Quantitative structure–activity relationship (QSAR) and molecular docking of xanthone derivatives as anti-tuberculosis agents

Emmy Yuanita a,b,*, Sudirman a, Ni Komang Tri Dharmayani a, Maria Ulfa a, Jufrizal Syahri b,*

a Department of Chemistry, Universitas Mataram, Majapahit Street 62, Mataram 83236, Indonesia
b Department of Chemistry, Universitas Muhammadiyah Riau, Tuanku Tambusai Street, Pekanbaru, Indonesia

ARTICLE INFO

Keywords:
Docking
QSAR
Xanthone
Anti-tuberculosis
kasA inhibitor

ABSTRACT

Quantitative structure–activity relationship (QSAR) and molecular docking approach were carried out to design novel anti-tuberculosis agents based on xanthone derivatives. QSAR designed new compounds were calculated by Austin Model 1 (AM1) methods and analysis of multi-linear regression (MLR). The result showed that the best model as follows: Log IC50 = 3.113 + 11.627 qC1 + 15.955 qC4 + 11.702 qC9, this result has appropriate some statistical parameters (PRESS = 2.11, r2 = 0.730, SEE = 0.3545, R = 0.6827, FCal/FTab = 4.68), and being used to design a potential anti-tuberculosis drugs with substituted amide, sulfoxide, and carboxylate group xanthone scaffold by a number of their inhibitory concentration (IC50). The mechanism action of sulfonamide substituted on the xanthone scaffold as anti-tuberculosis was carried out using molecular docking. Docking inhibition studies were carried out on MTB C171 Q receptor (4G6X.pdb) as kasA inhibitors using by the discovery studio. Based on the binding interaction showed, the sulfonamide substituted xanthone has potential being the anti-tuberculosis drugs by KasA inhibitor for target drug activity.

1. Introduction

Tuberculosis (TB) is an infectious bacterial disease caused by Mycobacterium tuberculosis (M. Tuberculosis), which known has been spread globally and has become chronic infectious disease from decades. In 2018, the World Health Organization (WHO) reported there are over 10 million people fell ill with TB, and about 1.2 million death among HIV-negative people [1]. One of the most common strategies against TB infectious is Directly Observed Treatment Short (DOTS). However, this method could ineffective and develop resistance toward anti-TB drugs if the treatment is done in inappropriate ways. It will cause Multidrug-resistant TB (MDR-TB) if the anti-TB drug regimens are a precept in inaccurate doses or poor-quality medicines. Furthermore, it will develop Extensive drug-resistant TB (XDR-TB) if used to treat patients with immune problems or patients with HIV infected [2]. Standard anti-TB drugs such as Isoniazid (INH), Rifampicin (RIF), Ethambutol (EMB), D-Cycloserine (DSC), and Pyrazinamide (PZA), Rifabutin (RBT) and Rifapentine (RPT) are also indicated resistant to prevent and cure the TB disease [1,3]. The drug-resistant to treat this disease is still the primary issue of why the number of deaths and lethality is hard to combat. Therefore the discovery and development of new effective anti TB drugs are extremely needed.

Drug discovery is an effort to address health problems, including TB. A practical approach that may be considered to discover the drug candidates rapidly is to modify natural products structure which activities are well established, and then to apply statistic analytical methods (e.g., the well-known quantitative structure–activity relationship (QSAR) study) to establish correlations between chemical structures and the corresponding biological activities [4–5]. QSAR analysis has been widely used to explore synthesized compounds and their derivatives as anti-TB. Recently reports that bedaquiline and its analogues compounds have an activity with the enzyme targets of M.Tuberculosis and may have effectiveness in MDR and XDR-TB through QSAR analysis [6]. This approach reduced the cost for the effective evaluation of synthesized compounds. Such compounds include 3-heteroaryl thioquinoline derivatives [7]; 8-methylquinolones [8]; 7-chloroquinoline derivatives [9] and substituted benzothiazole/benzimidazole analogs [10] cinnamic acid derivatives [11]. While the study of QSAR xanthone compounds as anti-TB has never been done, what has been published is the analysis of Q SAR xan thones as inhibition of monoamine oxidase (MAO) [12], α-glucosidase inhibitors [13], anti-malaria [14] and anti-cancer [15]. Furthermore, to find out the mechanism action of novel compounds

* Corresponding authors.
E-mail addresses: emmy_yuanita@unram.ac.id (E. Yuanita), jsyachri@umri.ac.id (J. Syahri).

https://doi.org/10.1016/j.jctube.2020.100203
has predicted by QSAR analysis carried out using molecular docking. Docking for anti-tuberculosis carried out on MTB C171Q receptor (4C6X.pdb) as KasA inhibitors. KasA inhibitors is one attractive target drugs [16]. The KasA enzyme is also essential for mycobacterial survival. This Homodimeric β-ketoacyl-ACP synthase (KAS) catalyzes a Claisen condensation reaction between acyl-AcpM and malonyl-AcpM in each elongation cycle of the FAS-II pathway [17], that is mycolic acid precursors [18]. Mycolic acid is one of essential compound has supporting the mycobacterium tuberculosis wall [19].

2. Experimental section

2.1. Materials

2.1.1. QSAR analysis

Xanthone derivative compounds were taken from Suksamrarn et al. (2003) [20]. The structures of xanthone derivatives, as well as their anti-tuberculosis activity, are presented in Table 1.

2.1.2. Docking analysis

The crystal structure of MTB C171Q used in the study was obtained from the protein data bank (www.rcsb.org/structure) with PDB code 4c6x as KasA inhibitors. KasA inhibitors are known as attractive target drugs for TB [14]. The KasA enzyme is also essential for mycobacterial survival. This homodimeric β-ketoacyl-ACP synthase (KAS) catalyzes a Claisen condensation reaction between acyl-AcpM and malonyl-AcpM in each elongation cycle of the FAS-II pathway [17]. One of component mycobacterial cell wall especially for Mycobacterium tuberculosis is dependent on mycolic acids [18], and the biosynthesis of mycolic acids requires the presence of two distinct fatty acid synthesis pathways, the mammalian-like type I (type I fatty acid synthesis (FAS-I)) and the bacterial type II (FAS-II) systems. The C16-18 acyl-CoA primers that were synthesized by the FAS-I system are elongated by the FAS-II pathway to C16-18 mycolic acid precursors [18].

---

Table 1

| Comp. id | Compound | pMIC |
|----------|----------|------|
| 1        | ![Image 1](image1.png) | 0.79588 |
| 2        | ![Image 2](image2.png) | 1.39794 |
| 3        | ![Image 3](image3.png) | 1.39794 |
| 4        | ![Image 4](image4.png) | 2 |
| 5        | ![Image 5](image5.png) | 0.79588 |
| 6        | ![Image 6](image6.png) | 1.09691 |
| 7        | ![Image 7](image7.png) | 1.39794 |
| 8        | ![Image 8](image8.png) | 2.30103* |
| 9        | ![Image 9](image9.png) | 1.39794 |
| 10       | ![Image 10](image10.png) | 0.79588* |

Table 1 (continued)

| Comp. id | Compound | pMIC |
|----------|----------|------|
| 11       | ![Image 11](image11.png) | 1.09691 |
| 12       | ![Image 12](image12.png) | 1.09691* |
| 13       | ![Image 13](image13.png) | 1.39794* |

*Tes set compound.
2.1.3. Instrumentation

2.1.3.1. QSAR and docking analysis. For this study, a PC was equipped with Intel® Core™ i5. All the compounds (Table 1) were calculated using package Gaussian® 09 W and statistical program IBM® SPSS® version 17.0.0 for Windows. The molecular docking was done on previous work [14,21] under the receptor-ligand interaction section in Discovery Studio 3.1 (Accelrys, Inc., San Diego, CA, USA). CHIMERA 1.9 and ChemOffice® 2015 software were also used to study the molecular modeling of the compound [14,21].

2.2. Procedure

2.2.1. QSAR analysis

2.2.1.1. Data set. The QSAR model’s preparation for 13 starts with dividing the data into two groups, namely training sets and test sets. The distribution of training sets (9 compounds) and test sets (4 compounds) was carried out randomly (Table 1). A number of compounds in the training set, analyzed using Multiple Linear Regression (MLR) with the backward method run on the statistical program (SPSS® Release 16.0.0) to obtain several models of the relationship between MIC logs with electronic descriptors or for generating QSAR models and a test set for validating the quality of the models.

2.3. Computational validations

Method validation is done by calculating the chemical shift value of 13C NMR xanthone compounds (as samples) using various methods, namely Austin Model 1 (AM1), Parameterized Model Number 3 (PM3), Hartree-Fock (HF), and Density Functional Theory (DFT). The chemical shift value of 13C NMR for each method is compared with the chemical shift value of 13C NMR from experimental results to find out the best method. Furthermore, r and PRESS’s value is determined from the difference in the value of the experimental and predicted data shifts. The method with r ~ 1 value and the smallest PRESS value is the best method, and it is used for further geometry optimization.

2.4. The development of QSAR model

The QSAR model was generated by the MLR Backward method using the SPSS package based on electronic parameters from a calculation using the Gaussian package with the selected method before. This method refers to the dependent variable y (biological activity) with a number of independent variables xi (electronic descriptors) using linear equations. Moreover, this regression method estimates the values of the regression coefficients by applying the least square curve fitting method. The model was chosen for QSAR calculation based on some statistical equations. Moreover, this regression method estimates the values of the regression coefficients by applying the least square curve fitting method. The best equation model obtained from the analysis using the MLR Backward method is selected for the final model. The model is said to be valid if the predicted value were chosen as a new antituberculosis compound. Herein, we describe an unprecedented QSAR study on a series of xanthone derivatives to provide comprehensive insight into the correlation between structures and activities of xanthones as anti TB. The QSAR model’s potential in guiding future effort in the rational design and synthesis of xanthone derivatives potentially have high-inhibitory activities toward anti TB. Based on the predictions of the QSAR Model, as shown in Table 2, the 3,6-dihydroxy and 1,3,6 trihydroxy xanthone derivatives are known to have good activity as a drug antituberculosis with the addition of amide, sulfoxide, and carboxylate group. Xanthone derivatives which have been successfully synthesized 3,6-dihydroxyxanthone and 1,3,6 trihydroxyxanthone is an intermediate compound [21,26].

| Position | δ Experiment | δ Calculation |
|----------|--------------|---------------|
| H-8      | 8.03         | 7.90          | 8.11          | 8.41          | 8.88          |
| H-5      | 6.85         | 7.52          | 7.48          | 6.60          | 7.04          |
| H-4      | 6.40         | 7.17          | 7.14          | 6.29          | 6.77          |
| H-7      | 6.93         | 6.25          | 6.16          | 5.71          | 6.09          |
| H-2      | 6.21         | 5.04          | 4.97          | 4.44          | 4.88          |
| C-1      | 163,98       | 161,96        | 160,00        | 156,53        | 161,84        |
| C-2      | 98,10        | 92,60         | 91,44         | 89,32         | 93,65         |
| C-3      | 165,10       | 162,81        | 161,72        | 151,62        | 156,30        |
| C-4      | 93,90        | 100,14        | 100,34        | 92,26         | 97,19         |
| C-4a     | 156,00       | 162,05        | 161,80        | 157,47        | 162,08        |
| C-5      | 102,30       | 106,95        | 106,77        | 100,91        | 105,84        |
| C-6      | 164,10       | 161,67        | 160,19        | 156,26        | 161,39        |
| C-7      | 113,80       | 93,46         | 105,23        | 103,65        | 106,66        |
| C-8      | 127,50       | 134,08        | 132,48        | 111,25        | 115,09        |
| C-9a     | 113,40       | 112,98        | 112,41        | 103,12        | 106,66        |
| C-9      | 179,80       | 180,29        | 174,96        | 177,09        | 187,15        |
| C-9a     | 102,50       | 104,13        | 105,22        | 98,23         | 102,57        |
| C-10a    | 158,00       | 157,77        | 156,88        | 131,45        | 135,04        |
| PRESS    | 585,88       | 297,16        | 1587,06       | 983,38        |

2.6. Design and activities prediction of the new compounds

The newly designed and predicted xanthone derivative compounds based on the validated data with the best inhibitory activity and a lower MIC50 value were chosen as a new antituberculosis compound. Herein, we describe an unprecedented QSAR study on a series of xanthone derivatives to provide comprehensive insight into the correlation between structures and activities of xanthones as anti TB. The QSAR model’s potential in guiding future effort in the rational design and synthesis of xanthone derivatives potentially have high-inhibitory activities toward anti TB. Based on the predictions of the QSAR Model, as shown in Table 2, the 3,6-dihydroxy and 1,3,6 trihydroxy xanthone derivatives are known to have good activity as a drug antituberculosis with the addition of amide, sulfoxide, and carboxylate group. Xanthone derivatives which have been successfully synthesized 3,6-dihydroxyxanthone and 1,3,6 trihydroxyxanthone is an intermediate compound [21,26].

3. Results and discussion

3.1. Validation method

The validation of computational methods conducted in this study is based on the accuracy of the chemical shift comparison of 1H and 13C NMR experimental data compared to chemical shift data obtained from calculations, namely ab initio HF, DFT, and semi-empirical AM1 and PM3. The chemical shift calculation was obtained from Gaussian 09 software with geometry optimization using the HF (6-31G), DFT (B3LYP, 6-31G), AM1, and PM3 methods and stored in file.log, and then the chemical shift was calculated using the GIAO SCF DFT method (B3LYP –6-31G). The selection method will be used based on the parameter correlation value (r) and the value of the Predictive Residual Sum of Square (PRESS). In theory, the higher the value of r and the lower the PRESS value indicates the results of the calculation of the chemical shift, the closer the experimental results. Linearity or correlation of experimental chemical shifts and chemical shift calculations are determined by finding the value of r2 by making a graph, as shown in Fig. 1 a-d. The PRESS value is the value of the difference in the results of calculations and experiments shown in Table 2. Therefore, the AM1 method has been selected as a calculation method for further modeling the anti-tuberculosis activity of xanthones derivatives.
independent variables, namely the electronic descriptors used on anti-tuberculosis activity by enter all descriptors simultaneously, and thereby maximizing all variable used which is formulated in the equation [27, 28]. Based on the statistical analysis results using the MLR (backward) analysis, the value of the Fcalculate and Ftable ratio (Fcal / Ftab). 2 QSAR equation models are generated, as shown in Table 3. The best QSAR equation is selected based on statistical parameters that indicate the significance level of the model, namely the coefficient of determination (r2), Standard Error of Estimate (SEE), and PRESS (predictive residual sum of the square) could be taken into account. Comparison of the mentioned parameters (r2, SEE, and PRESS) to the two models, presented in Table 4. It pointed out that it was also not easy to choose the best model because their values were not significantly different. However, it has to be continued validation against both equations is obtained to ensure that the equation with parameters most appropriate statistics.

### 3.2. Generation and selection of QSAR model

The preparation of the QSAR equation model begins with selecting training sets and test sets. A total of 13 xanthone derivative compounds used in this study were randomly divided into two data groups: the training set (9 compounds) and the test set (4 compounds). The training set compounds were analyzed statistically to produce the QSAR model, while the test set compounds were used to validate the QSAR model generated in the training set. The grouping of data is shown in Table 1. Preparation of the model (prior to regression analysis), the MIC value is converted to a logarithmic scale with the aim that the range of MIC values between compounds does not differ significantly, and the distribution of MIC values is getting better.

The electronic descriptors’ determination was carried out using the AM1 semi-empirical method obtained from the results of the method validation of all test compounds, and both included in the training set and test set. The electronic descriptor used is the net charge of atoms (q) in the parent frame of xanthone compounds, amounting to 15 atoms, namely qC1, qC2, qC3, qC4, qC5, qC6, qC7, qC8, qC9, qC10, qC11, qC12, qC13, qO14, and qO15. MLR (backward) analysis is then performed to get the maximum correlation between electronic descriptors on antituberculosis activity by enter all descriptors simultaneously, thereby maximizing all variable used which is formulated in the equation [27, 28]. Based on the statistical analysis results using the MLR (backward) method, 2 QSAR equation models are generated, as shown in Table 3. The best QSAR equation is selected based on statistical parameters that indicate the significance level of the model, namely the coefficient of determination (r2), Standard Error of Estimate (SEE), as well as the value of the Fcalculate and Ftable ratio (Fcal / Ftab).

From the data in Table 3, it can be seen that the 2 QSAR models produced have an R2 value greater than 0.6, which is in the range of 0.726–0.869 (fulfilling specified criteria). Thus illustrates the effect of independent variables, namely the electronic descriptors used on anti-tuberculosis activity, which is very large, which is more than 72–86%. Furthermore, observed from the SEE values for all models ranging from 0.192 to 0.246, it can be concluded that this value indicates the accuracy of the resulting model for predicting new anticancer compounds from the xanthone derivatives is approving (close to 0). This result indicated the determination of the best models among 2 QSAR models listed in Table 3 was not adequate only by comparing the size, while they had similar values. Therefore, other statistical parameters such as R2, standard estimation of error (SEE < 0.3), and also PRESS (predictive residual sum of the square) could be taken into account. Comparison of the mentioned parameters (r2, SEE, and PRESS) to the two models, presented in Table 4. It pointed out that it was also not easy to choose the best model because their values were not significantly different. However, it has to be continued validation against both equations is obtained to ensure that the equation with parameters most appropriate statistics.

### 3.3. Validation of QSAR models

QSAR Model (1) and (2) were applied to calculate the activity of 5 test set compounds, and we could see how good these models predict the anti-tuberculosis activity of the xanthone derivatives. Validation for searching the best model was performed using the calculation for each equation toward the test set. The PRESS value is the amount of quadrat difference between the predicted and observed Log MIC values, where the equation with the smallest difference was chosen as the best equation. PRESS value from the predicted and observed Log MIC of model 1 and 2 were listed in Table 4.

Based on the PRESS parameter, the smallest PRESS value was shown...
by model 2, with 9.2 indifference, compared with model 1 with 14.48 in value. This result indicated that it has a small difference in anti-tuberculosis activity as MIC values within the predicted and experimental. It could be decided from PRESS value that model 2 was better than model 1 as the QSAR model to predict the anti-tuberculosis activity, and also it could give a good structure of the predicted xanthone derivative compound.

Other parameters to convince the best model was with compared their slope and correlation coefficient (r²). As shown in Fig. 2, it was determined that r² of models 1 and 2 were 0.548 and 0.655, respectively.

According to PRESS and r² parameters of statistical analysis, model 1 would be the best QSAR model for generating the validation model toward the test set. Further validation for determining the best QSAR model was evaluated using ENTER statistical calculation. In the end, based on the calculation toward some statistical parameters, we found the result as follows: PRESS = 2.11, r² = 0.730, SEE = 0.3545, R = 0.6827, FCAL/FTAB = 4.68. Plots of the predicted versus experimental of anti-tuberculosis as Log MIC values were shown in Fig. 3.

The correlation coefficient (r²) was 0.6827, and it means there is a 68.27% similarity value between predicted and experimental anti-tuberculosis activity. Based on statistics obtained the renewal of the parameter values of equation 1, it could be generated the best equation model of QSAR regression as follows:

\[ \text{Log MIC} = 3.113 + 11.627 \, qC1 + 15.955 \, qC4 + 11.702 \, qC9 \]

Based on Equation 1 (Eq.1), electronic descriptors for electron charge of carbon atom in numbers 1, 4, and 9 (Fig. 4) are indicated as the position with the most effective to the anti-tuberculosis activity. The better anti-tuberculosis activity as MIC could be given by the more negative of the log MIC. The most negative value of log MIC can be achieved by sorting the negative net atomic charge of qC1, qC4 and qC9 or neighbor’s position should be occupied with electron-donating groups (EDG) such as hydroxyl, Sulfoxide, ester or amide and methoxy group to make the π system more nucleophile. So in this study, it has been conducted to design a new hydroxyl xanthone by modification of hydroxyl xanthone scaffold with sulfoxide groups, amide, or derivative of acid, to look out their predicted tuberculosis activity, as listed as Table 5.

### 3.4. Docking

Based on QSAR analysis, the best compound displayed has potency as anti-tuberculosis is sulfonamide substituted, hence how the mechanism of activity of this compound has to be analyzed by molecular docking, because this method is to complement the results of other studies that have been conducted on fluoroquinoline compounds using pharmacophore modeling showed hydrophobic hydrophobicity groups of substitutents, and hydrogen bond play an important role for the DNA gyrase inhibition [29]. Different results are shown by xanthones derived compounds. The first-line anti-tuberculosis drug isoniazid inhibits the enoyl ACP reductase InhA, but the isoniazid h, thereby validating the FAS-II pathway as a promising target for the development of novel antibiotics. FAS-II is one of the precursors for the biosynthesis of mycolic acid. Mycolic acid is one of the essential compounds that has supported antibiotics. FAS-II is one of the precursors for the biosynthesis of mycolic acid. Mycolic acid is one of the essential compounds that has supported

![](attachment:image1.png)

**Fig. 4.** Structure of xanthone.
inhibitory mechanisms similar to isoniazid drugs, namely as an Inhibitor KasA, whose role is to stop the formation of Mycobacterium tuberculosis cell walls.

4. Conclusion

Based on the experimental result and discussion that has explained, it can be concluded that the best QSAR equation was: \( \text{Log MIC} = 3.113 + 11.627 qC1 + 15.955 qC4 + 11.702 qC9 \). The 3,6-dihydroxy and 1,3,6-trihydroxy xanthone derivatives are known to have good activity as a drug anti-tuberculosis with the addition of amide, sulfoxide, and carboxylate groups. The docking analysis results showed the inhibitory mechanism as KasA inhibitor, the inhibitory mechanism in the cell wall of Mycobacterium tuberculosis.

CRediT authorship contribution statement

**Emmy Yuanita:** Conceptualization, Methodology, Software. **Sudirman:** Software, Validation. **Ni Komang Tri Dharmayani:** Writing - review & editing. **Maria Ulfa:** Supervision. **Jufrizal Syahri:** Data curation, Writing - original draft.
Acknowledgements

Financial support for these works from Penelitian Dasar (PD) and Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) of the ministry of research, technology, and Higher education-Indonesia is gratefully acknowledged for years 2019 with grand number 1830/UN18.L1/PP/2019. Similarly, the author is grateful to CADD Laboratory, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, for providing the Discovery Studio 3.1 licenses.

References

[1] WHO. Global tuberculosis report, 2018.
