

Single Nucleotide Polymorphism in the *rpoB* gene *Mycobacterium tuberculosis* from Papua-Indonesia and Its Impact on Rifampicin Resistance: A Whole-Genome Sequencing Analysis

YUSTINUS MALADAN*, TRI WAHYUNI, AND HANA KRISMAWATI

Center for Papua Health Research and Development, Jalan Ahmad Yani No. 48, Jayapura, Indonesia.

In the antibiotic era, Tuberculosis (TB) drug resistance, especially Rifampicin (RIF), is highly reported around the world. Resistance of RIF is caused by the mutation of genes associated with the RIF receptor. The aims of this study are detecting the Single Nucleotide Polymorphism (SNP) of Rifampicin resistant genes using Whole-Genome Sequencing (WGS) and analyzing the profile of protein changes caused by SNP. Twenty *Mycobacterium tuberculosis* culture samples were passed on the WGS procedure and 19 samples were adequate to further bioinformatics analysis. Single Nucleotide Polymorphisms Analysis was performed using TBProfiler. Based on TBProfiler, seventeen samples were resistant to rifampicin. The mutations were found in *rpoB* gene. The mutations that cause the resistance are S456L, D441Y, H451Y, 436P, Q438K. Other Single Nucleotide Polymorphisms H841R, V540M, and R230C were also found. The H841R mutants are present together with the S456L, V540M with S456L mutants, and R230C with Q438K mutants. Native protein for RNA Polymerase Subunit β used was the result of separation from the crystal structure of *Mycobacterium tuberculosis* H37Rv RNA polymerase (PDB: 5UHB). Binding affinity RIF to RNA Polymerase Subunit β calculated using AutoDock vina. Construction of mutant 3D structures using FoldX5. From the analysis, seventeen samples were resistant to RIF and two samples did not contain SNP which could cause resistance to RIF.

Key words: molecular docking, *Mycobacterium tuberculosis*, Rifampicin, *rpoB* gene

Di era antibiotik, resistensi obat Tuberkulosis (TB), terutama Rifampisin (RIF), sering dilaporkan di seluruh dunia. Resistensi RIF disebabkan oleh mutasi gen yang berhubungan dengan reseptor RIF. Penelitian ini bertujuan untuk mendeteksi Single Nucleotide Polymorphism (SNP) pada gen yang berhubungan dengan resisten Rifampisin menggunakan Whole-Genome Sequencing (WGS) dan menganalisis perubahan profil protein yang disebabkan oleh SNP tersebut. Dua puluh sampel kultur *Mycobacterium tuberculosis* yang telah melalui proses WGS dan 19 sampel diantaranya memenuhi untuk analisis bioinformatika lebih lanjut. Analisis SNP dilakukan dengan menggunakan TBProfiler. Berdasarkan TBProfiler, tujuh belas sampel resisten terhadap rifampisin. Mutasi ditemukan pada gen *rpoB*. Mutasi yang menyebabkan resistensi adalah S456L, D441Y, H451Y, 436P, Q438K. SNP lainnya H841R, V540M, dan R230C juga ditemukan. Mutan H841R hadir bersama dengan mutan S456L, V540M dengan S456L, dan R230C dengan mutan Q438K. Protein native untuk RNA Polymerase Subunit β yang digunakan merupakan hasil pemisahan dari struktur kristal RNA polimerase *Mycobacterium tuberculosis* H37Rv (PDB: 5UHB). *Binding affinity* RIF terhadap RNA Polymerase Subunit β dihitung menggunakan AutoDock vina. Konstruksi struktur 3D mutan menggunakan FoldX5. Dari hasil analisis, tujuh belas sampel resisten terhadap RIF dan dua sampel tidak mengandung SNP yang dapat menyebabkan resistensi terhadap RIF.

Kata kunci: gen *rpoB*, *molecular docking*, *Mycobacterium tuberculosis*, Rifampisin

Mycobacterium tuberculosis is a bacterium that causes tuberculosis (TB). Tuberculosis is a global burden disease that kills about 1,3 million people. Recently, Indonesia is placed third-ranked in the world of TB after India and China. Based on the Case Detection Rate (CDR), the provinces with the highest CDR are DKI Jakarta (122.2%), South Sulawesi (84.0%), Papua (78.5%) (Kemenkes RI 2019). Increasing drug resistance numbers were reported around the world and become the challenge of TB elimination. MDR-TB is caused by *M. tuberculosis*

resistance to isoniazid and Rifampicin (RIF) with or without resistance to other first-line TB drugs (Koch *et al.* 2018; Kemenkes RI 2019). In many studies, RIF resistance was reported (Brown *et al.* 2015). The mechanism of action of RIF is by inhibiting the transcription of DNA in *Mycobacterium tuberculosis* (Floss and Yu 2005; Lin *et al.* 2017).

The analysis of TB drug resistance is performed conventionally using phenotypic drug susceptibility testing (DST). However, conventional phenotypic DST is time-consuming from more than 2 weeks until 1 month (WHO 2018). Using PCR, the time is shorter but lacking sensitivity and specificity, since only certain gene regions can be detected. WGS is an approach to

*Corresponding author: Phone: +62-8114832859; E-mail: yustinus.maladan@litbang.kemkes.go.id

overcome many challenges both related to conventional phenotypes and the limitations of lacking comprehensive molecular tests. WGS providing detailed sequence information for whole genomes, especially genes associated with drug resistance [810]. One of the genes that can be identified by WGS is the *rpoB* gene. The *rpoB* gene functions to code for the formation of RNA Polymerase Subunit β (RNAP). Mutations in the gene can be associated with RIF resistance (Floss and Yu 2005).

WGS approaching possible to sequence whole part of *M. tuberculosis* effectively analyzes the resistance of RIF through SNP detection on the gene that associated with RIF resistance. WGS can be used to detect minor variants as well as subpopulations with low-level drug resistance as information in treatment to anticipate the emergence of drug resistance (Colman *et al.* 2015). In this study, we used *Mycobacterium tuberculosis* WGS data from Papua to detect the presence of SNPs associated with RIF resistance (Maladan *et al.* 2020). This study aims to detect the SNP of RIF resistant genes using WGS and analyzing the profile of protein amino acid substitution caused by SNP.

MATERIALS AND METHODS

Samples. The research design was cross-sectional. A total of 20 samples were obtained from TB patients in BSL level 2 Jayapura Regional Health Laboratory-Papua Province. Initial identification of drug resistance was carried out using GeneXpert. The resistant positive result sample were cultured on Lowenstein-Jensen's media as many as 20 samples. Isolation of *M. tuberculosis* DNA was performed using QIAamp DNA Mini Kit (REF:51306, Qiagen, German).

Library preparation and sequencing. Quantification of extracted *M. tuberculosis* DNA was measured using Qubit™ 3.0 (Thermo Fisher Scientific). NGS procedure followed Nextera XT DNA Library Prep Kit (Illumina 2016). The NGS process is carried out using the MiSeq 600 cycle Reagent Kit (V3). Tagging of *M. tuberculosis* genome was done using Nextera transposome. Amplification of the library was then performed using Nextera PCR Master Mix. Purification of the amplification product was performed using AMPure XP beads. After purification, using Nextera XT DNA Library Preparation Kits, the library was normalized. Pooling DNA was done on a single tube and dilution was run using the Bead-Based Normalization Method (Illumina 2017). The DNA was loaded on the cartridge and WGS was run on MiSeq.

All procedures were done in the Center of Papua Health Research and Development.

Bioinformatics analysis. Genome Analysis was done using TBProfiler software (Coll *et al.* 2015) on Linux 17.10 platform. Analysis of BAM file reads visualization, consensus creation, and alignment were done using Unipro Ugene NGS 1.3.1 (Okonechnikov *et al.* 2012).

Template of RNA Polymerase Subunit β *M. tuberculosis*. The crystal structure of *M. tuberculosis* RNA polymerase protein was obtained from PDB, with PDB ID 5UHB. Subunit β RNA Polymerase was then separated from the RNA Polymerase complex using the Discovery Studio Software. Construction of the L436P, Q438K, Q438K + R230C, D441Y, H451Y, S456L, S456L + V540M, and S456L + H841R mutants using foldX5.

Ligand Preparation. RIF is obtained by separating it from Subunit β RNA Polymerase in PDB format.

Molecular Docking. Docking analysis was performed by using AutoDock Vina (Trott and Olson 2010) which is integrated with PyRx (Dallakyan 2015) application. Docking results were visualized using PyMol 2.0 and Discovery Studio 2020.

RESULT

Reads nucleotide sequence that was produced by the MiSeq V3 reagent are around 300 bp. Alignment of reads was performed using *M. tuberculosis* H37Rv (NC_000962.3) as a reference sequence (NC_000962.3). The results of alignment read then produce a consensus which is a merging number of reads at the same location of the amplified gene. In this study, we analyze the SNPs in the *rpoB* genes *M. tuberculosis* from Papua that is associated with RIF resistance. The most common mutation was Thymine to Cytosine mutation at position 1349 bp. Other mutations are changes in Thymine to Cytosine at position 1.289, Cytosine to Adenine at position 1.294, Guanine to Thymine at position 1.303, and Cytosine to Thymine at position 1.333. Mutations in these genes are characterized by the color of the mutated nucleotides (Fig 1).

The *rpoB* gene was separated from the whole genome of *M. tuberculosis* for further analysis using the Unipro Ugene NGS 1.3.1 software. After separated, the genes are translated to see the changes in amino acids caused by nucleotide mutations (Fig 2). Changes in thymine to cytosine at position 1349 bp cause

changes in the amino acid serine to leucine (S450L). Changes in Thymine to Cytosine at position 1289 cause changes in Leucine to Proline (L430P). The change in Cytosine to Adenine at position 1294 causes the conversion of Glutamine to Lysine (Q432K). The change of Guanine to Thymine at position 1303 causes the conversion of aspartic acid to Tyrosine (D435Y). Changes in Cytosine to Thymine at position 1333 cause changes in Histidine to Tyrosine (H445Y). SNPs were also identified using TBProfiler to obtain mutation data in the *rpoB* genes. Five types of mutations that cause resistance to RIF were found, namely S456L, L436P, Q438K, D441Y, and H451Y (Table 1). Furthermore, the amino acid numbering will be adjusted according to the amino acid sequence in the *M. tuberculosis* H37Rv RNAP crystal structure by adding six to the amino acid numbering.

The effect of the combination of several mutations we analyzed using the Bioinformatics approach while providing an overview of the effects of mutations that have been confirmed resistant to RIF. The *M. tuberculosis* RNA Polymerase molecule consists of six chains and the β RNA Polymerase Subunit is contained in the C chain. For further analysis, the RNAP is separated from the RNA Polymerase molecule (Fig 3A). Binding sites of RNAP were predicted using Discovery Studio software (Fig 3B). These binding sites were used to adjust the RIF grid for docking analysis.

The results of the construction of mutants L436P, Q438K, Q438K + R230C, D441Y, H451Y, S456L, S456L + V540M, and S456L + H841R show that there are differences in topology on the surface of the RNAP when compared with native proteins (Fig 4). Mutants L436P, Q438K, D441Y, H451Y, and S456L at the binding site, while the various combinations of R230C, V540M, and H841R are outside the binding site. The presence of mutations in the binding site area can cause loss of contact between RIF and RNAP and interactions that form the binding site. The surface structure of the Q438K mutants and the Q438K + R230C mutants did not differ much (Figures 4B and 4C). Likewise, the S456L and S456L V540M mutants were not much different (Fig 4F and 4G), while the S456L mutants with the S456L + H841R mutants had differences in the surface around the binding site (Fig 4F dan 4H).

The molecular docking results showed different interactions between the native *M. tuberculosis* RNAP proteins (Fig 5). Interactions occur in the form of van der Waals bonds, conventional hydrogen bond, carbon

Hydrogen bond, Pi-Sigma, Alkyl, Pi-Alkyl, Pi-Pi T-shaped, Pi-Cation, unfavorable donor-donor. In the native RNAP protein, interaction with RIF forms hydrogen bonds with residues of S456, Q438, N493, H451, F439 (Fig5A).

In the D441Y mutant, there are only two hydrogen bonds, namely the R465 and R454 residues (Fig 4E). Hydrogen bonding to the R454 residue was also found in the crystal structure of the RNAP (Lin *et al.* 2017). In the H451Y mutant, hydrogen bonds are formed with the residue T451, R454, R465, Q438 (Fig 5F). In the S456L mutant, there are hydrogen bonds that are different from the native proteins, namely D441, H1035, Q614, R613, G491 (Fig 5G). The combination of the S456L and V450M mutants has almost the same hydrogen bonds as the S456L mutants but is added by the presence of the R454 hydrogen bonds (Fig 5H). Whereas the S456L + H841R mutant only has two hydrogen bonds, namely Q438 and N443 (Fig 5I).

The energy required in the interaction between native RNAP protein and mutants with RIF is calculated in the form of binding energy (Table 2). All mutants had an increase in binding energy compared to native protein RNAP which was -10.1 kcal/mol. In energy stability, almost all mutants experienced a decrease in energy stability compared to native protein RNAP, which was 841.36 kcal/mol.

DISCUSSION

WGS results showed twenty positive MDR-TB samples showed that 19 samples were of good quality and one sample of inadequate quality. Merge reads on WGS results using TBProfiler. The output of TBProfiler in the form of BAM format is further analyzed with Unipro Ugene software. Mutations in genes associated with isoniazid resistance can be identified by looking at the difference in nucleotides in each read produced by the WGS process.

The SNPs found in the *M. tuberculosis rpoB* gene from Jayapura were L436P, Q438K, Q438K + R230C, D441Y, H451Y, S456L, S456L + V540M, and S456L + H841R. In the TB284 sample, the S456 mutation was present together with the H641R mutation whereas in the TB690 sample the S456L mutation was present together with the V540M mutation (Table 1). In the TB153 sample, the Q438K mutation was present along with the R230C mutation. Mutations L436P, Q438K, D441Y, H451Y, S456L are mutations that have been confirmed to cause resistance to *M. tuberculosis* (Phelan *et al.* 2016; Dixit *et al.* 2019).

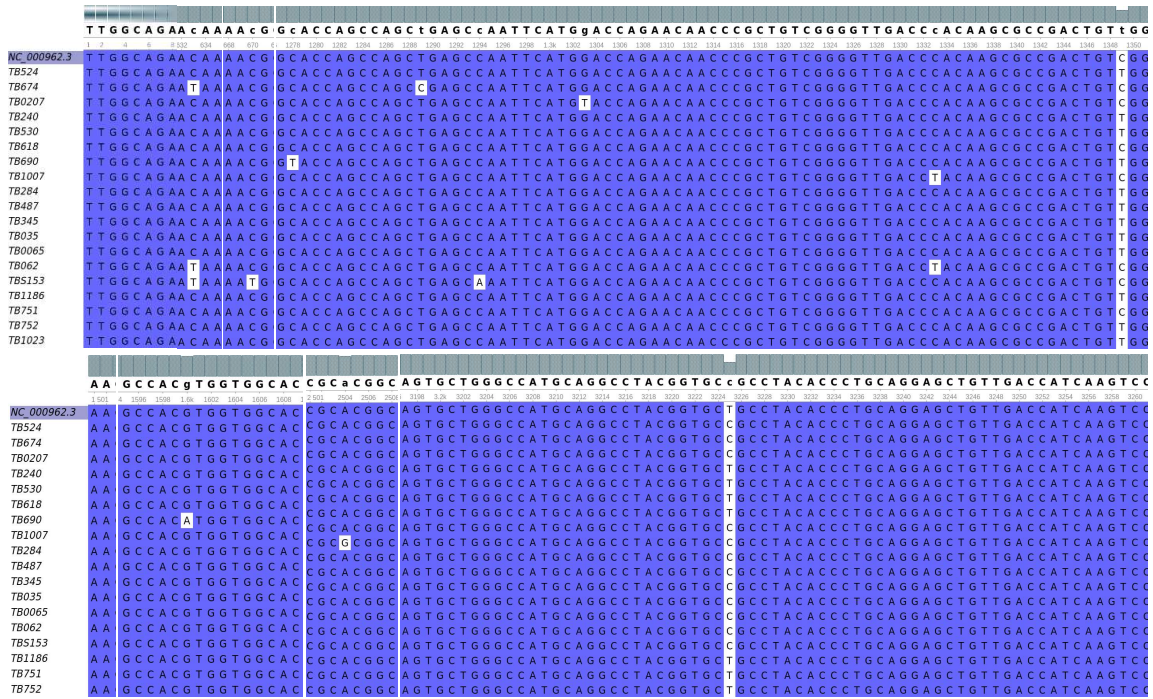


Fig 1 Mutations in the *rpoB* gene of *M. tuberculosis* of Papua from WGS. The most common mutation was T→C position 1349 bp. Other mutations are changes in T→C at position 1289, C→A at position 1294, G→T at position 1303, and C→T at position 1333. Some mutations that do not have the potential to cause resistance to RIF are C→T at position 670, G→A at position 1600, and A→G at position 2504.

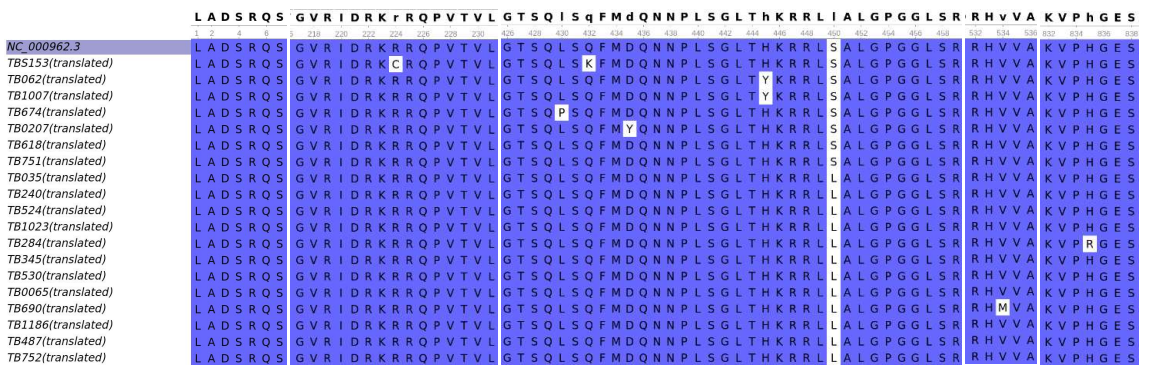


Fig 2 Gene Mutation that is associated with RIF resistance on *M. tuberculosis*. According to the TBProfiler, the mutations associated with RIF resistance are L430P (L436P), Q432K (Q438K), D435Y (D441 Y), H445Y (H451Y), and L450S (L456S).

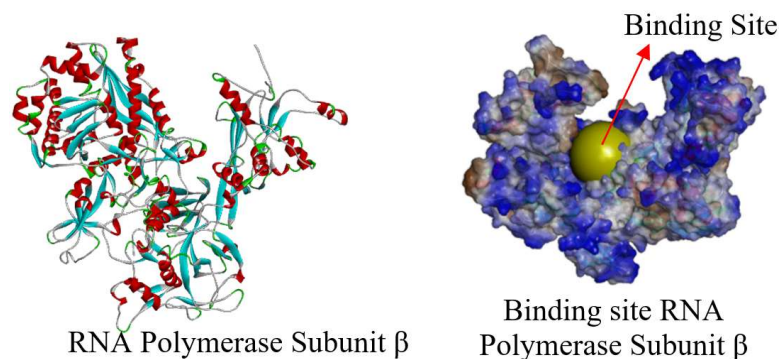


Fig 3 RNAP and topology RNAP. The RNAP binding site is marked with a yellow ball. This area is the target of RIF.

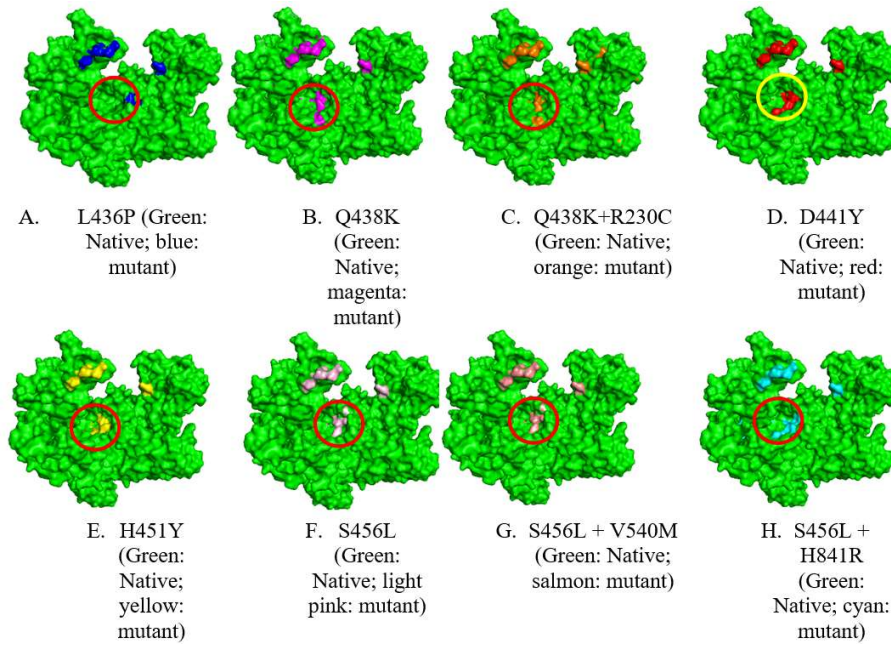


Fig 4 The surface structure of RNAP was compared with each mutant. The circled part is the binding site.

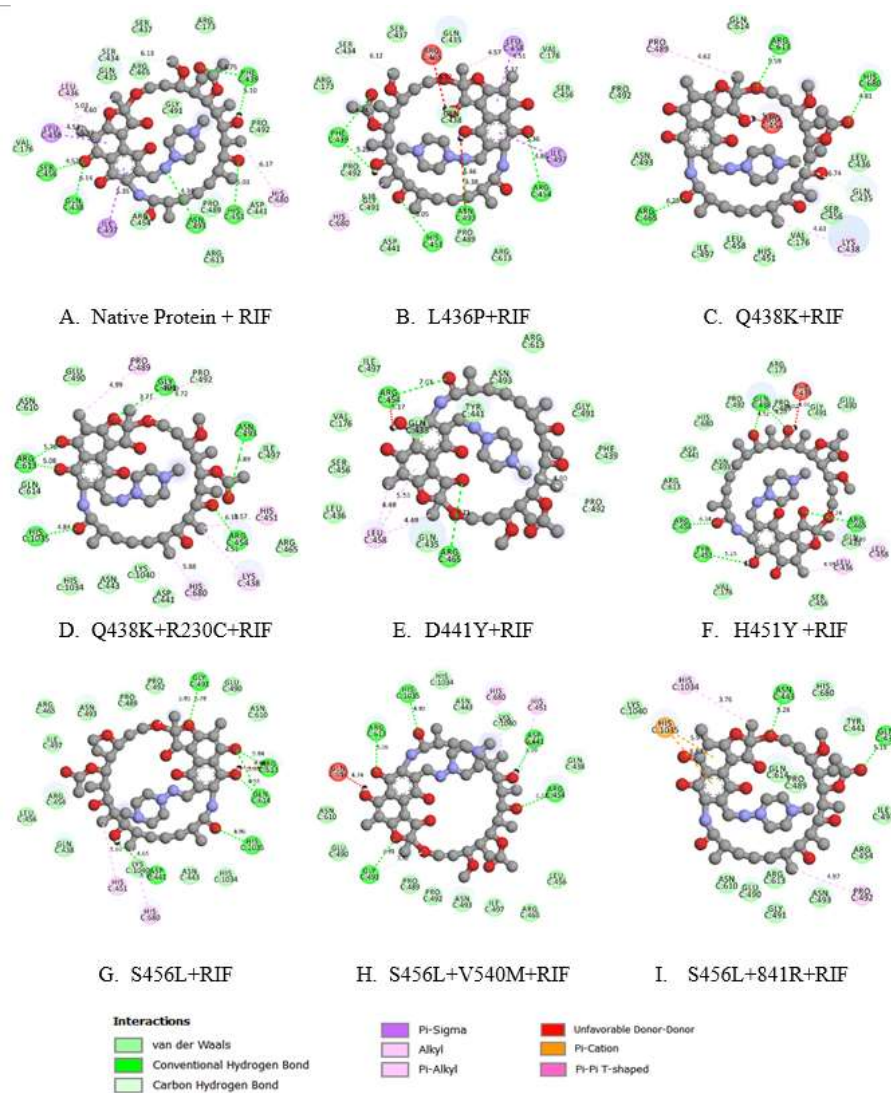


Fig 5 Map of native protein interactions and their respective mutants. RIF is displayed in ball and stick form, while the residue is shown in ball form.

Table 1 SNP mutation causes mutations as well as another mutation based on the TBprofiler database.

Code of Sample	SNP mutation		Effect mutation
	Gene	Mutation Position	
TB0065	S456L	resistant	-
TB0207	D441Y	resistant	-
TB035	S456L	resistant	-
TB062	H451Y	resistant	-
TB1007	H451Y	resistant	-
TB1023	S456L	-	-
TB1186	S456L	resistant	-
TB240	S456L	-	-
TB284	S456L	resistant	H841R
TB345	S456L	resistant	-
TB487	S456L	resistant	-
TB524	S456L	-	-
TB530	S456L	resistant	-
TB618	-	-	-
TB674	L436P	resistant	-
TB690	S456L	resistant	V540M
TB751	-	-	-
TB752	S456L	Resistant	-
TB153	Q438K	Resistant	R230C

Table 2 Stability energy RNA Polymerase and Binding Energy to RIF. All mutants had an increase in binding energy compared to native protein RNAP which was -10.1 kcal/mol. The combination of S456L with H841R has a greater increase of -6.8 kcal/mol. In energy stability, almost all mutants experienced a decrease in energy stability compared to native protein RNAP, which was 841.36 kcal/mol. The combination of S456L with H841R had a greater reduction of 824.06 kcal/mol.

Receptor	Binding Energy (kcal/mol)	Stability Energy (kcal/mol)
Native Protein RNAP	-10.1	841.36
S456L	-8.3	834.7
D441Y	-7.2	831.1
H451Y	-7.8	833.68
L436P	-9.9	843.88
Q438K	-7.2	833.75
S456L+H841R	-6.8	824.06
S456L+V534M	-8.3	833.34
Q438K +R230C	-8.1	831.34

This result is almost the same when compared to the crystal structure where hydrogen bonding occurs at the residues Q438, S456, Q435, R454, H451, F439 (Lin *et al.* 2017). The L436P mutant causes the loss of hydrogen bonds with residue S456 (Fig 5B). The Q438K mutant causes the loss of all key hydrogen bonds and is replaced by other hydrogen bonds such as R465, R613, and H680 (Fig 5C). The presence of the R230C mutation in the Q438K mutant indicates the presence of hydrogen bonding in N493 but at a further distance, namely 4.89 Å (Fig 4D). These results

probably suggest that the R230C mutation may slightly stabilize the Q438K mutant.

Native protein has a binding energy of -10.1 kcal/mol and stability energy of 841.36 kcal/mol. The S456L mutant and the combination of the S456L and V534M mutant have the same value, namely -8.3 kcal/mol. This value is greater than the native protein. This increase in binding energy can result in loss of contact with RIF. The combination of the S456 and H841R mutants causes greater binding energy than native RNAP proteins. The other mutants also

experienced increased binding energy with RIF, respectively, D441Y, H451Y, L436P, Q438K, and Q438K + R230C were -7.2, -7.8, -9.9, -7.2, -8.1 kcal/mol. Changes in hydrogen bonding to the key residue and increasing energy in all mutants are indicators of resistance to RIF. The results of this molecular docking analysis showed the same results as the results of the Drug Susceptibility Test (DST) which showed that mutants L436P, Q438K, D441Y, H451Y, S456L were resistant to RIF (Heep *et al.* 2001). These results illustrate how resistance to RIF is caused by mutations in the *M. tuberculosis rpoB* gene.

From the analysis, it was found that seventeen samples were resistant to rifampicin and two samples did not contain SNP which could cause resistance to rifampicin. The SNPs found in the *M. tuberculosis rpoB* gene from Jayapura were L436P, Q438K, Q438K + R230C, D441Y, H451Y, S456L, S456L + V540M, and S456L + H841R. The molecular docking results showed the same results as the results of the Drug Susceptibility Test (DST) as an illustration of how mutations in the *M. tuberculosis rpoB* gene could affect RIF. This method can be used to predict the effect of new mutations in the *M. tuberculosis rpoB* gene that have not been confirmed by the DST method.

ACKNOWLEDGEMENTS

Thanks to the Regional Health Laboratory of Jayapura for contributing to sample collection, culture and gene expert analysis.

REFERENCES

- Brown AC, Bryant JM, Einer-Jensen K, Holdstock J, Houniet DT, Chan JZ, Depledge DP, Nikolayevskyy V, Broda A, Stone MJ, Christiansen MT, Williams R, McAndrew MB, Tutill H, Brown J, Melzer M, Rosmarin C, McHugh TD, Shorten RJ, Drobniewski F, Speight G, Breuer J. 2015. Rapid whole-genome sequencing of *Mycobacterium tuberculosis* isolates directly from clinical samples. *J Clin Microbiol.* 53(7):2230-7. doi: 10.1128/JCM.00486-15.
- Coll F, McNerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill-Cawthorne G, Mallard K, Nair M, Miranda A, Alves A, Perdigão J, Viveiros M, Portugal I, Hasan Z, Hasan R, Glynn JR, Martin N, Pain A, Clark TG. 2015. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med.* 7(1):51. doi: 10.1186/s13073-015-0164-0.
- Colman RE, Schupp JM, Hicks ND, Smith DE, Buchhagen JL, Valafar F, Crudu V, Romancenco E, Noroc E, Jackson L, Catanzaro DG, Rodwell TC, Catanzaro A, Keim P, Engelthaler DM. 2015. Detection of low-level mixed-population drug resistance in *Mycobacterium tuberculosis* using high fidelity amplicon sequencing. *PLoS One.* 10(5):118. doi:10.1371/journal.pone.0126626.
- Dallakyan S. 2015. Small-molecule library screening by docking with PyRx. *Methods Mol Biol.* 1263:243250. doi:10.1007/978-1-4939-2269-7.
- Dixit A, Freschi L, Vargas R, Calderon R, Sacchetti J, Drobniewski F, Galea JT, Contreras C, Yataco R, Zhang Z, Lecca L, Kolokotronis SO, Mathema B, Farhat MR. 2019. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Sci Rep.* 9(1):5602. doi: 10.1038/s41598-019-41967-8.
- Floss HG, Yu T. 2005. Rifamycin mode of action, resistance, and biosynthesis. *Chem Rev.* 105:621632.
- Gygli SM, Keller PM, Ballif M, Blöchliger N, Hömke R, Reinhard M, Loiseau C, Ritter C, Sander P, Borrell S, Collantes Loo J, Avihingsanon A, Gnokoro J, Yotebieng M, Egger M, Gagneux S, Böttger EC. 2019. Whole-genome sequencing for drug resistance profile prediction in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother.* 63(4):e02175-18. doi: 10.1128/AAC.02175-18.
- Heep M, Brandstätter B, Rieger U, Lehn N, Richter E, Rüscher-Gerdes S, Niemann S. 2001. Frequency of *rpoB* mutations inside and outside the cluster I region in rifampin-resistant clinical *Mycobacterium tuberculosis* isolates. *J Clin Microbiol.* 39(1):107110. doi:10.1128/JCM.39.1.107-110.2001.
- Iketleng T, Lessells R, Dlamini MT, Mogashoa T, Mupfumi L, Moyo S, Gaseitsiwe S, Oliveira T De. 2018. *Mycobacterium tuberculosis* next-generation whole genome sequencing: Opportunities and challenges. *Tuberc Res Treat.* 2018:18.
- Illumina. 2016. Nextera® XT Library Prep Reference Guide. http://support.illumina.com/downloads/nextera_xt_sample_preparation_guide_15031942.html.
- Illumina. 2017. MiSeq system denature and dilute libraries guide. http://os.bio-protocol.org/attached/file/20171217/miseq_denature_dilute_libraries_guide_1503974003.pdf.
- Kemenkes RI. 2019. Profil Kesehatan Indonesia 2018. Jakarta. <http://www.depkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/Profil-Kesehatan-Indonesia-2016.pdf>.
- Koch A, Cox H, Mizrahi V. 2018. Drug-resistant tuberculosis: challenges and opportunities for

- diagnosis and treatment. *Curr Opin Pharmacol.* 42:715. doi:10.1016/j.coph.2018.05.013.
- Koser CU, Ellington MJ, Peacock SJ. 2014. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet.* 30(9):401-407. doi:10.1016/j.tig.2014.07.003.
- Lin W, Mandal S, Degen D, Liu Y, Ebright YW, Li S, Feng Y, Zhang Y, Mandal S, Jiang Y, Liu S, Gigliotti M, Talaue M, Connell N, Das K, Arnold E, Ebright RH. 2017. Structural basis of *Mycobacterium tuberculosis* transcription and transcription inhibition. *Mol Cell.* 66(2):169-179.e8. doi: 10.1016/j.molcel.2017.03.001.
- Maladan Y, Krismawati H, Oktavian A, Lestari CSW. 2020. Improving multidrug-resistance Tuberculosis Papuas management using whole genome sequencing. In: *Advances in Health Sciences Research.* Vol. 22. p. 17.
- Miotto P, Cabibbe AM, Borroni E, Degano M, Cirillo DM. 2018. Role of disputed mutations in the *rpoB* gene in interpretation of automated liquid MGIT culture results for rifampin susceptibility testing of *Mycobacterium tuberculosis*. *J Clin Microbiol.* 56(5):19. doi:10.1128/JCM.01599-17.
- Okonechnikov K, Golosova O, Fursov M, Varlamov A, Vaskin Y, Efremov I, German Grehov OG, Kandrov D, Rasputin K, Syabro M, et al. 2012. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics.* 28(8):1166-1167. doi:10.1093/bioinformatics/bts091.
- Phelan J, Coll F, Mc Nerney R, Ascher DB, Pires DEV, Furnham N, Coeck N, Hill-Cawthorne GA, Nair MB, Mallard K, Ramsay A, Campino S, Hibberd ML, Pain A, Rigouts L, Clark TG. 2016. *Mycobacterium tuberculosis* whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance. *BMC Med.* 14(31):113. doi:10.1186/s12916-016-0575-9.
- Satta G, Lipman M, Smith GP, Arnold C, Kon OM, Mchugh TD. 2017. *Mycobacterium tuberculosis* and whole-genome sequencing : how close are we to unleashing its full potential? *Clin Microbiol Infect.*:16. doi:10.1016/j.cmi.2017.10.030.
- Trott O, Olson A. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem.* 31(2):455-461. doi:10.1002/jcc.21334.AutoDock.
- WHO. 2018. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide. Geneva.