

Prediction of B-cell epitope by *in silico* analysis of *Mycobacterium tuberculosis* Ag85B antigen

Nihayatul Karimah, Sabar Pambudi*

Center for Pharmaceutical & Medical Technology, BPPT, Gedung 611 LAPTLAB, PUSPIPTEK, Serpong, Tangerang Selatan, Banten, Indonesia

Received 21st October 2019 / Accepted 12th February 2020

Abstract. *Mycobacterium tuberculosis* (*Mtb*) is a causative pathogen of tuberculosis (TB) that emerges as one of the deadliest communicable diseases in Indonesia. The quest for protein biomarkers for TB has been conducted in order to develop a TB diagnostic kit and a TB vaccine. One of the abundant biomarkers in the TB infected human serum is the Ag85B antigen. In this study, we employed immunoinformatic prediction tools such as Ellipro and VaxiJen to predict the B-cell epitopes of Ag85B wildtype and multidrug resistance type (mutant). We then performed molecular docking simulation to evaluate the predicted epitopes using HADDOCK. The screening of both continuous and discontinuous B-cell epitopes using criteria-based analysis resulted in the eight linear epitopes and two conformational epitopes in Ag85B with high antigenicity. The *in silico* analysis showed no major differences between Ag85B wildtype and Ag85B mutant, implying Ag85B a good target for TB vaccine candidates but not for a specific biomarker that differentiates wild-type and mutant TB.

Keywords: tuberculosis, Ag85B, B-cell epitope, *in silico*, mutant

INTRODUCTION

Tuberculosis (TB) is a significant public health threat worldwide, causing high morbidity and mortality (Zumla *et al.*, 2013). Detecting active TB cases is challenging and mostly relying predominantly on sputum analysis, including acid-fast *Bacillus* (AFB) smear microscopy (Ryu, 2015) and *Mycobacterium tuberculosis* (*Mtb*) culture, both of which have only moderate sensitivity and specificity and long turnaround time (Sharma *et al.*, 2012).

The emergence of multidrug-resistance tuberculosis (MDR-TB) represents a growing threat to public health and economic growth worldwide. The WHO has reported that the MDR-TB cases increased by more than 20% annually. The increasing TB cases, if not all, could be due to the improvements in rapid diagnostics

and phenotypic drug susceptibility testing (Lange *et al.*, 2018).

Ag85 is a virulence factor in TB, which appears as a complex to bind with fibronectin (Fn) protein in order to assist the attachment and invasion of *Mtb* into human macrophages (Kuo *et al.*, 2013). It is a mycolyltransferase that comprises of three proteins: Ag85A, Ag85B, and Ag85C. The first two proteins are the predominant component of the complex for about 60% of total culture fluid proteins from *Mtb* H37Rv to which the concentration of Ag85B higher than Ag85A (Wang *et al.*, 2012). The Ag85B can be detected in the culture fluid protein after three days (Zhang *et al.*, 2018). Therefore, focusing on Ag85B as the target biomarker is beneficial based on its quantity and its activity.

*Authors for correspondence: Sabar Pambudi, Center for Pharmaceutical & Medical Technology, BPPT, Gedung 611 LAPTLAB, PUSPIPTEK, Serpong, Tangerang Selatan, Banten, Indonesia. Email – sabra.pambudi@bppt.go.id

Previous study reported that specific mutation of Ag85B in BCG strain may be used to differentiate *Mycobacterium bovis* and BCG strains from *Mtb* strains (Jiang *et al.*, 2015). However, to the best of our knowledge, no information regarding the difference of B-cell epitopes of Ag85B wild type and mutant has been reported before. In this study, we predicted and evaluated the differences between B-cell epitopes of Ag85B from wildtype and mutant strain as well as their antigenicity using immunoinformatics tools. The result of this study will be beneficial in designing a new strategy of the developing detection and prevention methods of TB infection.

MATERIALS AND METHODS

Protein structure preparation

The sequences of Ag85B of *Mtb* for wild-type and mutant were retrieved from UniProt with the primary accession code P9WQP1 and Q847N4. We performed structure prediction using I-Tasser, which adopt multiple threading approaches to produce the full-length protein structure (Yang *et al.*, 2015). The BLAST search against the Protein Data Bank using the MRS web-server was used to find the correct template for structure prediction (Joosten *et al.*, 2010). The YASARA energy minimization was employed to fix the bad stereochemistry (Krieger *et al.*, 2009) of the structural model for wildtype and mutant Ag85B. On the other hand, the structure of the variable region of the heavy chain (VH) and light chain (VL) of ScFv α -Ag85B (KR101631054B1., 2019) was built using Prediction of Immunoglobulin Structure (PIGSpro) with the option of the same canonical structure to find the right template (Lepore *et al.*, 2017). The PDB code of 1W72 and 5MES were the chosen template for VH and VL, respectively. Energy minimization with YASARA was used once more to obtain the better structure of the ScFv anti-Ag85B antibody. We also identified the antibody binding region (ABR) using an automated tool called Paratome for both VH and VL using the modeled structure of ScFv anti-Ag85B (Kunik *et al.*, 2012). These regions are useful to annotate the antibody on the docking simulation, in which specific residues are

predicted to be responsible for antigen binding.

Identification of B-cell epitope

The main web tool, Ellipro, was used to predict the linear and discontinuous epitope of wildtype and mutant Ag85B based on 3D protein structure (Baloi, 2016). VaxiJen v.2 server and IEDB B-cell epitope prediction tools were employed in this study to confirm the immunogenicity of the predicted epitopes. The final score from the IEDB B-cell epitope prediction tool was determined by the average of score accumulation of Emini surface accessibility, Karplus–Schulz flexibility, Kolaskar–Tongaonkar antigenicity, and Parker hydrophilicity for each window (Emini *et al.*, 1985; Karplus, 1985; Kolaskar and Tongaonkar, 1990; Parker *et al.*, 1986). This score, along with epitope probability from BepiPred-2.0 (Jespersen *et al.*, 2017) and the presenting score from VaxiJen, was taken into account to predict the linear B-cell epitope of the wild-type and the mutant Ag85B *Mtb*.

Molecular docking simulation

The docking simulation to evaluate the correct binding and orientation between the antigen and antibody in each epitope was carried-out by using HADDOCK (van Zundert *et al.*, 2016). HADDOCK considers the importance of flexibility for which it allows conformational changes upon protein-protein binding. The high accuracy prediction is the docking model with i-RMSD or l-RMSD below 1 Å. The one with the i-RMSD below 2 Å and/or l-RMSD below 5 Å is a medium quality prediction, while the other with i-RMSD below 4 Å or l-RMSD below 10 Å are the acceptable prediction. We submitted the structure of the ScFv anti-Ag85B antibody as the first molecule and the structure of Ag85B as the second molecule. In consideration of the simulation restraint, the antibody binding region (ABR) in VH and VL was defined as the active residues of the ScFv anti-Ag85B antibody while the epitopes derived from Ellipro were designated as the active residues of Ag85B. We let HADDOCK define the passive residues automatically in 6.5 Å distance around the active residues. The parameters to be considered are the HADDOCK score, the RMSD from the overall lowest energy structure that represents i-RMSD and/or l-RMSD, and the z-score. After having the

complex Ag-Ab, we calculated the hydrogen bond energy between the two molecules as well as listed the residues of Ag85B, making H-bond using YASARA Structure (Krieger and Vriend, 2014).

RESULTS AND DISCUSSION

The wildtype and the mutant Ag85B of *Mycobacterium tuberculosis* have 325 amino acids. The difference between the two is that the mutant undergoes amino acid substitution of D218N and E270K in which the negatively charged aspartic acid 218 is replaced by asparagine and the negatively charged glutamic acid 270 is replaced by a positively charged lysine (Figure 1). The substitution into oppositely charged amino acids usually produces relatively major changes in the overall structure and function, especially, if the amino acid has a big size such as glutamic acid and lysine. Knowledge of such changes having implications for the binding of ScFv anti-Ag85B, as well as the resulting epitopes, is one of the objectives, that we want to highlight.

We built the 3D structure, using I-TASSER, that showed a slight difference between the two sequences as well as between the modeled structure and the template or reference structure represented by the root mean square deviation (RMSD) (Figure 2). The backbone RMSDs between the modeled structure and the reference

structure are lower than 2 Å, indicating that they have minor differences. The reference structure with PDB code 1F0N is a mature form of mycolyl-transferase Ag85B that does not have the signal peptide (residue 1-41) while the models are the full-length structure. The effect of threading using I-TASSER gave a relatively different result for the first 41 residues between the wildtype and the mutant, while the differences for the further residues are not substantial. The quality of the structure used in this study can be seen in the Ramachandran plot in Figure 3.

The structure assessment using MolProbity (Chen *et al.*, 2010) showed an excellent structural quality that the three structures have Ramachandran favored > 90% and MolProbity score < 2.0. A general structural quality requires chains to have a MolProbity score below 2.0. Therefore, all the modeled structures are eligible to be used for B-cell epitope mapping and biomolecular docking. Using the structure of ScFv anti-Ag85B, Paratome has identified the antigen-binding regions (ABRs): 134F^TFD^DFAMH142, 159T^WNSG^TIAY167, and 205R^GHYGLDV211 for the heavy chain; 25Y^SNI^GTNYVY34, 46L^VIQ^NTQRPS56, and 91W^DDSLS96 for the light chain. All these regions are predicted to have direct interaction with Ag85B. ABR is calculated based on the virtually antigen-binding residues that fall within regions of structural consensus.

Ag85B_MTb_wildt..	1	MTDVSRKIRAWGRRRLMIGTAAAVVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSPSMGR	60
Ag85B_MTb_mutant	1	MTDVSRKIRAWGRRRLMIGTAAAVVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSPSMGR	60
Ag85B_MTb_wildt..	61	DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYYQSGLSIVMPVGGQSSFYS	120
Ag85B_MTb_mutant	61	DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYYQSGLSIVMPVGGQSSFYS	120
Ag85B_MTb_wildt..	121	DWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTGSAAIGLSMAGSSAMI LAAYHP	180
Ag85B_MTb_mutant	121	DWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTGSAAIGLSMAGSSAMI LAAYHP	180
Ag85B_MTb_wildt..	181	QQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDP ^T QQI PKL	240
Ag85B_MTb_mutant	181	QQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAANMWGPSSDPAWERNDP ^T QQI PKL	240
Ag85B_MTb_wildt..	241	VANNTRLWVYCGNGTPNELGGANIPAEFLKNEFVRSSNLKFDAYNAAGGHNAVFNFPPNG	300
Ag85B_MTb_mutant	241	VANNTRLWVYCGNGTPNELGGANIPAEFLKNEFVRSSNLKFDAYNAAGGHNAVFNFPPNG	300
Ag85B_MTb_wildt..	301	THSWEYWGAQLNAMKGD ^L QSSLGAG	325
Ag85B_MTb_mutant	301	THSWEYWGAQLNAMKGD ^L QSSLGAG	325

Figure 1. The aligned residues between Ag85B wildtype (P9WQP1) and Ag85B mutant (Q847N4). Mutations occurred in residue D218N and residue E270K.

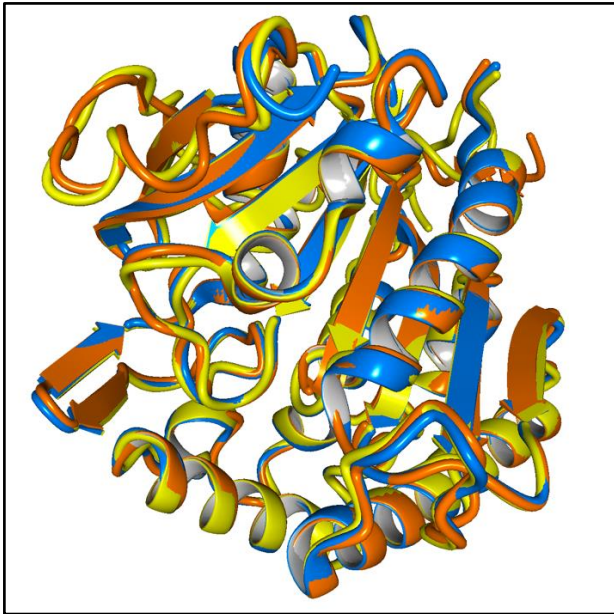


Figure 2. The aligned objects between the modeled Ag85B wildtype (yellow) and the reference structure with PDB ID 1F0N (blue) have an RMSD of 0.506 Å over 282 aligned residues with 100.00% sequence identity. The modeled Ag85B mutant (orange) and the reference structure 1F0N (blue) has an RMSD of 0.466 Å over 282 aligned residues with 99.29% sequence identity.

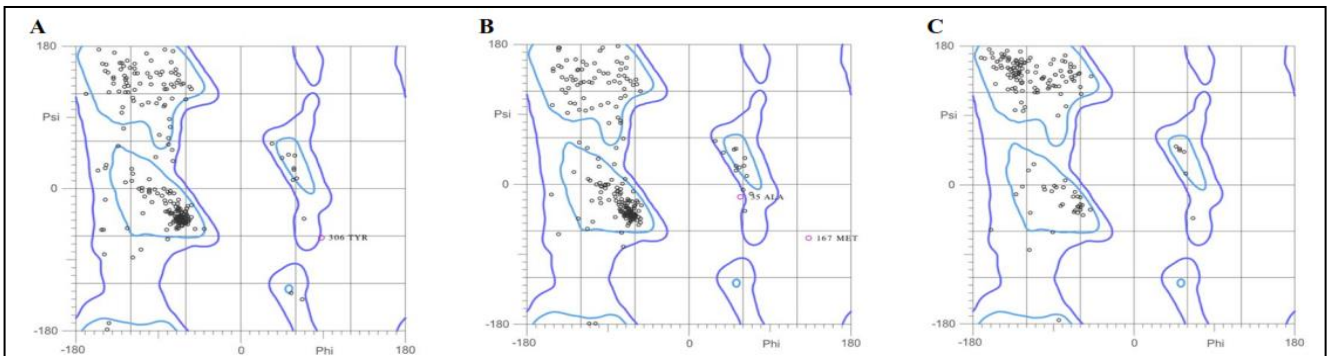


Figure 3. The Ramachandran plot of the modeled structures. (A) Ag85B wildtype with 90.4% of all residues was in favored regions, and 99.1% of all residues were in allowed regions. (B) Ag85B mutant with 91.0% of all residues were in favored regions and 98.5% of all residues were in allowed regions. (C) ScFv anti-Ag85B antibody 98.2% of all residues were in favored regions and 100.0% of all residues were in allowed regions.

The epitope prediction using Ellipro produced nine linear epitopes of Ag85B wildtype and ten linear epitopes of Ag85B mutant with the minimal six residues in each epitope (Table 1). Ellipro also predicted three conformational epitopes for both types of Ag85B (Table 2). Each of the epitopes has a protrusion index (PI) of more than 0.5. The protrusion index indicates the percentage of the molecular bulk where the atoms of the epitopes are protruding outside the region (Lata *et al.*, 2018). The higher the protrusion index, the higher the probability of the peptide to be found in experimentally determined continuous epitopes. The wildtype Ag85B and the mutant Ag85B did not show a major difference in terms of the predicted epitopes by Ellipro. The mutant Ag85B has one added epitope, 52QVPSPMGRD61,

which does not appear in Ag85B wildtype. The wildtype Ag85B and the mutant Ag85B only have one corresponding conformational epitopes, although the residues incorporated in all the conformational epitopes are almost identical. The highest protrusion index is 0.811 in Ag85B wildtype and 0.818 in Ag85B mutant for linear epitope 251CGNGTTPNELGGANIPAEF268. For the conformational epitope, the highest protrusion index is 0.713 in wildtype Ag85B and 0.708 in mutant Ag85B. The average protrusion index for all the predicted epitopes by Ellipro was also calculated that Ag85B wildtype has an average PI score of 0.677, while the Ag85B mutant has an average PI score of 0.669. This data indicates that the differences between Ag85B wildtype and Ag85B mutant are not substantial in

the case of linear epitope position based on their 3D structure.

For further analysis, the antigenicity for each resulting linear epitope from Ellipro was calculated using Vaxijen 2.0 and IEDB linear epitope prediction tools. The Vaxijen threshold was set to 0.5 that the epitopes scored above the threshold were predicted to be antigenic (Doytchinova and Flower, 2007). Three epitopes in both Ag85B showed Vaxijen score below the threshold and predicted to be nonantigenic.

Those three epitopes consist of the first 61 residues for the wildtype and the first 49 residues for the mutant, 296FPPNGTH302, and 99EWYYQSG105. For the reason that the first 41 residues are a signal peptide which undergoes cleavage in the mature form or after Ag85B is secreted, it was eliminated from epitope prediction. Moreover, Vaxijen and IEDB linear epitope prediction tools calculated that the predicted epitope containing this peptide has a low score, thus predicted to be nonantigenic.

Table 1. Predicted linear B-cell epitopes of Ag85B *Mycobacterium tuberculosis*.

	Peptide	PI by Ellipro	Vaxijen Score	Average score from IEDB linear epitope predictions	HADDOCK score	HADDOCK RMSD from the overall lowest-energy structure	HADDOCK Z-Score	Total H-bond energy (kJ/mol)	Residues of Ag85B with H-bond (*)
Wildtype									
WTL1	²⁵¹ CGNGTNPNEGGANIPAEF ₂₆₈	0.811	1.2546	1.276	-133.9 +/- 1.8	0.5 +/- 0.3	-2.3	292.05	W100, N257, E258, E267, A262, G261, N299, E305, A309, Q310
WTL2	²⁷⁸ LKFQDAYNAAGGHN ₂₉₁	0.729	1.5598	1.474	-157.6 +/- 0.8	0.6 +/- 0.4	-1.7	291.93	G198, E229, N253, R274, L278, K279, Q281, D282, N285, G288, G289, H290, F296
WTL3	¹ MTDVSRRKIRAWGRRRLMI GTAAAVVLPGLVGLAGGAATAGA FSRPLPVEYLQVPSMGRD ₆₁	0.722	0.1406	0.820	-110.5 +/- 7.6	0.9 +/- 0.5	-1.4	197.85	A21, L28, V29, G59, Q52, D61, Q147, W148, N152
WTL4	¹²² WYPACGKAGCQTY ₁₃₅	0.705	1.5224	1.426	-111.7 +/- 7.0	2.9 +/- 1.4	-1.1	255.85	G59, W123, S124, G128, K136, A217, D218, D225
WTL5	²⁰⁹ GDAGGYKAADMWGPSSDP ₂₂₉	0.66	0.7689	1.557	-103.5 +/- 4.2	5.3 +/- 0.8	0.1	242.05	S196, G198, S202, K215, A217, M219, S223, S224, W228, E229, Q235
WTL6	²⁹⁶ FPPNGTH ₃₀₂	0.646	0.1480	1.419	-125.0 +/- 10.1	0.5 +/- 0.3	-1.4	288.20	E258, L259, A262, E267, N299, E305, Q310, N312
WTL7	⁶⁸ SGGNNS ₇₃	0.625	2.8364	2.037	-103.4 +/- 20.0	0.9 +/- 0.6	-0.4	168.68	N71, Y102, K315, Q319, G323, A324
WTL8	⁹⁹ EWYYQSG ₁₀₅	0.597	0.3454	0.875	-162.8 +/- 8.0	0.7 +/- 0.5	0.0	309.30	S42, R43, S68, N71, E99, Y102, Q103, K315, G325
WTL9	¹⁹⁵ PSQMGPS ₂₀₂	0.596	1.1848	1.257	-173.0 +/- 2.9	0.6 +/- 0.3	-1.0	205.00	D194, M199, S202, G205, G221, W228, Q235, K279, D282
Mutant									
MTL1	²⁵¹ CGNGTNPNEGGANIPAEF ₂₆₈	0.818	1.2546	1.271	-121.6 +/- 1.7	5.5 +/- 0.6	-1.2	221.18	R83, M199, D210, N253, G254, N263, E267, N271, E305
MTL2	²⁷⁸ LKFQDAYNAAGGHN ₂₉₁	0.702	1.5598	1.474	-111.6 +/- 1.9	3.9 +/- 0.1	1.4	267.07	D194, Q197, E229, T234, S275, K279, D282
MTL3	¹ MTDVSRRKIRAWGRRRLMIG TAAAVVLPGLVGLAGGAAT AGAFSRPLPVE ₄₉	0.773	0.1413	0.785	-135.9 +/- 14.3	1.0 +/- 0.7	-1.5	321.07	R13, R14, G27, L28, L31, G33, G34, A35, V48, E49, Y50, Q147, S150, R153
MTL4	¹²² WYPACGKAGCQTY ₁₃₅	0.69	1.5224	1.426	-132.2 +/- 11.3	0.6 +/- 0.4	-1.8	223.97	D121, Y123, S124, G128, C132, K215, W220, S224, D225
MTL5	²⁰⁹ GDAGGYKAANWGPSSDP ₂₂₆	0.688	0.8104	1.426	-107.3 +/- 1.5	2.0 +/- 0.7	-0.6	187.98	S124, P125, G128, K129, S202, Y214, K215, N218, W220
MTL6	²⁹⁶ FPPNGTH ₃₀₂	0.655	0.1480	1.419	-115.7 +/- 14.5	1.9 +/- 1.2	-1.3	210.72	N253, G254, E258, E267, R274, N295, F296, N299, Q310
MTL7	⁶⁸ SGGNNS ₇₃	0.625	2.8364	2.036	-124.4 +/- 8.1	0.9 +/- 0.7	-1.1	236.78	R6, R13, G45, L46, S68, G69, N71, N72, Y102, G105, S150, N152
MTL8	⁹⁹ EWYYQSG ₁₀₅	0.608	0.3454	0.874	-135.8 +/- 3.8	0.8 +/- 0.5	-1.9	209.05	P44, N94, E99, Y102, E258, N312, K315, G316, S320
MTL9	¹⁹⁵ PSQMGPSLJGL ₂₀₆	0.586	0.9259	0.871	-135.1 +/- 6.8	0.7 +/- 0.5	-1.8	259.52	S196, P201, K215, S224, E267, N271, K279, D282
MTL10	⁵² QVPSMGRD ₆₁	0.543	1.1145	1.345	-117.3 +/- 2.0	4.9 +/- 0.2	-1.3	229.90	G33, Q52, S55, G59, R60, D61, Q147, W148, N152

(*)Residues in bold are part of the predicted linear B-cell epitopes.

Table 2. Predicted conformational B-cell epitopes of Ag85B *Mycobacterium tuberculosis*.

	Residues (*)	Number of residues	PI by Ellipro	HADDOCK K score	HADDOCK RMSD from the overall lowest-energy structure	HADDOCK Z-Score	Total H-bond energy (kJ/mol)	Residues of Ag85B with H-bond (**)
Wildtype								
WTC1	P238, V241, A242, N243, N244, T245, C251, G252, N253, G254, T255, P256, N257, E258, L259, G260, G261, A262, N263, I264, P265, A266, E267, N271, R274, S275, L278, Q281, D282, A283, Y284, N285, A286, A287, G288, G289, H290, N291, F296, P297, P298, N299, G300, T301, H302, S303, E305, Y306	48	0.713	-134.1 +/- 8.2	0.8 +/- 0.5	-1.8	349.88	N257, R274, Q281, D282, Y284, N285, G288, G289, H290, A292, N299, E305, A309, N312, D317
WTC2	M1, T2, D3, V4, S5, R6, K7, I8, A10, W11, G12, R13, R14, L15, M16, G18, T19, A20, A21, A22, V23, V24, L25, P26, G27, L28, V29, G30, L31, A32, G33, G34, A35, A36, T37, A38, G39, A40, F41, S42, R43, P44, G45, L46, P47, V48, E49, Y50, Q52, P54, P56, S57, D61, S68, G69, G70, N71, N72, S73, Q85, D86, D87, Y88, T95, P96, E99, W100, Y102, Q103, S104, G105, S143, Q147, W148, S150, A151, N152, R153, A154, V155, K156, T158, Q181, G316, Q319, S320, L322, G323, A324, G325	90	0.669	-92.5 +/- 15.0	0.7 +/- 0.4	-1.2	295.25	R6, L15, G34, L46, S68, N71, N72, Y102, Q103, Q147, R153, K156, T158
WTC3	M58, Q115, W122, S124, P125, A126, C127, G128, K129, A130, G131, C132, Q133, T134, Y135, K136, D194, P195, S196, Q197, G198, M199, P201, S202, G205, L206, G209, D210, A211, G212, G213, Y214, K215, A216, A217, D218, M219, W220, G221, P222, S223, S224, D225, P226, W228, E229, D232, T234, Q235, Q236	50	0.624	-149.6 +/- 7.6	0.5 +/- 0.3	-1.8	228.15	D194, S196, M199, S202, Y214, S223, W228, F272, S275, K279
Mutant								
MTC1	P238, V241, A242, N243, N244, T245, C251, G252, N253, G254, T255, P256, N257, E258, L259, G260, G261, A262, N263, I264, P265, A266, E267, F268, N271, F272, R274, S275, L278, K279, Q281, D282, A283, Y284, N285, A286, A287, G288, G289, H290, N291, F296, P297, P298, N299, G300, T301, H302, S303, E305, Y306	51	0.708	-130.5 +/- 15.7	0.8 +/- 0.5	-1.9	367.40	R246, N253, G254, T255, N257, R274, L278, Q281, D282, Y284, N285, F294, N295, F296, P298, A313
MTC2	M1, T2, D3, V4, S5, R6, K7, I8, R9, A10, W11, G12, R13, R14, L15, I17, G18, T19, A20, A21, A22, V23, V24, L25, P26, G27, L28, V29, G30, L31, A32, G33, G34, A35, A36, T37, A38, G39, A40, F41, S42, R43, P44, G45, L46, P47, V48, E49, Q52, P54, S55, P56, S57, M58, G59, R60, D61, S68, G69, G70, N71, N72, S73, Q85, D86, D87, Y88, Y101, S104, G105, L106, W122, S124, P125, A126, C127, G128, K129, A130, G131, C132, Q133, T134, Y135, K136, S143, W148, S150, A151, N152, R153, A154, V155, K156, Q181, D194, P195, S196, Q197, G198, M199, P201, S202, L203, G205, L206, G209, D210, A211, G212, G213, Y214, K215, A216, A217, N218, M219, G221, P222, S223, S224, D225, P226, E229, D232, T234, Q235, G316, Q319, S320, L322, G323, A324, G325	134	0.662	35.7 +/- 15.3	0.6 +/- 0.4	-0.9	362.32	K7, G18, T19, A22, V23, D121, T135, K136, S224, E229, Q235, K239
MTC3	P96, E99, W100, N312	4	0.583	-149.9 +/- 1.8	0.6 +/- 0.4	-1.7	228.40	R43, G45, N94, E99, W100, Y102, E258, N312, K315, Q319

(*)Residues in bold are part of the predicted linear B-cell epitopes.

(**)Residues in bold are part of the predicted conformational B-cell epitopes.

The next antigenicity calculation was carried out using IEDB linear epitope prediction (Emini surface accessibility prediction, Karplus–Schulz flexibility prediction, Kolaskar–Tongaonkar antigenicity, and Parker hydrophilicity prediction). The single score to predict a peptide antigenicity comes from the average of the score from the four methods. A higher score increases the probability of the protein to be antigenic. Again, the predicted epitope containing signal peptide has the lowest score, whereas the predicted epitope 68SGGNS73 is calculated to be the most antigenic, which corresponds with the VaxiJen result. Most of the result from IEDB B-cell linear epitope prediction corresponds with the result from VaxiJen, except for 296FPPNGTH302, which showed the opposite. Hence, it will be ambiguous to include this peptide as the predicted linear epitope because it has a low VaxiJen score. The average scores were calculated from both antigenicity prediction tools to discern the antigenicity difference between Ag85B wildtype and mutant. The average score of antigenicity denotes almost the same result between Ag85B wildtype (1.08 for VaxiJen and

1.35 for IEDB linear epitope prediction) and Ag85B mutant (1.07 for VaxiJen and 1.29 for IEDB linear epitope prediction). The similarity between the two is also exhibited on the similar pattern of antigenic propensity.

The predicted conformational epitope data from Ellipro was used to resolve the ambiguity of whether 296FPPNGTH302 is one of the predicted linear epitopes or not. It is found that 296FPPNGTH302 is listed on the top list of predicted conformational epitopes with a protrusion index 0.713 for Ag85B wildtype and 0.708 for Ag85B mutant. Considering this data, 296FPPNGTH302 is included as one of the predicted linear epitopes. In contrast to 99EWYYQSG105, it has a low score of protrusion index (calculated by Ellipro) and antigenic propensity (calculated by VaxiJen and IEDB B-cell linear epitope prediction tools). However, some residues of the peptide are listed on the predicted conformational epitope with protrusion index 0.669 for Ag85B wildtype and 0.662 for the Ag85B mutant. Having these average scores, 99EWYYQSG105 is delisted on the predicted linear epitope. The predicted

conformational epitope data were used to confirm whether the predicted linear epitopes become part of them. Both data were found to be closely related to each other, showing the appearance of most of the predicted linear epitopes on the predicted conformational epitopes. On the contrary, five out of seven (in Ag85B wildtype) and five out of eight (in Ag85B mutant) linear epitopes predicted by Bepipred 2.0 were observed in the linear and conformational epitopes predicted by Ellipro. This data reaffirms the antigenicity of some epitopes predicted by Ellipro.

The docking simulation was carried out in order to analyze the binding interaction between antigen and antibody. The value of i-RMSD or l-RMSD is equal to the value of HADDOCK RMSD from the overall lowest-energy structure. The brief results of the docking simulation are shown in Table 1 and Table 2. The docking simulation between Ag85B and ScFv anti-Ag85B using HADDOCK possess high accuracy with RMSD below 1 Å for 17 out of the 25 complexes. Six complexes have the medium quality prediction (RMSD < 5 Å), and two complexes have acceptable prediction (RMSD < 10 Å). These imply HADDOCK capability to find the correct position between Ag85B epitopes and the ABR of the ScFv anti-Ag85B antibody. The other parameter in the docking is the Z-score, which showed the number of standard deviations from the average of the scores that the cluster is located (Chen *et al.*, 2010). The more negative the Z-score, the better the cluster. Most of the result has negative Z-score of which the best clusters were successfully produced. The HADDOCK scoring function is a linear combination of van der Waals intermolecular energy, electrostatic intermolecular energy, desolvation energy, restraint violation energy, and buried surface area (Vangone *et al.*, 2017). HADDOCK cannot estimate the binding affinity of a complex but we used it to determine the possibility of the binding interaction between antigen in its epitope and antibody in its ABR in terms of the HADDOCK scoring function (Payandeh *et al.*, 2018). The lower the HADDOCK score, the higher the possibility of the pair to interact or make a complex. Most of the epitopes show HADDOCK score lower than -100.0, which indicates the high possibility of the molecules to interact with each other through the

given predicted epitopes and ABR (Xue *et al.*, 2016; de Vries *et al.*, 2010). Two conformational epitopes containing signal peptide scored higher than -100.0 with the worst score of 35.7 +/- 15.3 for the second conformational epitope of the Ag85B mutant. The positive score indicates that the ScFv anti-Ag85B antibody has relatively lousy contact to the given conformational epitope of Ag85B compared to the others. The restraint violation energy was recorded very high for the second conformational epitope, > 1000.0 kcal/mol, and > 2000.0 kcal/mol for Ag85B wildtype and Ag85B mutant, respectively. It occurred because the number of restraint residues calculated as active residues is relatively large, which makes the calculation was not satisfying. While the second conformational epitopes involve the signal peptide, and they were not taken into consideration.

The hydrogen bond energy calculated by the YASARA structure is inversely correlated to the HADDOCK score. The lower the HADDOCK score, the higher the H-bond energy. The higher the H-bond energy, the stronger the antigen binds to the antibody. One thing to note is that the docking result does not correspond to the protrusion index or antigenicity data. The docking only supports to give further annotation about the complex of antigen and antibody in a given predicted epitope and ABR, which then assists in predicting the correct conclusion in epitope mapping. We accordingly cannot eliminate 68SGGNNS73 in Ag85B wildtype from its possibility to be a predicted linear epitope for the reason that its HADDOCK score is lower than the others. 68SGGNNS73 in the Ag85B mutant has a higher HADDOCK score as well as its H-bond energy.

Moreover, the H-bond involves many residues of 68SGGNNS73 in Ag85B mutants such as S68, G69, N71, and N72. Only 68SGGNNS73 in Ag85B mutant (other than predicted epitopes containing signal peptide) was recorded to have the signal peptide making H-bond with the antibody while 68SGGNNS73 in Ag85B wildtype does not involve the signal peptide. In consequence, 68SGGNNS73 in Ag85B wildtype produced lower H-bond energy and a higher HADDOCK score. The main difference between 68SGGNNS73 data in Ag85B wildtype and Ag85B mutant is that they came

from different types of clusters. Therefore, the HADDOCK score, H-bond energy, and the residues involved in the H-bond are also different. The same thing applies to the other linear epitopes, e.g., 278LKFQDAYNAAGGHN₂₉₁, 251CGNGTPNELGGANIPAEF₂₆₈, 209GDAGGYKAADMWGPSSDPAWE₂₂₉, etc. This variability of the docking result is also the reason to separate it from antigenicity data. In

our analysis, most of the residues that make H-bond are probably important, especially if they are part of the linear and conformational epitopes. All the B-cell epitope candidates of Ag85B *Mycobacterium tuberculosis* are shown in Table 3. Overall, we can conclude that the differences between Ag85B wildtype and Ag85B mutant in terms of the predicted B-cell epitopes are not substantial.

Table 3. The candidates of the B-cell epitopes in Ag85B *Mycobacterium tuberculosis*.

Linear B-cell Epitope	Conformational B-cell Epitope
²⁵¹ CGNGTPNELGGANIPAEF ₂₆₈	P238, V241, A242, N243, N244, T245, C251, G252, N253, G254, T255, P256, N257, E258, L259, G260, G261, A262, N263, I264, P265, A266, E267, N271, R274, S275, L278, Q281, D282, A283, Y284, N285, A286, A287, G288, G289, H290, N291, F296, P297, P298, N299, G300, T301, H302, S303, E305, Y306
²⁷⁸ LKFQDAYNAAGGHN ₂₉₁	
¹²² WYSPACGKAGCQTY ₁₃₅	
²⁰⁹ GDAGGYKAANMWGPSSDP ₂₂₆ (or ²⁰⁹ GDAGGYKAADMWGPSSDPAWE ₂₂₉)	M58, Q115, W122, S124, P125, A126, C127, G128, K129, A130, G131, C132, Q133, T134, Y135, K136, D194, P195, S196, Q197, G198, M199, P201, S202, G205, L206, G209, D210, A211, G212, G213, Y214, K215, A216, A217, D218, M219, W220, G221, P222, S223, S224, D225, P226, W228, E229, D232, T234, Q235, Q236
²⁹⁶ FPPNGTH ₃₀₂	
⁶⁸ SGGNNS ₇₃	
¹⁹⁵ PSQGMGPSLIGL ₂₀₆	
⁵² QVPSPSMGRD ₆₁	

CONCLUSION

The B-cell epitope prediction of Ag85B *Mycobacterium tuberculosis* has been conducted using immunoinformatic tools. Further assessment of the antigenicity and the docking showed corresponding results to support the B-cell epitope prediction. The screening of both linear and conformational B-cell epitopes using criteria-based analysis resulted in the eight linear epitopes and two conformational epitopes in Ag85B with high antigenicity. There are no major differences between Ag85B wildtype and Ag85B mutant, which make this antigen the right candidate for the TB vaccine target. However, this biomarker cannot be used as the target to differentiate wildtype and mutant TB in a detection system by using a rapid test platform that employs a specific antibody against the antigen.

ACKNOWLEDGEMENTS

This work was supported by the INSINAS Pratama 2019 grant (018_6_IP_RISTEK_2019_34) from The Ministry of Research, Technology & Higher Education (RISTEKDIKTI), Indonesia. Nihayatul Karimah and Sabar Pambudi equally contributed to this work.

REFERENCES

- Baloi, B. 2016. In silico epitope prediction and 3D model analysis of Peste des petits ruminants virus nucleoprotein. *bioRxiv*: 095505.
- Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., & Richardson, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* 66(Pt 1): 12-21.
- de Vries, S. J., van Dijk, M., & Bonvin, A. M. J. J. 2010. The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols* 5(5): 883-897.

- Doytchinova, I. A. & Flower, D. R. 2007. Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 8(4): 1-7.
- Emini, E. A., Hughes, J. V., Perlow, D. S., & Boger, J. 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of Virology* 55(3): 836-9.
- Jespersen, M. C., Peters, B., Nielsen, M., & Marcatili, P. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research* 45(W1): w24-w29.
- Jiang, Y., Liu, H., Li, M., Li, G., Pang, H., Dou, X., Zhao, X., & Wan, K. 2015. Single nucleotide polymorphism in Ag85 genes of *Mycobacterium tuberculosis* complex: analysis of 178 clinical isolates from China and 13 BCG strains. *International Journal of Medical Sciences* 12(2): 126-134.
- Joosten, R. P., te Beek, T. A. H., Krieger, E., Hekkelman, M. L., Hooft, R. W. W., Schneider, R., Sander, C., & Vriend, G. 2010. A series of PDB related databases for everyday needs. *Nucleic Acids Research* 39 (suppl_1): D411-D419.
- Karplus, P. A. G. S. 1985. Prediction of chain flexibility in proteins. *Naturwissenschaften* 72: 212-213.
- Kolaskar, A. S. & Tongaonkar, P. C. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Letters* 276(1-2): 172-4.
- KR101631054B1, P. g. c. 2019. Antibody or Antigen Binding Fragment Binding to CFP-10 or Ag85B from *Mycobacteria*. <https://patents.google.com/patent/KR101631054B1/en>.
- Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., & Karplus, K. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* 77 Suppl 9: 114-22.
- Krieger, E. & Vriend, G. 2014. YASARA View - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics* 30(20): 2981-2.
- Kunik, V., Ashkenazi, S., & Ofran, Y. 2012. Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. *Nucleic Acids Research* 40: W521-4.
- Kuo, C. J., Ptak, C. P., Hsieh, C. L., Akey, B. L., & Chang, Y. F. 2013. Elastin, a novel extracellular matrix protein adhering to mycobacterial antigen 85 complex. *Journal of Biological Chemistry* 288(6): 3886-96.
- Lange, C., Chesov, D., Furin, J., Udhwadia, Z., & Dheda, K. 2018. Revising the definition of extensively drug-resistant tuberculosis. *Lancet Respiratory Medicine* 6(12): 893-895.
- Lata, K. S., Kumar, S., Vagharia, V., Sharma, P., Bhairappanvar, S. B., Soni, S., & Das, J. 2018. Exploring Leptospirosis proteomes to identify potential candidates for vaccine design against Leptospirosis using an immunoinformatics approach. *Scientific Reports* 8(1): 6935-6935.
- Lepore, R., Olimpieri, P. P., Messih, M. A., & Tramontano, A. 2017. PIGSPro: prediction of immunoglobulin structures v2. *Nucleic Acids Research* 45(W1): w17-w23.
- Parker, J. M., Guo, D., & Hodges, R. S. 1986. New hydrophobicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 25(19): 5425-32.
- Payandeh, Z., Rajabibazl, M., Mortazavi, Y., Rahimpour, A., & Taromchi, A. H. 2018. Ofatumumab monoclonal antibody affinity maturation through in silico modeling. *Iranian Biomedical Journal* 22(3): 180-192.
- Ryu, Y. J. 2015. Diagnosis of pulmonary tuberculosis: recent advances and diagnostic algorithms. *Tuberculosis and Respiratory Diseases* 78(2): 64-71.
- Sharma, S. K., Mohan, A. & Sharma, A. 2012. Challenges in the diagnosis & treatment of miliary tuberculosis. *Indian Journal of Medical Research* 135(5): 703-30.
- van Zundert, G. C. P., Rodrigues, J., Trellet, M., Schmitz, C., Kastiris, P. L., Karaca, E., Melquiond, A. S. J., van Dijk, M., de Vries, S. J., & Bonvin, A. 2016. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of Molecular Biology* 428(4): 720-725.
- Vangone, A., Rodrigues, J. P. G. L. M., Xue, L. C., van Zundert, G. C. P., Geng, C., Kurkcuoglu, Z., Nellen, M., Narasimhan, S., Karaca, E., van Dijk, M., Melquiond, A. S. J., Visscher, K. M., Trellet, M., Kastiris, P. L., & Bonvin, A. M. J. J. 2017. Sense and simplicity in HADDOCK scoring: Lessons from CASP-CAPRI round 1. *Proteins* 85(3): 417-423.
- Wang, C., Fu, R., Chen, Z., Tan, K., Chen, L., Teng, X., Lu, J., Shi, C., & Fan, X. 2012. Immunogenicity and protective efficacy of a novel recombinant BCG strain overexpressing antigens Ag85A and Ag85B. *Clinical & Developmental Immunology* 2012: 563838-563838.
- Xue, L. C., Rodrigues, J. P. G. L. M., Dobbs, D., Honavar, V., & Bonvin, A. M. J. J. 2016. Template-based protein-protein docking exploiting pairwise interfacial residue restraints. *Briefings in Bioinformatics* 18(3): 458-466.
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nature Methods* 12(1): 7-8.
- Zhang, W., Shu, Q., Zhao, Z., Fan, J., Lyon, C. J., Zelazny, A. M. & Hu, Y. 2018. Antigen 85B peptidomic analysis allows species-specific mycobacterial identification. *Clinical Proteomics* (15):1-1.
- Zumla, A., Ravighione, M., Hafner, R., & von Reyn, C. F. 2013. Tuberculosis. *New England Journal Medicine* 368(8): 745-55.