JJBS

Volume 19, Number 1,March 2026
ISSN 1995-6673
Pages 23 – 33
https://doi.org/10.54319/jjbs/190103

Jordan Journal of Biological Sciences

Structural and antigenic impact of RpfE (Rv2450c) mutations in *Mycobacterium tuberculosis* clinical isolates from South Sulawesi, Indonesia: insights from molecular docking, molecular dynamics, and epitope prediction

Sadiya Ramli[1], Israini Wiyulanda Iskandar[1], Astutiati Nurhasanah[2,*], Nihayatul Karimah[2], Najdah Hidayah[3], Muhammad Nasrum Massi[4,*], Doddy Irawan Setyo Utomo[2], Rizalinda Sjahril[4], Fadhilah Syamsuri[4], Irda Handayani[5]

[1]*Biomedical Science Master Program, Postgraduate School, Faculty of Medicine, Universitas Hasanuddin, Makassar, South Sulawesi 90245, Indonesia;* [2]*Centre for Vaccine and Drugs Research, Health Research Organization, National Research and Innovation Agency (BRIN), Banten 15314, Indonesia;* [3]*Institute for Research and Community Services, Universitas Hasanuddin, Makassar, South Sulawesi 90245, Indonesia;* [4]*Department of Microbiology, Faculty of Medicine, Universitas Hasanuddin, Makassar, South Sulawesi 90245, Indonesia;* [5]*Department of Clinical Pathology, Faculty of Medicine, Universitas Hasanuddin, Makassar, South Sulawesi 90245, Indonesia*

Received: April 18, 2025; Revised: June 26, 2025; Accepted: July 2, 2025

Abstract

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (Mtb), remains a global health threat. Resuscitation-Promoting Factor E (RpfE), encoded by *Rv2450c*, plays a key role in bacterial reactivation and is a promising vaccine and therapeutic target. However, mutations in *Rv2450c* may alter RpfE's structure, function, and antigenicity, affecting immune recognition.

This study analysed *Rv2450c* variations in 40 Mtb clinical isolates from South Sulawesi, Indonesia, using *in silico* approaches to predict structural and functional impacts. Structural modelling, molecular dynamics simulations, epitope binding predictions, and antigenicity analyses were conducted, alongside molecular docking with TLR4 and ligand-binding assessment with NAG3.

Three missense mutations were identified: Arg126Gln (ubiquitous), Thr20Arg (lineage-specific), and Pro35Ser (novel), resulting in three RpfE variants: M1 (containing double mutations Arg126Gln, Thr20Arg), M2 (containing double mutations Arg126Gln, Pro35Ser), and M3 (containing single mutation Arg126Gln). These mutations induced moderate conformational changes with minimal effects on stability or immunogenicity. VaxiJen analysis indicated moderate antigenicity across all variants. Notably, M2 showed the strongest binding to TLR4, while docking with NAG3 also showed affinity alterations.

These findings highlight the genetic diversity of *Rv2450c* and its potential implications for immune recognition and vaccine design. Experimental validation is needed to confirm these predictions.

Key words: molecular modelling, Mycobacterium tuberculosis, mutation, RpfE, Rv2450c

1. Introduction

Tuberculosis, caused by Mycobacterium tuberculosis (Mtb), is a significant global health concern, with more than ten million new cases and approximately 1.25 million deaths reported in 2023 (WHO, 2024). The pathogen's ability to persist in dormancy and reactivate underlies the chronicity and global burden of TB. Resuscitation-Promoting Factors (Rpfs) are crucial in reactivating dormant bacilli (Kana et al., 2008), functioning alongside partnering proteins in hydrolysis of peptidoglycan. Mtb possesses five *rpf* homologues, *rpfA-E* (Mukamolova et al., 2002), all of which are expressed during *in vitro* growth and detected in human tissues infected by the bacteria (Davies et al., 2008). In addition, *M. tuberculosis* H37Rv with combinations of three deleted *rpf*-like genes showed reduced ability to resuscitate from a 'non-culturable' state and exhibited differential growth attenuation in mice (Downing et al., 2005). RpfB and RpfE are also known to interact with Rpf-interacting protein A (RipA), colocalising at the septa of dividing cells, suggesting a role for the RipA–Rpf complex in peptidoglycan hydrolysis during cell division (Hett et al., 2007). RipA is known to hydrolyse several cell wall substrates and synergises with RpfB (Hett et al., 2008). However, RipA does not interact with RpfA, C, and D, suggesting that these proteins act on different pathways, perhaps involving other RipA-like proteins.

The crystal structure of the RpfE catalytic domain revealed that the protein adopts the characteristic Rpf fold, with narrower catalytic cleft in comparison to Mtb RpfB. Also, unlike RpfB, RpfE's predicted peptide-binding sites are more positively charged at neutral pH (Mavrici et al., 2014). The reversal of the electrostatic potential at the substrate-binding site indicates that RpfE may function optimally at a different pH than RpfB, or preferentially

* Corresponding author. e-mail: astutiati.nurhasanah@brin.go.id; nasrumm@unhas.ac.id

hydrolyse distinct micro-domains of the peptidoglycan (Mavrici et al., 2014).

RpfE's immunogenic properties, including T-cell activation and TLR4 binding (Choi et al., 2015; Park et al., 2021), make it a promising candidate for molecular adjuvants and vaccine development. However, since variations in other genes, such as *accD3* (Jalil et al., 2018), involved in Mtb cell wall biosynthesis, had been previously reported, it is also possible that the mutations in *rpfE* (also known as *Rv2450c)* gene may alter RpfE's structure, function, or antigenicity, potentially affecting bacterial persistence, immune evasion and potentially compromise vaccine efficacy.

Given that RpfE-treated dendritic cells induce PGE2 production through MAPK and COX-2 signalling, which in turn promotes Th1 and Th17 responses via EP4 receptor signalling (Choi et al., 2015), mutations in RpfE could potentially impair this immunostimulatory function. RpfE, like other members of the resuscitation-promoting factor (Rpf) family, possesses muralytic (peptidoglycan-hydrolysing) activity via its conserved catalytic domain, which is structurally similar to lysozymes (Mavrici et al., 2014). Mutations affecting key residues involved in peptidoglycan binding or hydrolysis—such as conserved glutamic acid residues essential for catalytic activity (Mavrici et al., 2014; Squeglia et al., 2013)—could reduce RpfE's ability to generate the muropeptide fragments that act as potent pathogen-associated molecular patterns (PAMPs). These muropeptides are thought to be crucial for dendritic cell activation and the subsequent induction of cytokines like PGE2 (Park et al., 2021). A mutation affecting the conserved catalytic residues could drastically impair muralytic activity, thereby reducing the availability of peptidoglycan-derived ligands for innate immune recognition, leading to suboptimal PGE2 production and consequently weaker Th1/Th17 polarisation. This could represent a potential mechanism of immune evasion by Mtb clinical isolates. Moreover, reduced muralytic activity due to such mutations may simultaneously hinder peptidoglycan remodeling, potentially impairing bacterial resuscitation from dormancy but favouring persistence under immune pressure—a classical trade-off seen in pathoadaptation (Kana et al., 2008). Thus, specific mutations in the catalytic domain of RpfE may confer a selective advantage by diminishing immune detection (via impaired PGE2-mediated Th1/Th17 priming) while modulating the bacterial life cycle balance between dormancy and active growth.

This study is the first to characterize RpfE mutations in clinical Mtb isolates from South Sulawesi, Indonesia, using in silico functional prediction. Indonesia is a high-incidence region reporting over one million TB cases in 2024 (Arlinta, 2024). Using in silico analyses, we identified mutations in *Rv2450c* and evaluated their effects on protein structure, antigenicity, and enzymatic activity. These findings provide insights into the role of RpfE variations in TB pathogenesis and implications for vaccine and therapeutic development.

2. Materials and Methods

2.1. Sample Collection

This study analysed 40 *M. tuberculosis* clinical isolates (Supplementary Table 1) stored at the HUMRC Laboratory, Universitas Hasanuddin, Makassar, Indonesia. The isolates originated from the South Sulawesi Referral Laboratory for Tuberculosis Culture and Drug Susceptibility Testing (DST), Faculty of Medicine, Universitas Hasanuddin, and were collected from patients with confirmed active pulmonary tuberculosis between 2016 and 2018. Each isolate was cultured, species identified, and DST-ed for first- and second-line anti-tuberculosis drugs. Based on DST results, isolates were categorised as drug-sensitive, rifampicin-resistant, multidrug-resistant (MDR), or extensively drug-resistant (XDR). Isolates that failed DNA quality standard were excluded from further molecular analysis. Ethical approval for the use of human-derived samples was granted by the Research Ethics Committee of Universitas Hasanuddin (Ethics Approval No: 678/UN4.6.4.5.3.1/PP36/2023). All procedures involving *M. tuberculosis* cultures were carried out in a certified Biosafety Level 3 (BSL-3) facility at the HUMRC Laboratory. Handling of live bacteria—including inoculation, subculturing, and sample processing—was strictly conducted within a Class II Biological Safety Cabinet (BSC) to minimise the risk of aerosol exposure. The laboratory was equipped with HEPA-filtered negative pressure airflow systems, and access was restricted to personnel with appropriate training.

2.2. DNA extraction

Genomic DNA was extracted using the gSYNC™ DNA Extraction Kit (Geneaid™), following the manufacturer's instructions with slight modifications to accommodate the characteristics of mycobacterial cells. Approximately 200 µL of bacterial pellet was collected from Mtb cultures grown in Löwenstein–Jensen (LJ) medium after 4 to 6 weeks of incubation at 37°C. The pellet was transferred into a sterile 1.5 mL microcentrifuge tube and resuspended in 200 µL of PBS. The suspension was heat-inactivated at 95°C for 30 minutes to ensure biosafety, then centrifuged at 13,000 rpm for 5 minutes. The supernatant was discarded, and the pellet underwent enzymatic lysis by adding 200 µL of GS1 buffer and 20 µL of Proteinase K, followed by incubation at 60°C for 30 minutes.

Following lysis, 200 µL of GS2 buffer was added, mixed thoroughly, and incubated at room temperature for 10 minutes. DNA binding was achieved by adding 200 µL of absolute ethanol, mixing well, and transferring the mixture into the gSYNC™ DNA mini column. After centrifugation at 14,000 rpm for 1 minute, the column was washed with 400 µL of W1 buffer and twice with 600 µL of wash buffer (ethanol-based). The column was dried through additional centrifugation to remove any residual ethanol. DNA was eluted by adding 50 µL of pre-warmed elution buffer (provided in the kit) directly to the center of the column membrane, incubated at room temperature for 3 minutes, and centrifuged at 14,000 rpm for 1 minute. The eluted genomic DNA was stored at -20°C until further analysis. DNA purity and concentration were assessed using a NanoDrop™ 2000 spectrophotometer (Thermo Scientific™).

2.3. PCR Amplification of RpfE Gene

The Rv2450c (RpfE) gene was amplified using PCR with primers designed via Primer-BLAST based on the Mtb H37Rv reference genome (NC_000962.3). The PCR was performed with an annealing temperature of 58°C, resulting in an amplicon size of 683 bp. The full primer sequences, PCR cycling parameters, and electrophoresis conditions are detailed in Table 1.

Table 1. *Mtb* RpfE PCR amplification primers

Gene	Identifier	Size	Primer sequence	Tm (°C)	Amplicon size
RpfE	Rv2450c	519	F 5' GTCACGGGTTTGACGCTACT 3'	58°C	683
			R 5' AATGCACCTGGCAGCTAACC 3'		
PCR Conditions		95°C (3 min); 30 cycles of 95°C (30 s), 58°C (15 s), 72°C (1 min); final extension at 72°C (10 min)			
Gel Electrophoresis		0.8% agarose gel, 1X TAE buffer, 5 - 7V/cm, 40 mins; UV transilluminator			

2.4. DNA Sequencing and Data Analysis

Amplicons were purified and subjected to bidirectional Sanger sequencing using the BigDye® Terminator v3.1 Cycle Sequencing Kit on an ABI PRISM® 3730xl DNA Analyzer platform at 1st BASE Laboratories (Selangor, Malaysia). The same primers as used for PCR amplification (Table 1) were used for the sequencing. Resulting chromatograms were manually checked for quality and contigs were assembled using BioEdit v7.2.5 (Hall, 1999). Contigs were aligned to the H37Rv reference genome (NC_000962.3) and compared with GenBank database sequences (Supplementary Table 2). Homologous sequences and protein variants were identified using BLASTp (Altschul et al., 1990).

2.5. Phylogenetic Analysis

Phylogenetic trees were constructed using the Tamura 3-parameter (T92) model in MEGA 11 (Tamura et al., 2021), incorporating six *Rv2450c* sequences from different Mtb lineages (Supplementary Table 2)

2.6. Construction and Validation of the RpfE Model Structure

RpfE protein structures were modelled using I-TASSER (http://zhang.bioinformatics.ku.edu/I-TASSER) (Yang & Zhang, 2015) and AlphaFold (https://alphafold.ebi.ac.uk/) (Jumper et al., 2021), then validated using SAVES v6.0 server (https://saves.mbi.ucla.edu/) (Laskowski et al., 1993).

2.7. Structural Alignment

Structural similarity was assessed by aligning modelled structures, with Root Mean Square Deviation (RMSD) and coverage calculated using PyMOL 3.0.3.

2.8. Prediction of Protein Stability and Solubility

The stability and solubility of the proteins were computed using Protparam (https://web.expasy.org/protparam/) (Gasteiger E. et al., 2005).

2.9. Prediction of T cell epitopes

Cytotoxic T Lymphocyte (CTL) and Helper T Lymphocyte (HTL) epitopes were predicted using IEDB tools (http://tools.iedb.org/mhci/ and http://tools.iedb.org/mhcii/), with NetMHCPan 4.1 BA (HLA-I) and NetMHCIIPan 4.0 BA (HLA-II) (Reynisson et al., 2020) methods, and HLA reference sets covering >97% (HLA-I) (Weiskopf et al., 2013) and 99% (HLA-II) (Greenbaum et al., 2011) of the population. IC50 thresholds <500 nM for binding affinity, <50 nM for strong binders were used.

2.10. Prediction of Epitope Antigenicity

Antigenicity was assessed using VaxiJen 2.0 (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) (Doytchinova & Flower, 2007), with 0.4 threshold indicating antigenicity, and scores > 1.0 indicating high antigenicity, whereas 0.4 < scores < 1.0 indicating intermediate antigenicity.

2.11. Prediction of Signal Peptide

Signal peptides were identified using SignalP 6.0 (https://services.healthtech.dtu.dk/services/SignalP-6.0/) (Teufel et al., 2022).

2.12. TLR4 – RpfE Molecular Docking

Potential binding interfaces of TLR4 (PDB ID: 3FXI)-H37Rv RpfE was initially predicted by ClusPro (Kozakov et al., 2017) blind docking. The interfaces were then used as restraints for guided TLR4-RpfE (H37Rv and the mutants M1, M2, M3) dockings in HADDOCK 2.4 (Honorato et al., 2024). LRR positions were from UniProt TLR4_Human (Accession Number O00206). Interaction interfaces were analysed with PDBsum Generate (Laskowski, 2022) and the Protein–Ligand Interaction Profiler (PLIP) (Adasme et al., 2021), with visualisations prepared using PyMOL.

2.13. RpfE Catalytic Domain – NAG3 Ligand Molecular Docking

Docking of RpfE with its ligand, NAG3 (PDB ID: 4KPM) (Mavrici et al., 2014), was performed using AutoDock Vina v1.2.3 (parameters in Supplementary Table 3). The docking results were analysed for binding energy, pose or conformation relative to the binding site, and molecular interactions involving residues within 5Å. Validation was conducted by superimposing the re-docked NAG3 structure with the original 4KPM (RMSD <2Å). Models were visualised using BIOVIA Discovery Studio 2021.

2.14. Molecular Dynamics (MD) Simulations

MD simulations (100 ns) were conducted using GROMACS v2023.2 (Abraham et al., 2023), with TIP3P water and ionic conditions. Binding free energy was calculated using gmx_MMPBSA from trajectory data.

3. Results

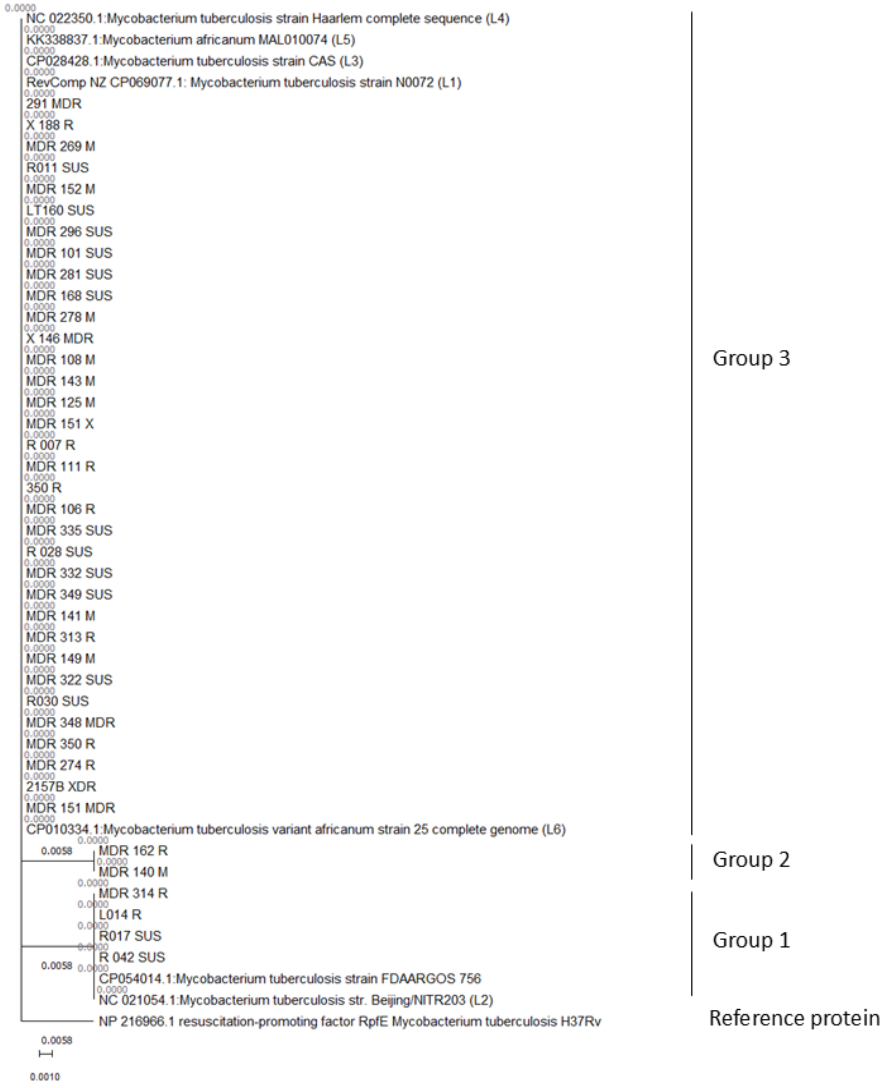
3.1. PCR Amplification and Sequencing of the Amplicons

PCR amplification of the RpfE gene from 40 clinical isolates produced 683-bp amplicons. Sanger sequencing identified three nucleotide substitutions: 377G>A (all

isolates), 59C>G (4 isolates), and 103C>T (2 isolates) (Figure 1, Supplementary Figure 1). These substitutions resulted in three missense mutations (Supplementary Figure 1D)—Thr20Arg, Pro35Ser, and Arg126Gln. All isolates carried Arg126Gln, with subsets carrying either Thr20Arg (M1, hence Arg126Gln, Thr20Arg) or Pro35Ser (M2, Arg126Gln, Pro35Ser), or neither (M3, Arg126Gln). Representative isolates of each mutant were submitted to GenBank with the following accession numbers M1 (isolate R042SUS, Accession: PV075057), M2 (MDR162R, Accession: PV075056), M3 (MDR322SUS, Accession: PV075058).

3.2. Sequence Diversity

Sequence alignment with six Mtb lineages (L1–L6) revealed lineage-specific mutation patterns (Supplementary Figure 2). The Arg126Gln mutation (M3) was observed in lineages L1, L3, L4, L5, and L6 but absent in H37Rv. M1 (Arg126Gln, Thr20Arg) was found in Beijing/NITR203 (L2) and FDAARGOS_756, while M2 (Arg126Gln, Pro35Ser) was unique to this study. BLASTp confirmed that Pro35Ser is novel, not reported in GenBank.



Isolate code	Nucleotide Substitution	Amino acid Substitution	Percentage of Clinical Isolates Containing the Sequence Variation
All clinical isolates	377G>A	*Arg*126*Gln*	100
R017SUS R042SUS L014R MDR314R	59C>G	*Thr*20*Arg*	10
MDR162R MDR140M	103C>T	*Pro*35*Ser*	5

Figure 1. Mutation sites (Arg: Arginine; Gln: Glutamine; Thr: Threonine; Pro: Proline; Ser: Serine; A: Adenine; C: Cytosine; G: Guanine; T: Thymine) found in the clinical Mtb isolates and phylogenetic tree of the RpfE gene, constructed using sequences from H37Rv, the mutant sequences identified in this study, and previously reported sequences from the NCBI-GenBank database. The divergence scale at the bottom of the tree represents the genetic divergence in terms of the number of substitutions per site. A value of 0.0010 indicates a genetic difference of 0.0010 substitutions per DNA site between the compared strains.

3.3. Phylogenetic Analysis

Maximum Likelihood phylogenetic analysis grouped isolates into four clusters (Figure 1). M1 was linked to FDAARGOS_756 and Beijing/NITR203 (L2), M2 to a distinct cluster, and M3 to N0072 (L1), CAS (L3), Haarlem (L4), *M. africanum* MAL010074 (L5), and *M. tuberculosis* africanum strain 25 (L6). H37Rv formed a separate cluster, confirming evolutionary divergence.

3.4. Impact of Nucleotide Variations on RpfE Protein Structure and Antigenicity

3.4.1. RpfE Modelling and Validation

I-TASSER-derived 3D models (Supplementary Figure 3) demonstrated higher reliability than AlphaFold models (Supplementary Table 5). MD simulations (100 ns) confirmed stability, with Ramachandran plots indicating that the majority of residues were located within favoured regions.

3.4.2. Effect of Amino Acid Variations on RpfE Protein Structure, Stability, Solubility, and Signal Peptide Position

Secondary structure analysis (Supplementary Figure 4) showed no significant changes in the number of helices (three) or beta strands (three) across the reference and mutant proteins, though beta strand lengths varied. M1 retained the same beta strand positions as the reference (6–13, 17–19, and 123–125), while M2 exhibited slightly shorter strands (6–12, 17–18). M3 displayed an extended N-terminal beta strand (5–13), with other strand positions matching those of the reference and M1.

Three-dimensional structural alignments (Supplementary Figure 5) revealed RMSD scores and coverage of 1.70 Å (74.84%) for M1, 1.32 Å (56.79%) for M2, and 0.92 Å (64.65%) for M3, reflecting moderate deviations from the reference.

According to the Protparam results (Table 2), all the proteins might be considered unstable, with Instability Indexes exceeding 40, and they are likely hydrophilic, based on the negative GRAVY scores.

SignalP 6.0 predicted the signal peptide position (amino acids 1–28) and cleavage site (between residues 28 and 29) to be unchanged across all mutants compared to the reference RpfE protein. Detailed results are provided in Supplementary Figure 6.

Table 2. Instability index and GRAVY score of RpfE reference and variant proteins, as predicted using Protparam

Protein	Instability Index	Grand average of hydropathicity (GRAVY) score
Reference protein (H37Rv)	47.84	-0.269
M1 (Thr20Arg, Arg126Gln)	50.73	-0.285
M2 (Pro35Ser, Arg126Gln)	49.29	-0.259
M3 (Arg126Gln)	48.17	-0.263

3.4.3. Effect of Amino Acid Variations on Protein's Antigenicity

The VaxiJen scores for the RpfE reference protein and its M1, M2, and M3 mutants were 0.8023, 0.7946, 0.8068, and 0.7996, respectively, indicating medium antigenicity (0.4 < VaxiJen scores < 1.0).

3.4.4. Effect of Amino Acid Variations on T-cell Epitopes

The overall antigenicity of mutated epitopes is largely unchanged, even though binding affinities with MHC-I and MHC-II vary with mutations. Amino acid variations in the RpfE protein influenced predicted MHC-I and MHC-II T-cell epitopes affinity towards HLA alleles (Supplementary Files 1–8), although the effect seems to be allele and mutation specific, and to a varying degree.

The reference RpfE protein (H37Rv) contained ten high-affinity MHC-I epitopes (IC50 < 50 nM), none of which includes the positions of amino acid variations (Thr20, Pro35, or Arg126). Increasing the IC50 threshold to <500 nM, showed Thr20 and Arg126 as part of medium-affinity epitopes, but not Pro35. The Thr20Arg mutation (Supplementary Table 6) reduced binding affinity for three epitopes across five alleles but enhanced seven others, generating two new strong-binding epitopes (IC50 < 50 nM) with medium antigenicity (0.4 < VaxiJen scores < 1.0). The Arg126Gln mutation (Supplementary Table 7) created three intermediate-affinity epitopes but no strong-binding ones. The novel Pro35Ser mutation, however, had limited impact on binding affinity, indicating a minimal role in MHC-I-mediated recognition.

The reference RpfE protein contained 108 high-affinity (IC50<50 nM) MHC-II epitopes, 25 of which carry Thr20, and two carry Arg126. Pro35 was found in intermediate-affinity (IC50 < 500 nM), highly antigenic epitopes (VaxiJen scores > 1.0). The Thr20Arg mutation reduced epitope-MHC-II combinations from 300 to 289, generated 17 new high-affinity epitopes (Supplementary Table 8), and reduced the binding of a strong-binding epitope, with little impact on antigenicity. The Arg126Gln mutation (Supplementary Table 9) converted two high-affinity epitopes into intermediate binders but preserved their antigenicity. The novel Pro35Ser mutation (Supplementary Table 10) caused two intermediate-affinity epitopes to lose binding capability, though antigenicity remained high.

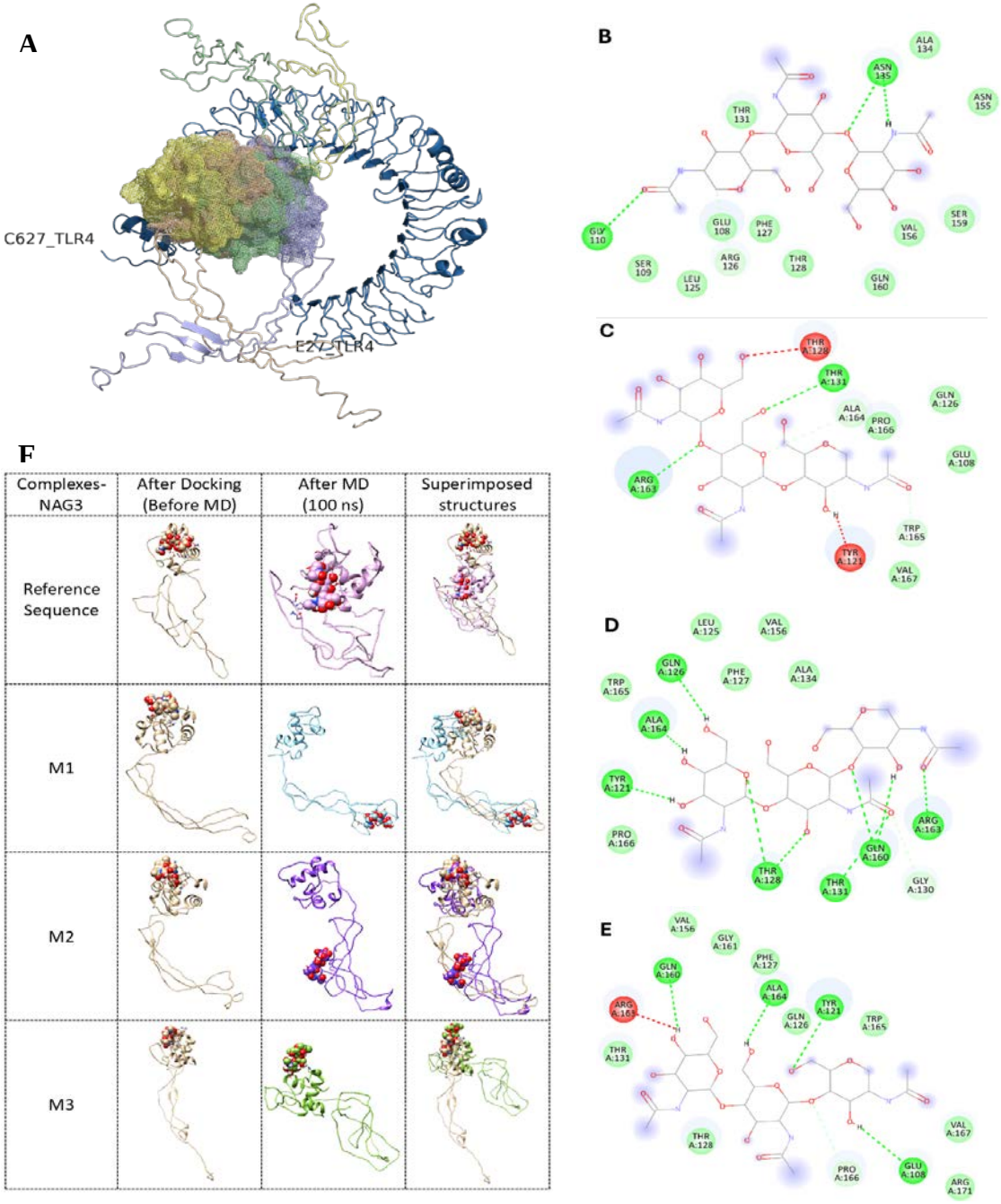


Figure 2. A. Complex of TLR4 (blue ribbon) - RpfE from H37Rv (wheat), M1 (pale green), M2 (pale yellow), M3 (light blue). RpfE residues interacting with NAG3 in B. Reference protein, C. M1, D. M2, and E. M3. Red colour indicates residues showing unfavourable interactions, whereas green indicates favourable interactions.

3.4.5. Effect of Amino Acid Variations on TLR4 Binding

Molecular docking revealed notable differences in RpfE-TLR4 binding, interaction residues (Table 3), and complex structures (Figure 2A). HADDOCK scores indicated stronger binding for M2 and M3 compared to the reference and M1. M2 showed the highest structural stability with the lowest RMSD, while the reference and M3 exhibited greater flexibility, indicating less stable interactions.

Van der Waals energy was most favourable for M3, while M2 and M1 had the strongest electrostatic interactions, surpassing the reference. M3 also had the lowest desolvation energy, enhancing complex stability. Buried surface area (BSA) was larger in all mutants compared to the reference, correlating with stronger binding.

Hydrogen bond analysis revealed that both reference and mutant RpfE proteins interact with TLR4's LRR14–LRRCT regions (amino acids 456–606), primarily via the RpfE catalytic domain (residues 98–172). Mutations did not significantly alter these interactions, although M3 showed a slightly expanded binding interface, including His431 in LRR13.

Despite the absence of negative control or decoys, the docking result revealed that all RpfE proteins interact with TLR4's LRR14–LRRCT regions through hydrogen bonds, in which the interaction with LRR region of TLR4 is known to mediate interactions with several microbial ligands.

Table 1. Results of molecular docking in HADDOCK and residues interaction between TLR4 and RpfE proteins from reference strain (H37Rv), and strains with M1, M2, and M3 mutations

	H37Rv			M1			M2			M3		
HADDOCK score	-118.2 +/- 5.2			-116.8 +/- 10.2			-131.0 +/- 9.7			-124.7 +/- 2.5		
Cluster size	7			6			9			19		
RMSD from the overall lowest-energy structure	23.7 +/- 0.2			14.8 +/- 0.1			0.8 +/- 0.7			19.3 +/- 0.1		
Van der Waals energy	-65.9 +/- 11.6			-51.0 +/- 7.3			-53.6 +/- 3.1			-81.6 +/- 3.3		
Electrostatic energy	-191.5 +/- 34.3			-353.7 +/- 56.0			-386.4 +/- 37.7			-132.3 +/- 19.8		
Desolvation energy	-16.4 +/- 2.5			2.9 +/- 6.1			-4.3 +/- 1.8			-21.0 +/- 1.3		
Restraints violation energy	24.4 +/- 19.6			21.4 +/- 20.9			42.0 +/- 14.9			44.2 +/- 29.6		
Buried Surface Area	1965.6 +/- 121.4			2268.9 +/- 131.0			2144.9 +/- 106.6			2245.1 +/- 140.1		
Z-Score	-2			-2.2			-1.8			-1.6		
	Receptor	Ligand	Bond Length (Å)	Receptor	Ligand	Bond Length (Å)	Receptor	Ligand	Bond Length (Å)	Receptor	Ligand	Bond Length (Å)
Hydrogen Bonds	Tyr-46	Glu-66	2.78	Gly-363	Glu-88	2.67	Lys-362	Pro-80	2.99	Glu-42	Leu-33	2.68
	His-456	Arg-170	2.7	Asn-365	Glu-88	2.8	His-529	Asn-155	2.76	His-431	Arg-163	2.94
	Gln-507	Ser-159	3.22	Lys-388	Asp-29	2.6	Ala-610	Arg-163	2.66	His-456	Arg-163	2.95
	Asn-531	Arg-158	3.21	His-529	Gln-106	2.78	Gln-616	Arg-163	2.95	Asp-502	Asn-119	3.34
	Leu-553	Arg-158	2.87	His-555	Gln-106	3.16	Gln-616	Gln-160	2.9	Gln-505	Glu-108	2.69
	Arg-606	Ser-98	3.15	Gln-578	Ala-103	3.29				His-529	Cys-107	3.24
	Glu-608	Ala-96	3.23	Arg-606	Arg-171	2.76				His-555	Arg-171	3.24
										Gln-578	Ser-109	2.94
										Arg-606	Ser-109	3.04
Salt Bridges				Lys-388	Asp-29	2.6	Arg-264	Asp-81	2.64	His-529	Glu-108	3.4
				Lys-435	Asp-42	2.89	Arg-264	Asp-81	2.73			
				Glu-608	Arg-171	2.64	Lys-341	Asp-30	3.2			
							Lys-362	Asp-30	2.57			
							Arg-382	Glu-88	3.79			
							His-431	Glu-89	3.88			
							His-456	Glu-89	2.68			
							Glu-608	Arg-133	2.76			

Table 2. Binding affinities from molecular docking, ligand displacement and structural stability metrics from MD simulations, and MM_PBSA calculation. [a] All units are in Å (Amstrong); [b] Radius of Gyration, [c] All units are in kJ/mol.

	Molecular Docking	Molecular Dynamics						MM_PBSA Calculation			
Variant	Binding Affinity (kcal/mol)	Ligand Displacement[a]	Average[a] $RMSD_{Backbone}$	Average[a] $RMSD_{Ligand}$	RMSD Stability	Average[a] RMSF	Average[a] RoG[b]	Van der Walls Energy[c]	Electrostatic Energy[c]	Polar Solvation Energy[c]	ΔG Binding Free Energy[c]
Reference (H37Rv)	-6.05	3.9 → 2.4	0.84	1	Stable after 20 ns	0.65 (Lowest fluctuations)	1.93 (Stable)	-33.48 ± 1.24	-42.55 ± 12.00	-4.05 ± 0.21	-14.53 ± 6.40
M1	-5.28	2.7 → 57.4	1.57	1.07	Unstable	1.10 (Fluctuations observed)	2.66 (Unstable)	-29.49 ± 1.28	-37.15 ± 11.62	-3.67 ± 0.06	-8.07 ± 5.28
M2	-5.02	2.2 → 31.3	1.59	1.05	Unstable	1.16 (Fluctuations observed)	2.85 (Unstable)	-18.76 ± 1.43	-4.12 ± 1.84	-2.60 ± 0.21	-5.78 ± 2.98
M3	-4.65	1.7 → 12.7	1.82	1.17	Stabilised early, but slightly fluctuated after 80 ns	0.99 (Moderate fluctuations)	2.34 (Partial stability)	-31.19 ± 2.50	-63.51 ± 3.93	-4.29 ± 0.07	-16.52 ± 4.27

3.4.6. Effect of Amino Acid Variations on NAG3 Binding to RpfE Catalytic Domain

Molecular docking confirmed NAG3 binding to the RpfE catalytic domain across all variants, with the reference protein (H37Rv) exhibiting the strongest binding affinity (-6.05 kcal/mol), followed by M1, M2, and M3 (Table 4). Hydrogen bonding analysis revealed variations in binding interactions (Figure 2B - E), with M2 showing the highest number of hydrogen bonds (10), while M1 and M3 displayed unfavourable interactions that may weaken binding stability.

MD simulations revealed varying ligand stability (Figure 2F). The metrics of ligand displacement and structural stability are presented in Table 4.

RMSD analysis (Supplementary Figure 7A) confirmed that the reference protein stabilised after 20 ns with minimal fluctuations, while M1 and M2 exhibited continuous instability. Ligand RMSD fluctuations (Supplementary Figure 7B) confirms weak interactions particularly between NAG3 and M2. RMSF analysis (Supplementary Figure 7C) indicated that key fluctuations occurred at Asp42, Gly62, Asn119, Ser140, and Gly172, with mutation sites (residues 20, 35, 126) remaining stable. RoG analysis (Supplementary Figure 7D) showed structural stability in the reference protein, followed by M3, which stabilised after ~20 ns.

MM-PBSA calculations (Table 4) revealed lower binding free energies for the reference and M3 complexes, indicating higher binding stability compared to M1 and M2.

4. Discussion

This study examined the structural, immunological, and molecular implications of naturally occurring RpfE (Rv2450c) mutations in *Mycobacterium tuberculosis*, focusing on three key polymorphic variants. All observed mutants shared the Arg126Gln substitution, with additional mutations at Thr20Arg (M1) and Pro35Ser (M2), highlighting both lineage-specific and novel variations. M2 variation was observed in two of 40 isolates (Figure 1). If this proportion reflects its true frequency in the population, this might suggest that this unique 5% of the Mtb population comprises a locally evolved strain.

Notably, Arg126Gln was found across diverse lineages (L1, L3, L4, L5, and L6), in H37Rv derivatives (Santoro et al., 2017), African samples (Osei-Wusu et al., 2021), and laboratory strains (Ioerger et al., 2010), though absent in the H37Rv reference, suggesting it may represent an evolutionary adaptation with broad geographical persistence. Its presence in isolates from patients undergoing prolonged isoniazid and rifampicin treatment (Li et al., 2018) suggests a possible selective advantage. Since the increase in multidrug-resistant (MDR) microorganisms triggering infections is growing worldwide and becoming more serious in developing countries (Al-Ouqaili et al., 2025; Owaid & Al-Ouqaili, 2024), we attempted to correlate the presence of this mutation with DST data (Supplementary Table 1). However, the mutation was not restricted to MDR, XDR, or rifampicin-resistant isolates. The M1-type double mutation (Thr20Arg, Arg126Gln) was primarily associated with the Beijing/NITR203 (L2) strain, consistent with previous reports identifying Thr20Arg as a lineage-defining marker (Tantivitayakul et al., 2020), found in 28.7% of Indonesian strains (Atavliyeva et al., 2024).

4.1. Structural Impact and TLR4 Interactions

Molecular modelling revealed that M2 displayed the strongest affinity to TLR4, as evidenced by HADDOCK scores and docking RMSD values, indicating a potential increase in innate immune engagement. There is no definitive biological threshold for HADDOCK score. The score was used primarily for guiding the comparative purpose between a set of simulations, in which the lower the score, the more favourable the binding.

The enlargement of buried surface area and enhanced electrostatic interactions across all mutants support this observation. Interestingly, M3's docking revealed an expanded interface involving LRR13, which could represent a distinct binding mode for RpfE–TLR4 interaction. These findings are consistent with the critical role of TLR4 in recognising Mtb proteins through leucine-rich repeat domains independently of MD2, as previously described (Kleinnijenhuis et al., 2011; Park et al., 2021). Such enhanced receptor interactions may influence the early immune activation cascade, including proinflammatory cytokine production and antigen processing. As a known TLR4 agonist, RpfE enhances dendritic cell maturation and cytokine production (Park et al., 2021).

4.2. MHC Presentation and Antigenicity

Our study indicated that the overall antigenicity of mutated epitopes is unchanged, despite variations of MHC-I and MHC-II binding affinities observed with the mutations. These findings suggest that while mutations alter epitope presentation and binding affinity, they do not drastically enhance the immunogenic profile of the RpfE protein.

The epitope analysis showed that Thr20 and Arg126 are embedded within several MHC-II epitopes, with M1 notably generating new high-affinity binders while slightly reducing overall MHC-II epitope diversity. Importantly, this mutation also gave rise to new strong-binding MHC-I epitopes, suggesting an immunogenic shift that may enhance T-cell responses. Previous studies have demonstrated that single amino acid polymorphisms can dramatically influence epitope presentation and recognition (Bui et al., 2007). The Pro35Ser substitution in M2, which is novel and unique to this study, had minimal impact on MHC-binding affinity but preserved high antigenicity, indicating a potentially conserved immunological function. Meanwhile, M3's Arg126Gln substitution created intermediate-affinity MHC-I epitopes and modestly weakened MHC-II affinity but retained immunogenic potential. These observations underscore the functional plasticity of RpfE epitopes and the importance of characterising lineage-specific mutations in vaccine design (Comas et al., 2010; Panda et al., 2023).

VaxiJen was used as part of integrative strategies to define immunological properties of each variants, to distinguish the antigenicity between the variants and, to specifically define the range of their antigenicity. Although the result showed only slight differences between variants, it gave us information that they possessed medium range of antigenicity. Based on the Vaxijen score and Haddock

result, M2 variant is the most antigenic and M1 is the least antigenic.

4.3. Ligand Binding to NAG3 and Structural Stability

Molecular dynamics simulations, performed following docking of NAG3 to RpfE variants, revealed notable differences in ligand-binding stability and structural behaviour. While the reference RpfE protein maintained structural integrity and stabilised after 20 ns of simulation, all mutant variants displayed reduced stability, with M1 and M2 showing the most pronounced fluctuations. These findings suggest that the observed mutations may compromise the receptor's conformational stability upon ligand binding, possibly affecting biological functions such as peptidoglycan hydrolysis and reactivation from dormancy (Rosser et al., 2017). Interestingly, despite M2 forming the highest number of hydrogen bonds, it exhibited weak overall interactions and instability, highlighting that increased hydrogen bonding alone does not ensure structural resilience. M3 showed partial stability but reduced affinity, suggesting a potential compensatory adaptation. Overall, the lower MM-PBSA binding free energies across mutants reflect a trend of weakened RpfE–ligand interactions, which may impair physiological ligand recognition and catalytic efficiency.

These mutations may disrupt RpfE's catalytic function, which is critical for bacterial cell wall remodelling and resuscitation. Arg126, previously identified as part of RpfE's catalytic cleft (Mavrici et al., 2014), facilitates glycan hydrolysis in the peptidoglycan (PG) matrix (Romano et al., 2023)**,** protecting bacteria from environmental stress. RpfE also interacts with RipA, a key peptidase in PG remodelling and cell division (Hett et al., 2008). Mutations affecting substrate binding could impair these functions, potentially altering bacterial dormancy and resuscitation.

4.4. Impact on Strain Tracking and Epidemiology

The presence of polymorphisms in antigen-encoding genes, such as rpfE, may alter epitope structures, potentially affecting immune recognition. This variation could lead to reduced vaccine efficacy if the candidate vaccine is based on the reference strain H37Rv, which may not fully represent circulating clinical isolates in South Sulawesi. Our findings highlight the importance of incorporating region-specific genetic data in the rational design of next-generation TB vaccines, ensuring the inclusion of conserved and immunodominant epitopes prevalent in local strains, hence, the importance of strain tracking. The polymorphisms could serve as molecular markers for strain tracking and surveillance, aiding in the identification of local transmission patterns and the emergence of drug-resistant *M. tuberculosis* lineages.

5. Conclusions

Overall, our findings suggest that while these mutations contribute to sequence diversity and lineage specificity, their impact on protein structure and antigenicity is relatively moderate. The enhanced TLR4 binding in M2 and M3 could influence immune recognition, whereas the observed changes in ligand binding may have functional consequences for bacterial physiology.

Further experimental validation, including enzymatic activity assays and immune response studies, is required to elucidate the biological significance of these mutations and their potential role in Mtb pathogenesis. In addition, the observation of variation of antigen-encoding genes in clinical isolates emphasises the necessity of strain tracking as well as genetic surveillance, to ensure new tuberculosis vaccines have suitable population coverage.

Acknowledgement

The authors sincerely thank the National Research and Innovation Agency (BRIN) for funding this research, and StudiBio for their valuable computational resources and continuous support throughout the study.

Data Availability Statement

All supplementary data can be downloaded from the following links:

- Supplementary Tables and Figures
- Supplementary File 1
- Supplementary File 2
- Supplementary File 3
- Supplementary File 4
- Supplementary File 5
- Supplementary File 6
- Supplementary File 7
- Supplementary File 8

Funding Statement

This work was supported by the National Research and Innovation Agency (BRIN) through the grants RIIM No.12/II.7/HK/2023 (granted to Astutiati Nurhasanah) and RP-ORK No.23/III.9/HK/2023 (granted to Nihayatul Karimah).

Conflict of Interest Disclosure

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

Ethics Approval Statement

The study protocol complies with the Declaration of Helsinki and has been approved by Health Research Ethics Committee of Faculty of Medicine, Universitas Hasanuddin (No. 678/UN4.6.4.5.3.1/PP36/2023, September 13, 2023).

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT (Free Online Version) in order to rephrase some sentences in parts of the articles, to improve readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Abraham, M., Alekseenko, A., Bergh, C., Blau, C., Briand, E., Doijade, M., Fleischmann, S., Gapsys, V., Garg, G., Gorelov, S., Gouaillardet, G., Gray, A., Irrgang, M. E., Jalalypour, F., Kutzner, C., Lemkul, J. A., Lundborg, M., Merz, P., Miletic, V., … Lindahl, E. (2023). GROMACS 2023.2 Manual. ***GROMACS Archive, Version 2023.2***, 1–848. https://doi.org/https://doi.org/10.5281/zenodo.8134388

Adasme, M. F., Linnemann, K. L., Bolz, S. N., Kaiser, F., Salentin, S., Haupt, V. J., & Schroeder, M. (2021). PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Res.*, *49*(W1), W530–W534. https://doi.org/10.1093/nar/gkab294

Al-Ouqaili, M. T. S., Ahmad, A., Jwair, N. A., & Al-Marzooq, F. (2025). Harnessing bacterial immunity: CRISPR-Cas system as a versatile tool in combating pathogens and revolutionizing medicine. *Front. Cell. Infect. Microbiol.*, *15*. https://doi.org/10.3389/fcimb.2025.1588446

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Arlinta, D. (2024, February 8). Estimated TB cases in Indonesia have increased to more than 1 million cases. *Kompas.id*. https://www.kompas.id/baca/english/2024/02/07/en-estimasi-kasus-tbc-di-indonesia-naik-jadi-lebih-dari-1-juta-kasus

Atavliyeva, S., Auganova, D., & Tarlykov, P. (2024). Genetic diversity, evolution and drug resistance of Mycobacterium tuberculosis lineage 2. *Front. Microbiol.*, *15*. https://doi.org/10.3389/fmicb.2024.1384791

Bui, H.-H., Sidney, J., Li, W., Fusseder, N., & Sette, A. (2007). Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics*, *8*(1), 361. https://doi.org/10.1186/1471-2105-8-361

Choi, H., Kim, W. S., Back, Y. W., Kim, H., Kwon, K. W., Kim, J., Shin, S. J., & Kim, H. (2015). *Mycobacterium tuberculosis* RpfE promotes simultaneous Th1- and Th17-type T-cell immunity via TLR4-dependent maturation of dendritic cells. *Eur. J. Immunol.*, *45*(7), 1957–1971. https://doi.org/10.1002/eji.201445329

Comas, I., Chakravartti, J., Small, P. M., Galagan, J., Niemann, S., Kremer, K., Ernst, J. D., & Gagneux, S. (2010). Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nat. Genet.*, *42*(6), 498–503. https://doi.org/10.1038/ng.590

Davies, A. P., Dhillon, A. P., Young, M., Henderson, B., McHugh, T. D., & Gillespie, S. H. (2008). Resuscitation-promoting factors are expressed in Mycobacterium tuberculosis-infected human tissue. *Tuberculosis*, *88*(5), 462–468. https://doi.org/10.1016/j.tube.2008.01.007

Downing, K. J., Mischenko, V. V., Shleeva, M. O., Young, D. I., Young, M., Kaprelyants, A. S., Apt, A. S., & Mizrahi, V. (2005). Mutants of *Mycobacterium tuberculosis* Lacking Three of the Five *rpf* -Like Genes Are Defective for Growth In Vivo and for Resuscitation In Vitro. *Infect. Immun.*, *73*(5), 3038–3043. https://doi.org/10.1128/IAI.73.5.3038-3043.2005

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*(1), 4. https://doi.org/10.1186/1471-2105-8-4

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., & Bairoch A. (2005). Protein Identification and Analysis Tools on the Expasy Server. In J. M. Walker (Ed.), ***The Proteomics Protocols Handbook*** (pp. 571–607). Humana Press .

Greenbaum, J., Sidney, J., Chung, J., Brander, C., Peters, B., & Sette, A. (2011). Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics*, *63*(6), 325–335. https://doi.org/10.1007/s00251-011-0513-0

Hall, T. (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, *41*, 95–98.

Hett, E. C., Chao, M. C., Deng, L. L., & Rubin, E. J. (2008). A Mycobacterial Enzyme Essential for Cell Division Synergizes with Resuscitation-Promoting Factor. *PLoS Pathog.*, *4*(2), e1000001. https://doi.org/10.1371/journal.ppat.1000001

Hett, E. C., Chao, M. C., Steyn, A. J., Fortune, S. M., Deng, L. L., & Rubin, E. J. (2007). A partner for the resuscitation-promoting factors of *Mycobacterium tuberculosis*. *Mol. Microb.*, *66*(3), 658–668. https://doi.org/10.1111/j.1365-2958.2007.05945.x

Honorato, R. V., Trellet, M. E., Jiménez-García, B., Schaarschmidt, J. J., Giulini, M., Reys, V., Koukos, P. I., Rodrigues, J. P. G. L. M., Karaca, E., van Zundert, G. C. P., Roel-Touris, J., van Noort, C. W., Jandová, Z., Melquiond, A. S. J., & Bonvin, A. M. J. J. (2024). The HADDOCK2.4 web server for integrative modeling of biomolecular complexes. *Nat. Protoc.*, *19*(11), 3219–3241. https://doi.org/10.1038/s41596-024-01011-0

Ioerger, T. R., Feng, Y., Ganesula, K., Chen, X., Dobos, K. M., Fortune, S., Jacobs, W. R., Mizrahi, V., Parish, T., Rubin, E., Sassetti, C., & Sacchettini, J. C. (2010). Variation among genome sequences of H37Rv strains of Mycobacterium tuberculosis from multiple laboratories. *J. Bacteriol.*, *192*(14), 3645–3653. https://doi.org/10.1128/JB.00166-10

Jalil, A. A. A., Al-khafaji, Z. M., & Al-ouqaili, M. T. (2018). INVESTIGATION OF ACCD3 GENE OF MYCOBACTERIUM TUBERCULOSIS IRAQI ISOLATES. *Asian J. Pharm. Clin. Res.*, *11*(8), 208. https://doi.org/10.22159/ajpcr.2018.v11i8.25269

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kana, B. D., Gordhan, B. G., Downing, K. J., Sung, N., Vostroktunova, G., Machowski, E. E., Tsenova, L., Young, M., Kaprelyants, A., Kaplan, G., & Mizrahi, V. (2008). The resuscitation-promoting factors of *Mycobacterium tuberculosis* are required for virulence and resuscitation from dormancy but are collectively dispensable for growth *in vitro*. *Molecular Microbiology*, *67*(3), 672–684. https://doi.org/10.1111/j.1365-2958.2007.06078.x

Kleinnijenhuis, J., Oosting, M., Joosten, L. A. B., Netea, M. G., & Van Crevel, R. (2011). Innate Immune Recognition of *Mycobacterium tuberculosis*. *Clin. Dev. Immunol.*, *2011*, 1–12. https://doi.org/10.1155/2011/405310

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nat. Protoc.*, *12*(2), 255–278. https://doi.org/10.1038/nprot.2016.169

Laskowski, R. A. (2022). PDBsum1: A standalone program for generating PDBsum analyses. *Prot. Sci.*, *31*(12), e4473. https://doi.org/10.1002/pro.4473

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, *26*(2), 283–291. https://doi.org/10.1107/S0021889892009944

Li, H., Guo, H., Chen, T., Yu, L., Chen, Y., Zhao, J., Yan, H., Chen, M., Sun, Q., Zhang, C., Zhou, L., & Chen, L. (2018). Genome-wide SNP and InDel mutations in Mycobacterium tuberculosis associated with rifampicin and isoniazid resistance. *Int. J. Clin. Exp. Pathol.*, *11*(8), 3903–3914.

Mavrici, D., Prigozhin, D. M., & Alber, T. (2014). *Mycobacterium tuberculosis* RpfE crystal structure reveals a positively charged catalytic cleft. *Prot. Sci.*, *23*(4), 481–487. https://doi.org/10.1002/pro.2431

Mukamolova, G. V., Turapov, O. A., Young, D. I., Kaprelyants, A. S., Kell, D. B., & Young, M. (2002). A family of autocrine growth factors in *Mycobacterium tuberculosis*. *Mol. Microb.*, *46*(3), 623–635. https://doi.org/10.1046/j.1365-2958.2002.03184.x

Osei-Wusu, S., Otchere, I. D., Morgan, P., Musah, A. B., Siam, I. M., Asandem, D., Afum, T., Asare, P., Asante-Poku, A., Kusi, K. A., Gagneux, S., & Yeboah-Manu, D. (2021). Genotypic and phenotypic diversity of Mycobacterium tuberculosis complex genotypes prevalent in West Africa. *PLOS ONE*, *16*(8), e0255433. https://doi.org/10.1371/journal.pone.0255433

Owaid, H. A., & Al-Ouqaili, M. T. S. (2024). Molecular and bacteriological investigations for the co-existence CRISPR/Cas system and β-lactamases of types extended-spectrum and carbapenemases in Multidrug, extensive drug and Pandrug-Resistant Klebsiella pneumoniae. *Saudi J. Biol. Sci.*, *31*(7), 104022. https://doi.org/10.1016/j.sjbs.2024.104022

Panda, S., Morgan, J., Cheng, C., Saito, M., Gilman, R. H., Ciobanu, N., Crudu, V., Catanzaro, D. G., Catanzaro, A., Rodwell, T., Perera, J. S. B., Chathuranga, T., Gunasena, B., DeSilva, A. D., Peters, B., Sette, A., & Lindestam Arlehamn, C. S. (2023). *Identification of differentially recognized T cell epitopes in the spectrum of* Mtb *infection*. In *bioRxiv*. https://doi.org/10.1101/2023.04.12.536550

Park, H.-S., Choi, S., Back, Y.-W., Lee, K.-I., Choi, H.-G., & Kim, H.-J. (2021). Mycobacterium tuberculosis RpfE-Induced Prostaglandin E2 in Dendritic Cells Induces Th1/Th17 Cell Differentiation. *Int. J. Mol. Sci.*, *22*(14), 7535. https://doi.org/10.3390/ijms22147535

Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.*, *48*(W1), W449–W454. https://doi.org/10.1093/nar/gkaa379

Romano, M., Squeglia, F., Kramarska, E., Barra, G., Choi, H.-G., Kim, H.-J., Ruggiero, A., & Berisio, R. (2023). A Structural View at Vaccine Development against M. tuberculosis. *Cells*, *12*(2), 317. https://doi.org/10.3390/cells12020317

Rosser, A., Stover, C., Pareek, M., & Mukamolova, G. V. (2017). Resuscitation-promoting factors are important determinants of the pathophysiology in *Mycobacterium tuberculosis* infection. *Crit. Rev. Microbiol.*, *43*(5), 621–630. https://doi.org/10.1080/1040841X.2017.1283485

Santoro, F., Guerrini, V., Lazzeri, E., Iannelli, F., & Pozzi, G. (2017). Genomic polymorphisms in a Laboratory Isolate of Mycobacterium tuberculosis Reference Strain H37Rv (ATCC27294). *New Microbiol.*, *40*(1), 62–69.

Squeglia, F., Romano, M., Ruggiero, A., Vitagliano, L., De Simone, A., & Berisio, R. (2013). Carbohydrate Recognition by RpfB from Mycobacterium tuberculosis Unveiled by Crystallographic and Molecular Dynamics Analyses. *Biophysic. J.*, *104*(11), 2530–2539. https://doi.org/10.1016/j.bpj.2013.04.040

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Tantivitayakul, P., Juthayothin, T., Ruangchai, W., Smittipat, N., Disratthakit, A., Mahasirimongkol, S., Tokunaga, K., & Palittapongarnpim, P. (2020). Identification and in silico functional prediction of lineage-specific SNPs distributed in DosR-related proteins and resuscitation-promoting factor proteins of Mycobacterium tuberculosis. *Heliyon*, *6*(12), e05744. https://doi.org/10.1016/j.heliyon.2020.e05744

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.*, *40*(7), 1023–1025. https://doi.org/10.1038/s41587-021-01156-3

Weiskopf, D., Angelo, M. A., De Azeredo, E. L., Sidney, J., Greenbaum, J. A., Fernando, A. N., Broadwater, A., Kolla, R. V., De Silva, A. D., De Silva, A. M., Mattia, K. A., Doranz, B. J., Grey, H. M., Shresta, S., Peters, B., & Sette, A. (2013). Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *Proc. Natl. Acad. Sci. U. S. A.*, *110*(22). https://doi.org/10.1073/pnas.1305227110

WHO. (2024). ***Global tuberculosis report 2024***. https://www.who.int/publications/i/item/9789240101531

Yang, J., & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.*, *43*(W1), W174–W181. https://doi.org/10.1093/nar/gkv342.