasser Foda, Faculty of Science, Mansoura University, for his help in MD simulation and its data analysis.

We also like to thank Mr. Mohanad M. Faried, Nanomaterials Research and Synthesis Unit, Animal Health Research Institute (AHRI), Agricultural Research Center (ARC), for his help in protein data collection and curation of previous publications.

References

1. Farhab M, Aziz MW, Shaukat A, Cao M-X, Hou Z, Huang S-Y, et al. Review of Toxoplasmosis: What We Still Need to Do. Vet Sci. 2025;12:772. <https://doi.org/10.3390/vetsci12080772>.
2. Silva LA, Andrade RO, Carneiro ACA V., Vitor RWA. Overlapping *Toxoplasma gondii* Genotypes Circulating in Domestic Animals and Humans in Southeastern Brazil. PLoS One. 2014;9:e90237. <https://doi.org/10.1371/journal.pone.0090237>.

3. Razali K, Issad NA, Mebkhouf F, Boudjellaba S, Khelef D. High seroprevalence and associated risk factors of *Toxoplasma gondii* infection in aborted ewes from Tebessa Province, Northeastern Algeria: A One Health perspective. *Vet World*. 2025;18:3367–77. <https://doi.org/10.14202/vetworld.2025.3367-3377>.
4. Ayub I, Sani SS, Zaman A, Shahid MR, Bashir W, Ijaz M, et al. Prevalence of *Toxoplasma gondii* in Human and Animals and its Economic Impacts: A Review. *The Pakistan Veterinary Journal*. 2024. <https://doi.org/10.29261/pakvetj/2024.221>.
5. Howe DK, Sibley LD. *Toxoplasma gondii* Comprises Three Clonal Lineages: Correlation of Parasite Genotype with Human Disease. *Journal of Infectious Diseases*. 1995;172:1561–6. <https://doi.org/10.1093/infdis/172.6.1561>.
6. Su C, Khan A, Zhou P, Majumdar D, Ajzenberg D, Dardé M-L, et al. Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proceedings of the National Academy of Sciences*. 2012;109:5844–9. <https://doi.org/10.1073/pnas.1203190109>.
7. Meeusen ENT, Walker J, Peters A, Pastoret P-P, Jungersen G. Current Status of Veterinary Vaccines. *Clin Microbiol Rev*. 2007;20:489–510. <https://doi.org/10.1128/CMR.00005-07>.
8. Attias M, Teixeira DE, Benchimol M, Vommaro RC, Crepaldi PH, De Souza W. The life-cycle of *Toxoplasma gondii* reviewed using animations. *Parasit Vectors*. 2020;13:588. <https://doi.org/10.1186/s13071-020-04445-z>.
9. Dunay IR, Gajurel K, Dhakal R, Liesenfeld O, Montoya JG. Treatment of Toxoplasmosis: Historical Perspective, Animal Models, and Current Clinical Practice. *Clin Microbiol Rev*. 2018;31. <https://doi.org/10.1128/CMR.00057-17>.
10. Innes EA, Hamilton C, Garcia JL, Chrysafidis A, Smith D. A one health approach to vaccines against *Toxoplasma gondii*. *Food Waterborne Parasitol*. 2019;15:e00053. <https://doi.org/10.1016/j.fawpar.2019.e00053>.
11. Buxton D, Innes EA. A commercial vaccine for ovine toxoplasmosis. *Parasitology*. 1995;110:S11–6. <https://doi.org/10.1017/S003118200000144X>.
12. Dubey JP. Toxoplasmosis in sheep—The last 20 years. *Vet Parasitol*. 2009;163:1–14. <https://doi.org/10.1016/j.vetpar.2009.02.026>.
13. Chu K-B, Quan F-S. Advances in *Toxoplasma gondii* Vaccines: Current Strategies and Challenges for Vaccine Development. *Vaccines (Basel)*. 2021;9:413. <https://doi.org/10.3390/vaccines9050413>.
14. Zaher MR, El-Husseiny MH, Hagag NM, El-Amir AM, El Zowlaty ME, Tammam RH. A novel immunoinformatic approach for design and evaluation of heptavalent multiepitope foot-and-mouth disease virus vaccine. *BMC Vet Res*. 2025;21:152. <https://doi.org/10.1186/s12917-025-04509-1>.
15. Alsaiari AA, Hakami MA, Alotaibi BS, Alkhalil SS, Hazazi A, Alkhorayef N, et al. Rational design of multi-epitope-based vaccine by exploring all dengue virus serotypes proteome: an immunoinformatic approach. *Immunol Res*. 2024;72:242–59. <https://doi.org/10.1007/s12026-023-09429-6>.
16. Margaroni M, Agallou M, Tsanaksidou E, Kammona O, Kiparissides C, Karagouni E. Immunoinformatics Approach to Design a Multi-Epitope Nanovaccine against *Leishmania* Parasite: Elicitation of Cellular Immune Responses. *Vaccines (Basel)*. 2023;11:304. <https://doi.org/10.3390/vaccines11020304>.
17. Dey J, Mahapatra SR, Singh PK, Prabhushwamimath SC, Misra N, Suar M. Designing of multi-epitope peptide vaccine against *Acinetobacter baumannii* through combined immunoinformatics and protein interaction-based approaches. *Immunol Res*. 2023;71:639–62. <https://doi.org/10.1007/s12026-023-09374-4>.
18. Hajissa K, Zakaria R, Suppian R, Mohamed Z. Epitope-based vaccine as a universal vaccination strategy against *Toxoplasma gondii* infection: A mini-review. *J Adv Vet Anim Res*. 2019;6:174. <https://doi.org/10.5455/javar.2019.f329>.
19. Zhang Y, Li D, Lu S, Zheng B. Toxoplasmosis vaccines: what we have and where to go? *NPJ Vaccines*. 2022;7:131. <https://doi.org/10.1038/s41541-022-00563-0>.
20. Foroutan M, Zaki L, Tavakoli S, Soltani S, Taghipour A, Ghaffarifar F. Rhomboid antigens are promising targets in the vaccine development against *Toxoplasma gondii*. *EXCLI J*. 2019;18:259–72. <https://doi.org/10.17179/excli2018-1993>.
21. Tian X, Yang Z, Wan G, Xie T, Wang M, Sun H, et al. Vaccination with recombinant *Toxoplasma gondii* bradyzoite-formation deficient 1 (rTgBFD1) antigen provides partial protective immunity against chronic *T. gondii* infection. *Front Vet Sci*. 2022;9:957479. <https://doi.org/10.3389/fvets.2022.957479>.
22. Mévélec M-N, Lakhfir Z, Dimier-Poisson I. Key Limitations and New Insights Into the *Toxoplasma gondii* Parasite Stage Switching for Future Vaccine Development in Human, Livestock, and Cats. *Front Cell Infect Microbiol*. 2020;10. <https://doi.org/10.3389/fcimb.2020.607198>.
23. Harb OS, Kissinger JC, Roos DS. ToxoDB: the functional genomic resource for *Toxoplasma* and related organisms. In: *Toxoplasma gondii*. Elsevier; 2020. p. 1021–41. <https://doi.org/10.1016/B978-0-12-815041-2.00023-2>.
24. Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins: Structure, Function, and Bioinformatics*. 2006;65:40–8. <https://doi.org/10.1002/prot.21078>.
25. Chatterjee R, Sahoo P, Mahapatra SR, Dey J, Ghosh M, Kushwaha GS, et al. Development of a Conserved Chimeric Vaccine for Induction of Strong Immune Response against *Staphylococcus aureus* Using Immunoinformatics Approaches. *Vaccines (Basel)*. 2021;9:1038. <https://doi.org/10.3390/vaccines9091038>.
26. Nilsson JB, Greenbaum J, Peters B, Nielsen M. NetMHCpan-4.2: improved prediction of CD8+ epitopes by use of transfer learning and structural features. *Front Immunol*. 2025;16. <https://doi.org/10.3389/fimmu.2025.1616113>.
27. Nilsson JB, Kaabinejadian S, Yari H, Kester MGD, van Balen P, Hildebrand WH, et al. Accurate prediction of HLA class II antigen presentation across all loci using tailored data acquisition and refined machine learning. *Sci Adv*. 2023;9. <https://doi.org/10.1126/sciadv.adj6367>.
28. Bui H-H, 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. <https://doi.org/10.1186/1471-2105-7-153>.
29. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press; 2005. p. 571–607. <https://doi.org/10.1385/1-59259-890-0:571>.

30. Buchan DWA, Moffat L, Lau A, Kandathil SM, Jones DT. Deep learning for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* 2024;52:W287–93. <https://doi.org/10.1093/nar/gkae328>.
31. Stolfi P, Castiglione F, Mastrostefano E, Di Biase I, Di Biase S, Palmieri G, et al. In-silico evaluation of adenoviral COVID-19 vaccination protocols: Assessment of immunological memory up to 6 months after the third dose. *Front Immunol.* 2022;13. <https://doi.org/10.3389/fimmu.2022.998262>.
32. Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, et al. The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc.* 2021;16:5634–51. <https://doi.org/10.1038/s41596-021-00628-9>.
33. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022;19:679–82. <https://doi.org/10.1038/s41592-022-01488-1>.
34. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein–protein docking. *Nat Protoc.* 2017;12:255–78. <https://doi.org/10.1038/nprot.2016.169>.
35. Xue LC, Rodrigues JP, Kastitis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics.* 2016;32:3676–8. <https://doi.org/10.1093/bioinformatics/btw514>.
36. Krissinel E. Crystal contacts as nature's docking solutions. *J Comput Chem.* 2010;31:133–43. <https://doi.org/10.1002/jcc.21303>.
37. Kutzner C, Páll S, Fechner M, Esztermann A, de Groot BL, Grubmüller H. Best bang for your buck: GPU nodes for GROMACS biomolecular simulations. *J Comput Chem.* 2015;36:1990–2008. <https://doi.org/10.1002/jcc.24030>.
38. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1–2:19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.
39. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, et al. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J Chem Theory Comput.* 2012;8:3257–73. <https://doi.org/10.1021/ct300400x>.
40. Zhang X, Jiang L, Weng G, Shen C, Zhang Q, Liu M, et al. HawkDock version 2: an updated web server to predict and analyze the structures of protein–protein complexes. *Nucleic Acids Res.* 2025;53:W306–15. <https://doi.org/10.1093/nar/gkaf379>.
41. Erazo Flores BJ, Knoll LJ. *Toxoplasma gondii* at the Host Interface: Immune Modulation and Translational Strategies for Infection Control. *Vaccines (Basel).* 2025;13:819. <https://doi.org/10.3390/vaccines13080819>.
42. Rezaei F, Sarvi S, Sharif M, Hejazi SH, Pagheh A sattu, Aghayan SA, et al. A systematic review of *Toxoplasma gondii* antigens to find the best vaccine candidates for immunization. *Microb Pathog.* 2019;126:172–84. <https://doi.org/10.1016/j.micpath.2018.11.003>.
43. Zhang M, Zhao L, Song J, Li Y, Zhao Q, He S, et al. DNA vaccine encoding the *Toxoplasma gondii* bradyzoite-specific surface antigens SAG2CDX protect BALB/c mice against type II parasite infection. *Vaccine.* 2013;31:4536–40. <https://doi.org/10.1016/j.vaccine.2013.07.065>.
44. Lu G, Zhou J, Zhou A, Han Y, Guo J, Song P, et al. SAG5B and SAG5C combined vaccine protects mice against *Toxoplasma gondii* infection. *Parasitol Int.* 2017;66:596–602. <https://doi.org/10.1016/j.parint.2017.06.002>.
45. Du R, He J, Meng J, Zhang D, Li D, Wang H, et al. Vaccination with a DNA vaccine cocktail encoding TgROP2, TgROP5, TgROP9, TgROP16, TgROP17, and TgROP18 confers limited protection against *Toxoplasma gondii* in BALB/c mice. *Parasitol Res.* 2024;123:420. <https://doi.org/10.1007/s00436-024-08435-3>.
46. Zhang N-Z, Xu Y, Wang M, Petersen E, Chen J, Huang S-Y, et al. Protective efficacy of two novel DNA vaccines expressing *Toxoplasma gondii* rhomboid 4 and rhomboid 5 proteins against acute and chronic toxoplasmosis in mice. *Expert Rev Vaccines.* 2015;14:1289–97. <https://doi.org/10.1586/14760584.2015.1061938>.
47. Han Y, Zhou A, Lu G, Zhao G, Wang L, Guo J, et al. Protection via a ROM4 DNA vaccine and peptide against *Toxoplasma gondii* in BALB/c mice. *BMC Infect Dis.* 2017;17:59. <https://doi.org/10.1186/s12879-016-2104-z>.
48. Ghaffari AD, Dalimi A, Ghaffarifar F, Pirestani M, Majidiani H. Immunoinformatic analysis of immunogenic B- and T-cell epitopes of MIC4 protein to designing a vaccine candidate against *Toxoplasma gondii* through an in-silico approach. *Clin Exp Vaccine Res.* 2021;10:59. <https://doi.org/10.7774/cevr.2021.10.1.59>.
49. Onile OS, Ojo GJ, Oyeyemi BF, Agbowuro GO, Fadahunsi AI. Development of multiepitope subunit protein vaccines against *Toxoplasma gondii* using an immunoinformatics approach. *NAR Genom Bioinform.* 2020;2. <https://doi.org/10.1093/nargab/lqaa048>.
50. Yang C-S, Yuk J-M, Lee Y-H, Jo E-K. *Toxoplasma gondii* GRA7-Induced TRAF6 Activation Contributes to Host Protective Immunity. *Infect Immun.* 2016;84:339–50. <https://doi.org/10.1128/IAI.00734-15>.
51. Alijani M, Saffar B, Yosefi Darani H, Mahzounieh M, Fasihi- Ramandi M, Shakshi-Niaei M, et al. Immunological evaluation of a novel multi-antigenic DNA vaccine encoding SAG1, SAG3, MIC4, GRA5, GRA7, AMA1 and BAG1 against *Toxoplasma gondii* in BALB/c mice. *Exp Parasitol.* 2023;244:108409. <https://doi.org/10.1016/j.exppara.2022.108409>.
52. Moghadamzad Z, Dalimi A, Pirestani M, Ghafarifar F. Designing a multi-epitope vaccine using *Toxoplasma* ROP5, ROP7, and SAG1 epitopes and immunogenicity evaluation against acute and chronic toxoplasmosis in BALB/c mice. *Microb Pathog.* 2025;204:107567. <https://doi.org/10.1016/j.micpath.2025.107567>.
53. Naserifar R, Ghaffarifar F, Dalimi A, Sharifi Z, Solhjoo K, Hosseini Khosroshahi K. Evaluation of Immunogenicity of Cocktail DNA Vaccine Containing Plasmids Encoding Complete GRA5, SAG1, and ROP2 Antigens of *Toxoplasma gondii* in BALB/C Mice. *Iran J Parasitol.* 2015;10:590–8.
54. Chyb M, Ferra BT, Kawka M, Skwarecka M, Dziadek B, Gatkowska J. Immunogenicity and protective efficacy of recombinant chimeric antigens based on surface proteins of *Toxoplasma gondii*. *Front Immunol.* 2024;15. <https://doi.org/10.3389/fimmu.2024.1480349>.

55. Dodangeh S, Fasihi-Ramandi M, Daryani A, Valadan R, Asgarian-Omran H, Hosseini Z, et al. Protective efficacy by a novel multi-epitope vaccine, including MIC3, ROP8, and SAG1, against acute *Toxoplasma gondii* infection in BALB/c mice. *Microb Pathog*. 2021;153:104764. <https://doi.org/10.1016/j.micpath.2021.104764>.
56. Wu Q, Zhang Z, Chu H, Xia B, Li W, Ding J, et al. A novel combined quadrivalent self-amplifying mRNA-LNP vaccine provokes protective immunity against acute and chronic toxoplasmosis in mice. *Infect Dis Poverty*. 2025;14:55. <https://doi.org/10.1186/s40249-025-01332-6>.
57. Sun X-M, Zou J, AA ES, Yan W-C, Liu X-Y, Suo X, et al. DNA vaccination with a gene encoding *Toxoplasma gondii* GRA6 induces partial protection against toxoplasmosis in BALB/c mice. *Parasit Vectors*. 2011;4:213. <https://doi.org/10.1186/1756-3305-4-213>.
58. Majidiani H, Pourseif MM, Kordi B, Sadeghi M-R, Najafi A. TgVax452, an epitope-based candidate vaccine targeting *Toxoplasma gondii* tachyzoite-specific SAG1-related sequence (SRS) proteins: immunoinformatics, structural simulations and experimental evidence-based approaches. *BMC Infect Dis*. 2024;24:886. <https://doi.org/10.1186/s12879-024-09807-x>.
59. Hammed-Akanmu M, Mim M, Osman AY, Sheikh AM, Behmard E, Rabaan AA, et al. Designing a Multi-Epitope Vaccine against *Toxoplasma gondii*: An Immunoinformatics Approach. *Vaccines (Basel)*. 2022;10:1389. <https://doi.org/10.3390/vaccines10091389>.
60. Vargas-Montes M, Valencia-Jaramillo MC, Valencia-Hernández JD, Gómez-Marín JE, Arenas AF, Cardona N. In silico identification and ex vivo evaluation of *Toxoplasma gondii* peptides restricted to HLA-A*02, HLA-A*24 and HLA-B*35 alleles in human PBMC from a Colombian population. *Med Microbiol Immunol*. 2025;214:5. <https://doi.org/10.1007/s00430-024-00815-x>.
61. McMurtrey C, Trolle T, Sansom T, Remesh SG, Kaever T, Bardet W, et al. *Toxoplasma gondii* peptide ligands open the gate of the HLA class I binding groove. *Elife*. 2016;5. <https://doi.org/10.7554/eLife.12556>.
62. Ghaffari AD, Rahimi F. Immunoinformatics studies and design of a novel multi-epitope peptide vaccine against *Toxoplasma gondii* based on calcium-dependent protein kinases antigens through an in-silico analysis. *Clin Exp Vaccine Res*. 2024;13:146. <https://doi.org/10.7774/cevr.2024.13.2.146>.
63. Xiang W, Qiong Z, Li-peng L, Kui T, Jian-wu G, Heng-ping S. The location of invasion-related protein MIC3 of *Toxoplasma gondii* and protective effect of its DNA vaccine in mice. *Vet Parasitol*. 2009;166:1–7. <https://doi.org/10.1016/j.vetpar.2009.08.014>.
64. Fereig RM, Omar MA, Alsayeqh AF. Exploiting the Macrophage Production of IL-12 in Improvement of Vaccine Development against *Toxoplasma gondii* and *Neospora caninum* Infections. *Vaccines (Basel)*. 2022;10:2082. <https://doi.org/10.3390/vaccines10122082>.
65. Pritam M, Singh G, Swaroop S, Singh AK, Pandey B, Singh SP. A cutting-edge immunoinformatics approach for design of multi-epitope oral vaccine against dreadful human malaria. *Int J Biol Macromol*. 2020;158:159–79. <https://doi.org/10.1016/j.ijbiomac.2020.04.191>.
66. Viallard JF, Pellegrin JL, Ranchin V, Schaeffer T, Dehais J, Longy-Boursier M, et al. Th1 (IL-2, interferon-gamma (IFN- γ)) and Th2 (IL-10, IL-4) cytokine production by peripheral blood mononuclear cells (PBMC) from patients with systemic lupus erythematosus (SLE). *Clin Exp Immunol*. 2001;115:189–95. <https://doi.org/10.1046/j.1365-2249.1999.00766.x>.
67. Pan X, Ke H, Niu X, Li S, Lv J, Pan L. Protection Against *Helicobacter pylori* Infection in BALB/c Mouse Model by Oral Administration of Multivalent Epitope-Based Vaccine of Cholera Toxin B Subunit-HUUC. *Front Immunol*. 2018;9:1003. <https://doi.org/10.3389/fimmu.2018.01003>.
68. Bourgeois AL, Scott DA, Baqar S, Tribble DR, Svennerholm A-M, Walker RI. The case for evaluation of double mutant heat-labile toxin as an adjuvant to improve oral cholera vaccine immunogenicity. *Front Immunol*. 2026;17. <https://doi.org/10.3389/fimmu.2026.1805998>.
69. Zhou P, Yu Y, Qi W, Wang X, Yu Y, Wang J, et al. PLGA nanoparticles as an efficient carrier in *Toxoplasma* GAP45: a more effective vaccine against acute toxoplasmosis than traditional ones. *Front Immunol*. 2025;16. <https://doi.org/10.3389/fimmu.2025.1600399>.
70. Hosseini Z, Daryani A, Fasihi-Ramandi M, Asgarian-Omran H, Nayeri T, Dodangeh S, et al. Design and immunological evaluation of a multi-epitope vaccine candidate against *Toxoplasma gondii* incorporating MIC13, GRA1, and SAG1 antigens in BALB/c mice. *Food Waterborne Parasitol*. 2025;40:e00269. <https://doi.org/10.1016/j.fawpar.2025.e00269>.
71. Viegas C, Seck F, Fonte P. An insight on lipid nanoparticles for therapeutic proteins delivery. *J Drug Deliv Sci Technol*. 2022;77:103839. <https://doi.org/10.1016/j.jddst.2022.103839>.
72. Wang X, Li H, Chen C, Liang Z. Understanding of endo/lysosomal escape of nanomaterials in biomedical application. *Smart Molecules*. 2025;3. <https://doi.org/10.1002/smo.20240017>.
73. Zhou ZP, Chen ZR, Bandara RA, Duan R, Cao H, Liu J, et al. DNA-Based Vaccines: Advances, Applications, and Future Prospects. *Genes Dis*. 2026;10:2025. <https://doi.org/10.1016/j.gendis.2025.102025>.

Figures

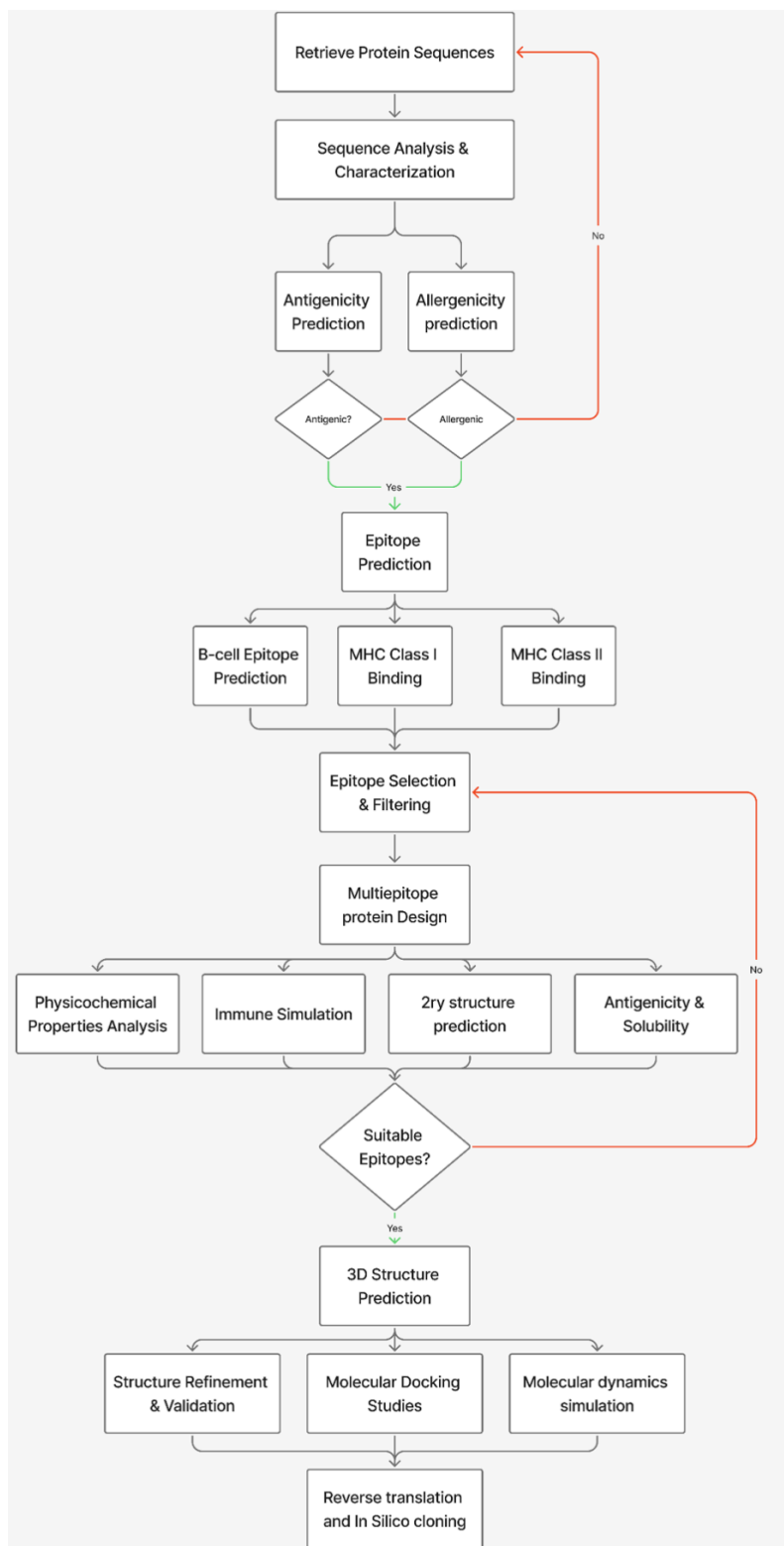


Figure 1

Vaccine computational design workflow.

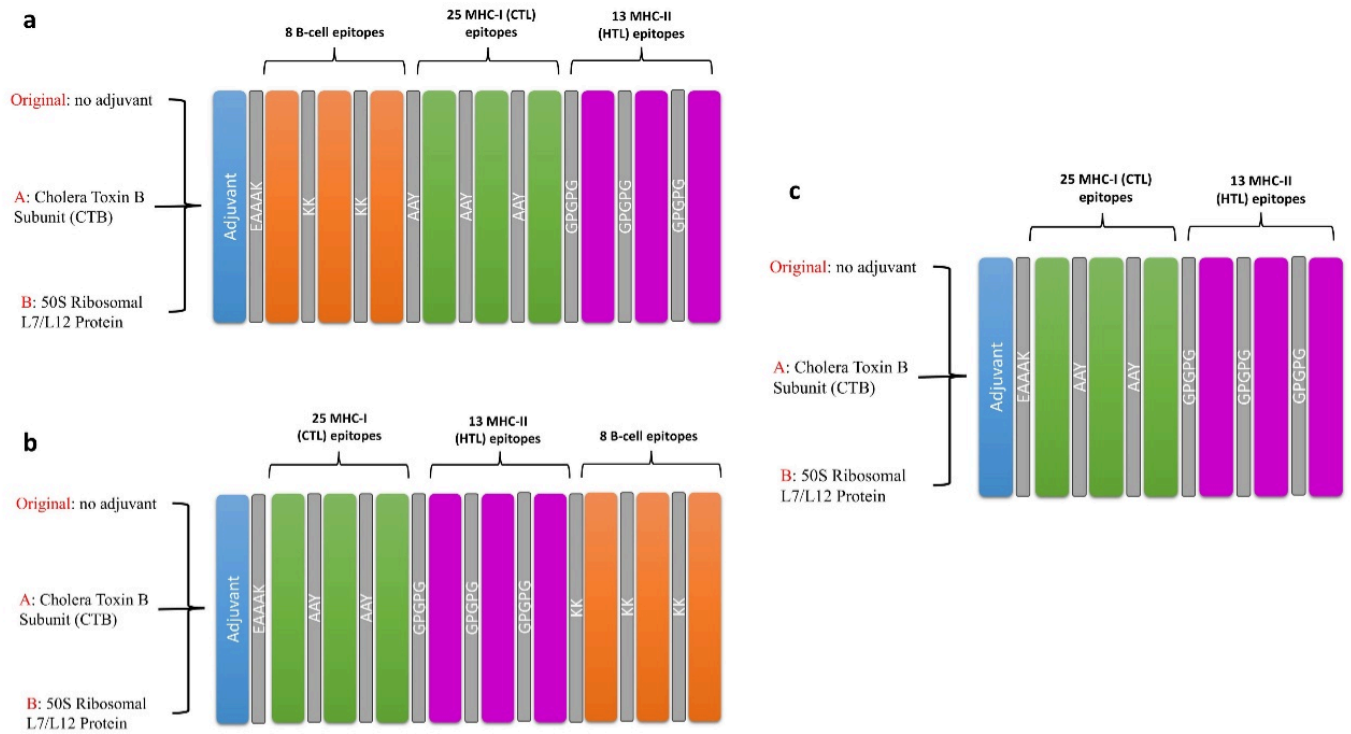


Figure 2

Vaccine construct design for A) MEV-1 variants, B) MEV-2 variants, and C) MEV-3 variants. The figure shows the organization of B, MHC-I, and MHC-II epitopes, and the adjuvants of A and B construct variants.

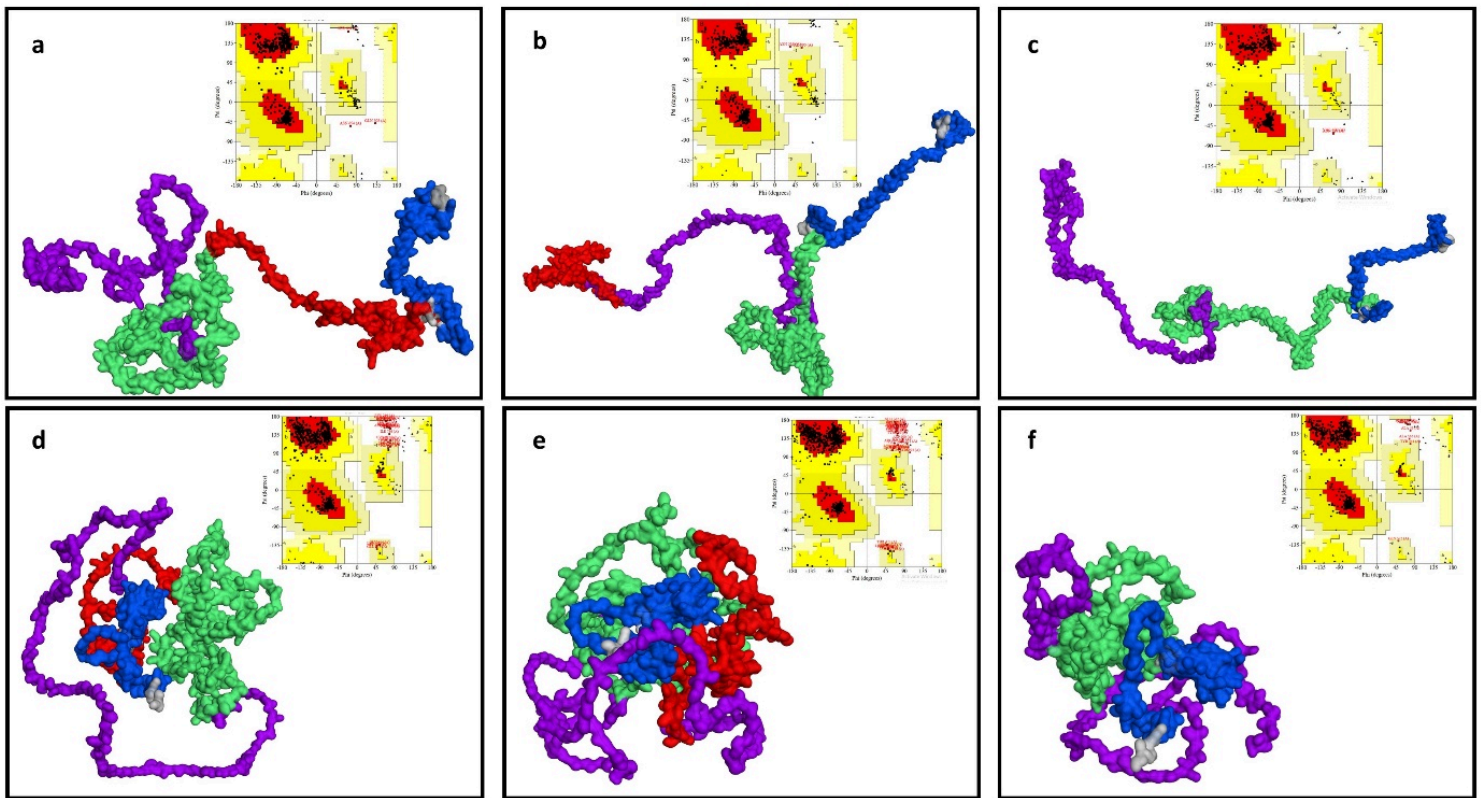


Figure 3

Tertiary structure prediction of predicted vaccines with Ramachandran plots of each construct. A) MEV-1A using trRossetta, B) MEV-2A using trRossetta, C) MEV-3A using trRossetta, D) MEV-1A using ColabFold, E) MEV-2A using ColabFold, and F) MEV-3A using ColabFold. The color scheme of the structures represents epitopes where Blue represents the adjuvant, Green represents MHC-I epitopes, Magenta represents MHC-II epitopes, and Red represents B-cell epitopes.

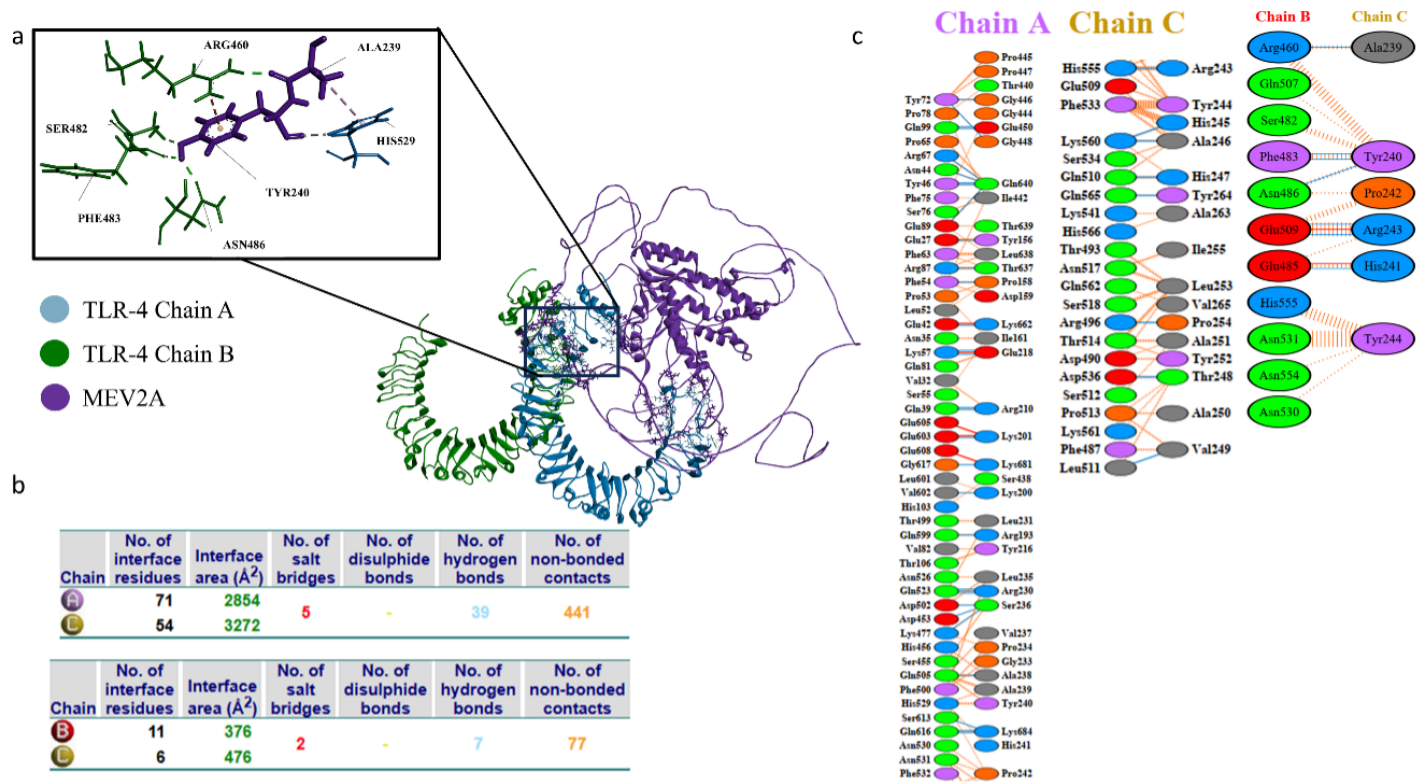


Figure 4

Molecular docking analysis for MEV-2A docked models with TLR4. A) Visualization of docked MEV-2A- TLR4 complex using Discovery Studio. B) the number of interacting residues between TLR4 chain A and MEV-2A, and TLR4-chain B and MEV-2A. C) The visualization of interacting residues using PDBsum between TLR4 and MEV-2A

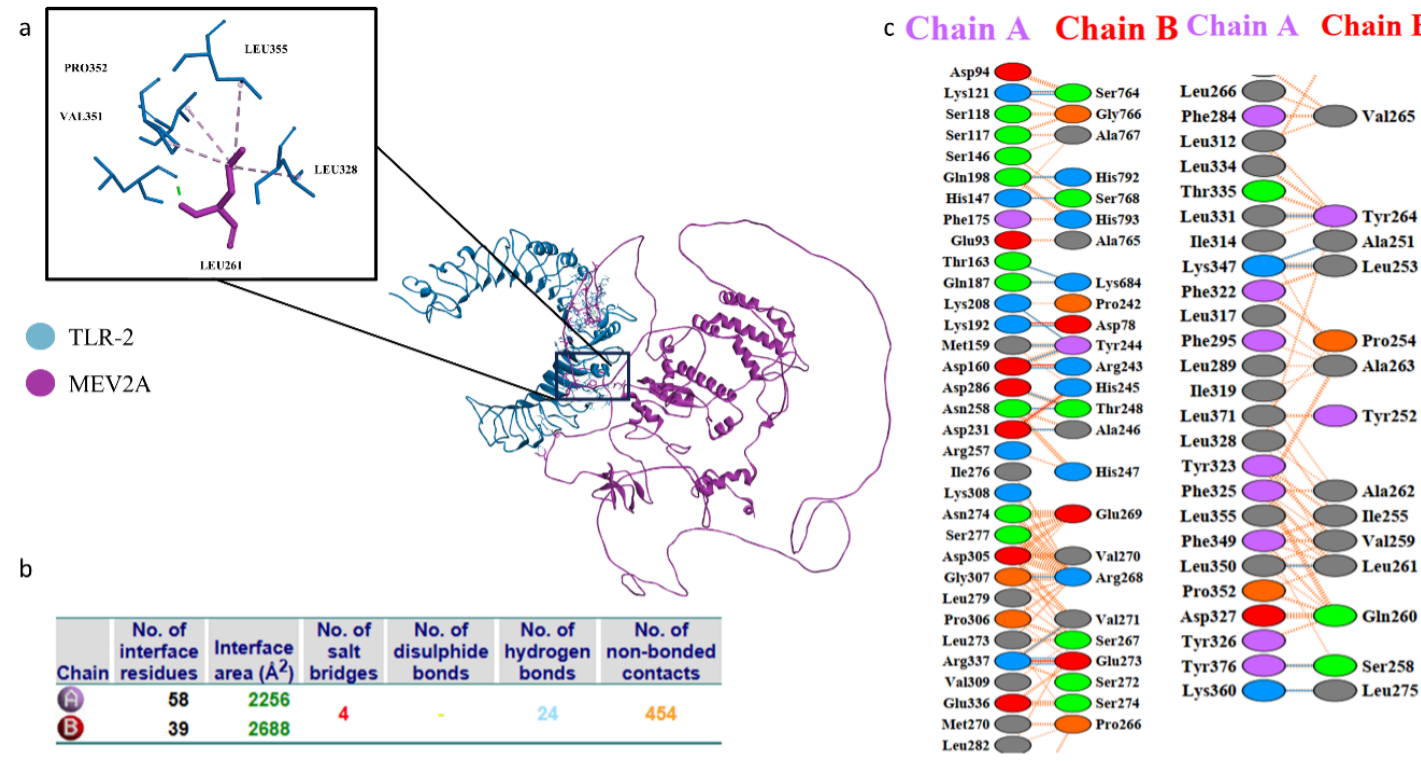


Figure 5

Page 21/24

Molecular docking analysis for MEV-2A docked models with TLR2. A) Visualization of docked MEV-2A- TLR2 complex using Discovery Studio. B) the number of interacting residues between TLR2 and MEV-2A. C) The visualization of interacting residues using PDBsum between TLR2 and MEV-2A.

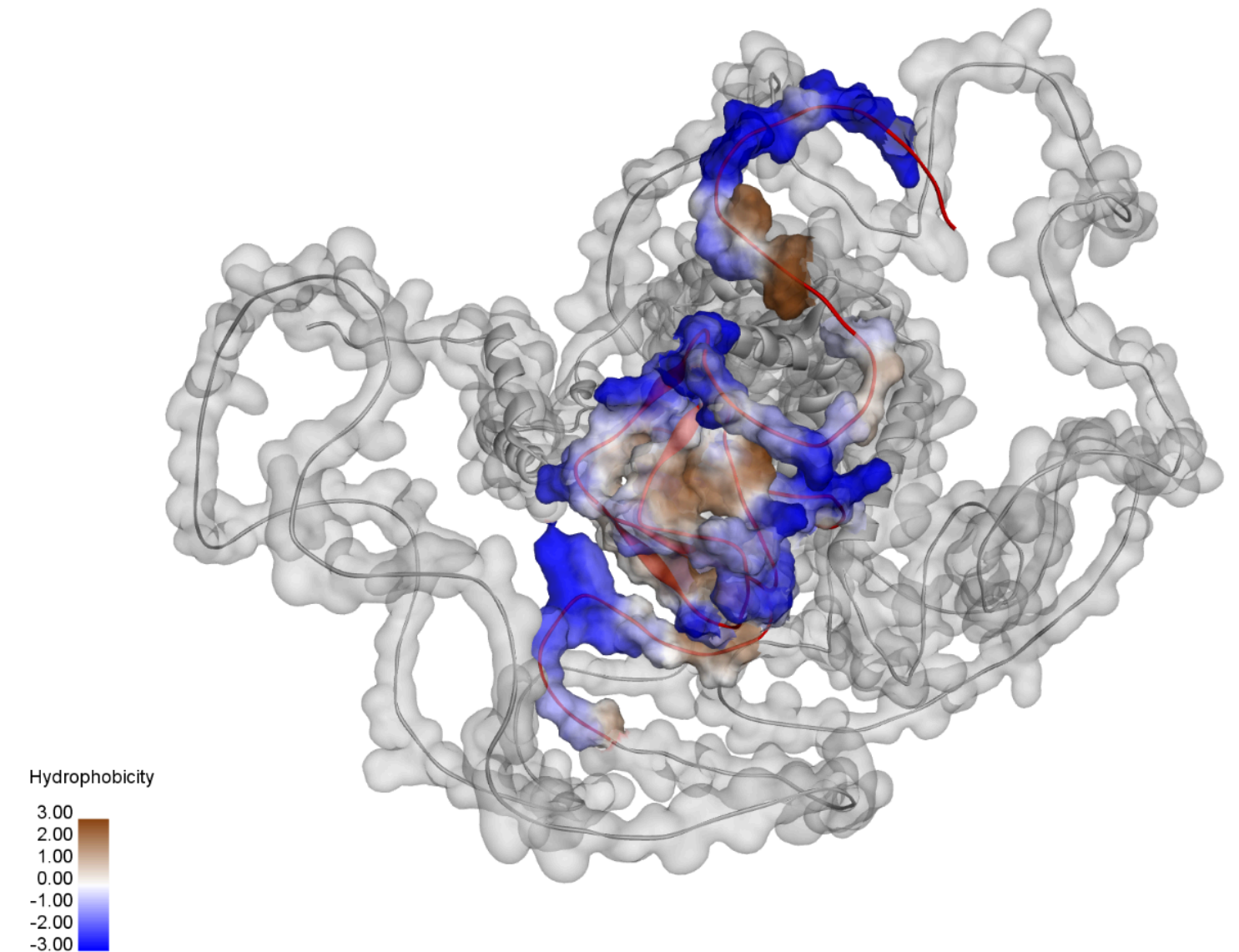


Figure 6

Hydrophobicity and surface exposure analysis of the selected B-cell epitopes. Structural profiling of the 8 linear B-cell epitopes mapped onto the tertiary structure of the final vaccine chimera. The color gradient indicates the hydrophobicity index, shifting from brown (highly hydrophobic) to blue (highly hydrophilic). The localized predominance of blue surface patches indicates that the selected epitope domains possess exceptional hydrophilicity and high solvent accessibility, validation features essential for robust humoral immune recognition.

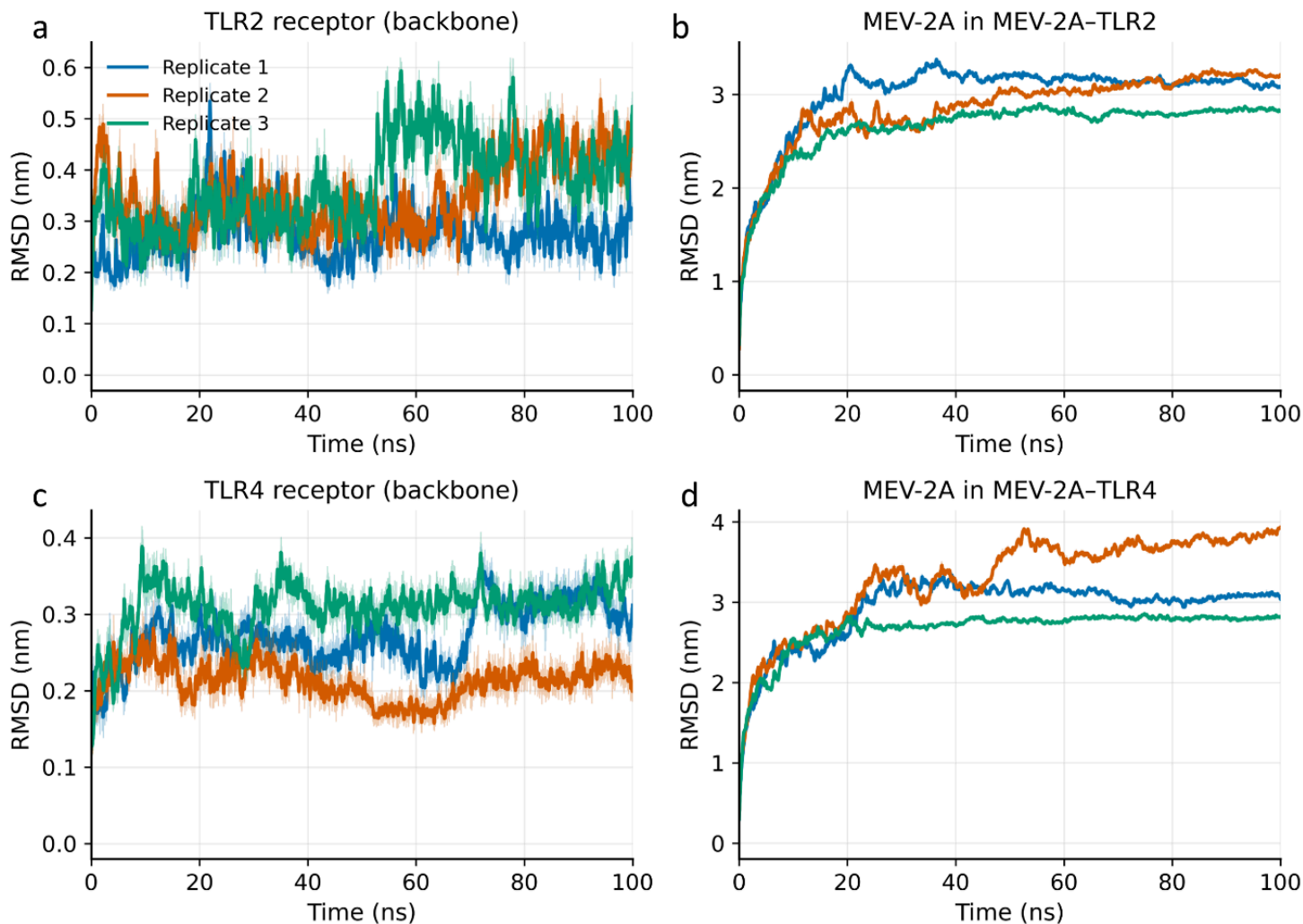


Figure 7

Backbone root-mean-square deviation (RMSD) versus simulation time for the three independent replicates of each complex. (A, C) TLR2 and TLR4 receptor backbones; (B, D) the MEV-2A construct within the MEV-2A-TLR2 and MEV-2A-TLR4 complexes, respectively. Bold lines, 25-frame running averages; faint traces, raw data.

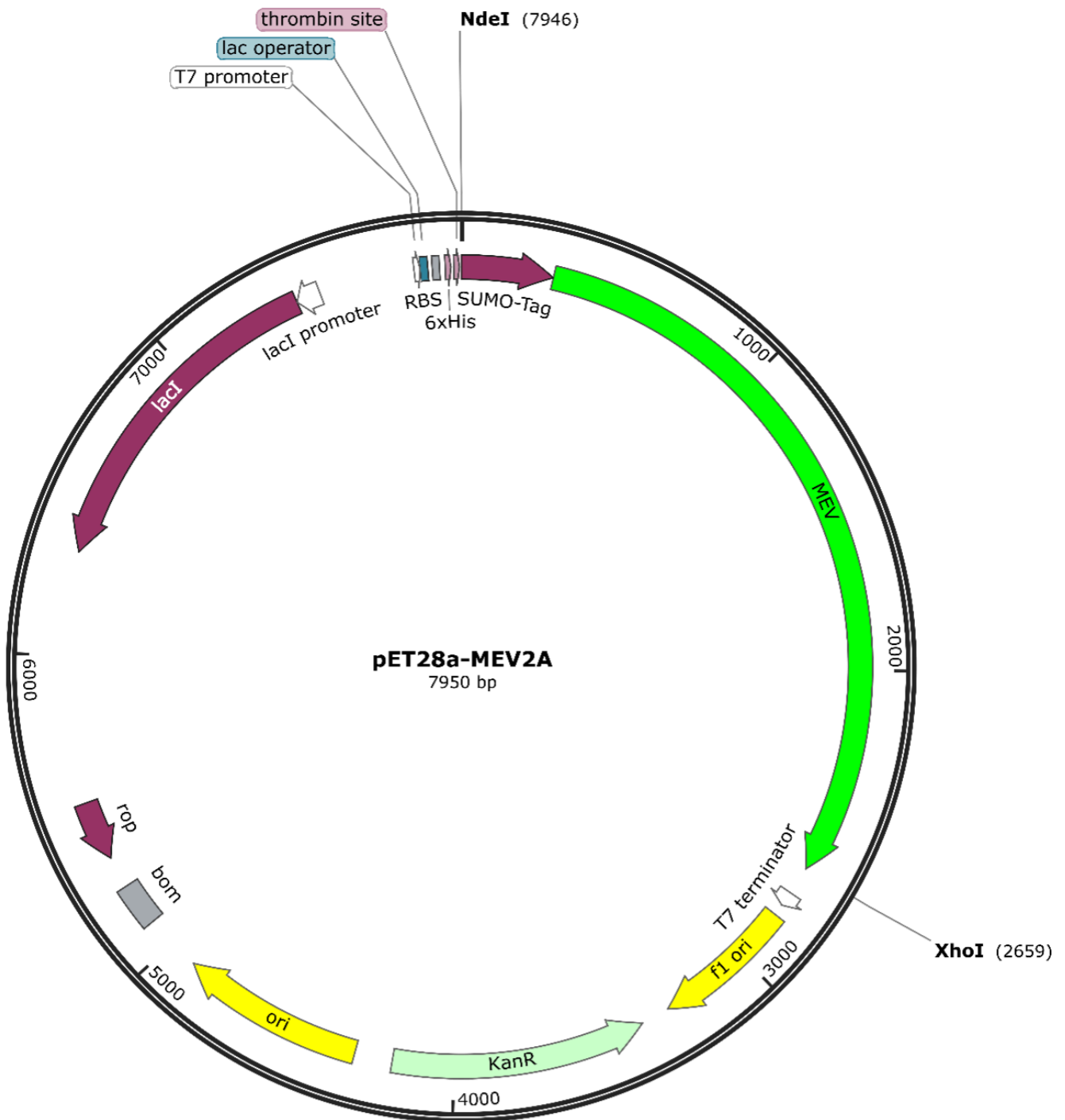


Figure 8

In silico cloning of MEV-2A into the pET28a vector with SUMO and 6X-His-tags in the N-terminal.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial1.xlsx](#)
- [Supplementarymaterial2.xlsx](#)
- [supplementarymaterial3.docx](#)