Walisongo Journal of Chemistry
Walisongo Journal of Chemistry Vol. 5 Issue 2 (2022), 194-201
ISSN: 2621-5985 (online); 2549-385X (print)
DOI: https://doi.org/10.21580/wjc.v5i2.13412

Analysis of Quantitative Structure-Activity Relationship (QSAR) Of 1,8-Napthalymide-4-Aminoquinoline Derivatives as Antimalarial Compounds

Ika Nur Fitriani[1*], Muhammad Sholahuddin[1], Adi Tiara Zikri[2]

[1]Department of Chemistry, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang, Indonesia
[2]Department of Material Science and Engineering, Gyeongnam National University, South Korea

*Corresponding author: ikanurf@walisongo.ac.id

Received: 31 October 2022; Accepted: 20 December 2022; Published: 25 December 2022

Abstract

QSAR analysis from derivative compounds of 1,8-naphthalimide-4-aminoquinoline has been carried out with the aim of knowing the quantitative relationship between structure and activity based on the log IC50 value. The modelling was made with two dimensions and geometric optimization is carried out. The AM1 semi-empirical method was used to optimize the geometry of the total energy of 25 compounds. The descriptors in the QSAR calculation were dipole moment, atomic net charge, ELUMO, EHOMO, SAG, SAA, hydration energy, polarizability, and log P. The descriptors were used to find quantitative equations of the relationship between the structure and antimalarial activity.

Keywords: *AM1; antimalarial; QSAR; 1,8-naphthalimide-4-aminoquinoline; derivative compounds*

Introduction

Malaria is a disease that receives very serious attention worldwide. The number of reported human deaths from malaria is still relatively high. The World Health Organization (WHO) stated that from 2016 to 2020, the number of cases worldwide increased from 226 million to 241 million, and the number of deaths increased from 566,000 to 627,000 (WHO, 2021).

The general mechanism action of malaria drugs is to form a strong binding complex with hematin, inhibiting the formation of hemozoin that accumulates in the digestive vacuoles of the malaria parasite. This cuts off the food supply and kills parasites (Tilley, Loria and Foley, 2003)

For decades, one of the malaria drugs (prophylaxis) is chloroquine. The drug is widely used to treat malaria worldwide (White et al., 2014). The drug chloroquine is also used to treat patients infected with COVID-19, but the Food and Drug Administration (FDA) cancelled its emergency use permit in July 2020 (FDA, 2020).

In recent years, the parasite has become resistant to antimalarial drugs, such as *Plasmodium falciparum* resistance to chloroquine and artemisinin, and is widespread in various malaria-affected areas (MAG, 2018). The existence of resistance to standard antimalarial drugs indicates the need to find new antimalarial compounds with higher or at least comparable therapeutic efficacy.

Synthesis of 1,8-naphthalimido-4-aminoquinoline derivatives and obtained 25 conjugated derivatives, 19 compounds of them had an inhibitory concentration value

of 50 (IC50) lower than that of chloroquine (Shalini *et al.*, 2020). The IC50 for chloroquine as a standard antimalarial drug was reported to be 232.65 nM by means of in-vitro test. IC50 is the drug concentration required for 50% inhibition (Swinney, 2011). The IC50 value indicates that a compound is effective as an antimalarial agent by inhibiting biological processes by half.

The study of 1,8-naphthalimido-4-aminoquinoline derivatives can be carried out using QSAR method. Computational chemistry combines theoretically predicted data with laboratory experimental data. QSAR is known as structure-property relationship which is visualized as biological activity (Siswanta and Nugraha, 2017).

QSAR is used to design new drugs and to study computationally the relationship between activity and structure of compounds. The application of the QSAR method begins with geometry optimization or structural modeling to obtain the molecular descriptors, *i.e.* using the Hyperchem program package (Hamzah, Rauf and Anam, 2014). The computation of molecular descriptors allows computational prediction of biological activity (Armunanto and Sudiono, 2010). To the best of our knowledge, there is no research of 1,8-naphtalimido-4-aminoquinoline as antimalarial using QSAR. In this research the compound derivative of QSAR was tested using the semi-empirical Austin model 1 (AM1).

Research methods

Tools and materials

The software used was Hyperchem 8.0.5, BuildQSAR 2.1.0, Avogadro 1.2.0 and Notepad++ 8.1.2. All materials (data) used in the form of data set structure and antiplasmodial activity (IC50) from the results of research in the form of the basic compound 1,8-naphthalimide-4-aminoquinoline (Shalini *et al.*, 2020) including 25 derivative compounds of 1,8 - substituted naphthalimide-4-aminoquinoline. QSAR analysis was done using the multilinear regression method.

Procedure

Structure modeling and geometry optimization

The two-dimensional structure was added with H atoms and formed into a three-dimensional structure and then geometry was optimized using the AM1 semi-empirical method. The convergence limit is determined based on orientation observations, namely the gradient limit used is 0.1 kcal/(Å.mol) based on the Polak-Ribiere algorithm (Rakhman *et al.*, 2019).

Determination and calculation of descriptors

Calculation of descriptor using the AM-1 computational semiempirical method using Hyperchem 8.0.5 in the form of an electronic descriptor consisting of HOMO energy, LUMO energy, dipole moment and total atomic charge.

Multiple Linear Regression (MLR) Analysis

MLR analysis was carried out using the BuildQSAR software based on the genetic algorithm method that was available in the software to create an equation model. This process also includes a cross-validation step which excludes one from validating each accepted equation. The independent variables were hydrophobicity, steric and electronic parameters and the dependent variable was log IC50.

Data collection was divided into training sets and test sets. A training set of 20 compounds selected at random was used to develop the QSAR equation model and a test set of 5 compounds was used to validate the developed QSAR equation model.

The best validation of the equation model is the data by considering the values of R, R^2, and PRESS. The best equation model obtained can be used to predict the value of theoretical antimalarial activity for each compound 1,8-naphthalimido-4-aminoquinoline.

 | ISSN 2621-5985 (online) | ISSN 2549-385X (print)
Volume 5, Issue 2, 2022

Results and Discussion

In this study, 25 derivatives of 1,8-naphthalimido-4-aminoquinoline were geometrically optimized. Geometry optimization is a conformational transformation process of a compound until a conformation with the lowest potential energy is obtained. The geometry optimization process is carried out to obtain the compound structure in a steady state to minimize energy calculated using AM1. The semi-empirical approach AM1 (Austin Model 1) was developed in the 1980s as a revision of MNDO (Modified Neglect of Diatomic Overlap), resolving its main flaw, specifically the inability to form hydrogen bonds, without increasing computation time (Piekuś-Słomka, Zapadka and Kupcewicz, 2022). Several previous study used this level of theory in QSAR modelling had been published (Garcia *et al.*, 2022; Piekuś-Słomka, Zapadka and Kupcewicz, 2022).

The potential energy is lower due to the influence of the distance between the atoms making up the stable compound. The constituent atoms have a specific electronic charge, and the atoms are placed at a distance that minimizes interactions between atoms to prevent collisions of potential charges due to the induction of electrons from other constituent atoms. This is the main reason why optimized compounds have larger interatomic distances and lower energies than non-optimized compounds.

The compounds derived from 1,8-naphthalimide-4-aminoquinoline which have been geometrically optimized are calculated using the calculation of chemical-physical properties. This calculation is needed to get the descriptor value for each compound. The descriptors used in this study were Hansch parameters consisting of electronic parameters, steric parameters, and hydrophobic parameters. The log IC50 value data as the dependent variable and descriptors as independent variables, were processed by MLR analysis using BuildQSAR 2.1.0 software. In calculating the QSAR, the method used to find the best equation model uses a genetic algorithm that can be processed in terms of the highest correlation coefficient or F-test equation and the lowest standard deviation equation (De Oliveira and Gaudio, 2001). The genetic algorithm produces some of the best variable models from the data set that has been entered. Calculations using the genetic algorithm method resulted in the best 4 equation models. The four equation models can be seen in Table 1.

The equation model obtained by the BuildQSAR software was validated using a training set and a test set. The training set serves to develop the QSAR equation model that meets the statistical criteria values (R, R^2, and PRESS). A test set is carried out to confirm the QSAR equation model. Furthermore, 25 compounds derived from 1,8-naphthalimide-4-aminoquinoline were divided into 20 compounds for the training set and 5 compounds for the test set which were randomly selected.

In Table 1, each model has different predictions of antimalarial activity (log IC50), therefore it is necessary to choose the best equation that can predict log IC50 accurately. The selection of the best equation model begins by looking at the involvement of the Hansch parameters in the 4 equation models. Hansch equation put forward the concept that the relationship between chemical structure and biological activity of a compound can be expressed quantitatively through electronic, steric, and hydrophobic parameters (Yuliato, 2014). Based on this concept, it is known that each parameter has a relationship with one another in determining the biological activity of a compound. Table 1 shows that the four equation models have descriptors that represent the Hansch parameters.

 | ISSN 2621-5985 (online) | ISSN 2549-385X (print)
Volume 5, Issue 2, 2022

Table.1 MLR result set training descriptors and equation models

Model	Descriptor	Equation Model
1	qC2 qC14 Log P SAA	Log (IC50) = 1.3983 (± 0.7192) - 8.2044 (± 4.5362) qC2 + 3.4984 (± 2.3215) qC14 + 0.4657 (± 0.1191) log P + 0 .0019 (± 0.0012) SAA
2	qC2 qC14 Log P SAG	Log (IC50) = 1.2453 (± 0.8405) - 6.0202 (± 3.8202) qC2 + 2.9555 (± 2.2089) qC14 + 0.4310 (± 0.1126) Log P + 0.0017 (± 0.0011) SAG
3	qC4 qC14 Log P SAG	Log (IC50) = 1.1287 (± 0.8922) - 2.0067 (± 1.2983) qC4 + 3.2819 (± 2.3593) qC14 + 0.4455 (± 0.1180) Log P+ 0.0020 (± 0.0013) SAG
4	qC5 qC14 Log P SAG	Log (IC50) = 1.7016 (± 0.7413) + 4.3638 (± 2.8541) qC5 + 2.8868 (± 2.2245) qC14 + 0.4404 (± 0.1171) Log P + 0.0017 (± 0.0011) SAG

The analysis of the best equation model in this study begins with the statistical criteria of the correlation coefficient (R value). This measure describes the linearity of the equation model. The increase in the value of the dependent variable is proportional to the increase in the value of the independent variable when the R value is high. An acceptable value of R must be above 0.9 (Kesar *et al.*, 2019).

Table 2 R and R^2 values in the MLR training set equation model

Model	Descriptor				R	R^2
1	qC2	qC14	Log P	SAA	0.9140	0.8354
2	qC2	qC14	Log P	SAG	0.9100	0.8281
3	qC4	qC14	Log P	SAG	0.9090	0.8263
4	qC5	qC14	Log P	SAG	0.9080	0.8245

Table 2 shows that the four equation models have R values above 0.9. The descriptors (independent variables) in each equation model have a strong relationship with antimalarial activity (log IC50) as the dependent variable. The largest R value is found in equation 1 model was 0.9140. The R value proves that there is a strong relationship between the independent variable and the dependent variable.

The next statistical standard for the best equation model is the value of R^2. The R^2 value shows the percentage of antimalarial activity whose relationship can be explained through the descriptors in the study. The value of R^2 is more representative of the actual influence value. In MLR, when a new independent variable is entered into the equation model, the value of R^2 will experience an increase in value, even though the variable is not very influential.

The R^2 value has standard that must be met in order to get good results in the R^2 value of the training set and the R^2 value of the test set. The range of R^2 values in the training set must be above 0.8 while the test set must have a value above 0.6 (Kesar *et al.*, 2019). The R^2 value in Table 2 is the R^2 value of the training set. The R^2 value for all equation models above is 0.8. The largest R^2 value is found in the equation 1 model was 0.8354.

A test set test was carried out to validate the equation model, after it was known that the R^2 value of the training set in the 4 equation models had a value above 0.8. In Table 3, the four equation models have met the R^2 value above 0.6 which indicates that the equation model can be used as the best QSAR equation model.

Table 3. The value of R2 test set against the training set equation model

Model	R2
1	0.8965
2	0.9090
3	0.9002
4	0.9099

The R2 value in this study was calculated using the BuildQSAR 2.1.0 software which has undergone several adjustments based on the working method used. The two parameters, both R and R^2, are a measure of the linearity of an equation model, therefore it is necessary to consider other statistical criteria.

Each of the statistical criteria parameters used to analyze the four equation models, provides a different equation model as an equation model that meets the relevant parameter criteria. It is not enough to determine the equation model just by looking at the completion of the statistical parameter criteria.

The analysis needs to be continued by using the PRESS parameter to test the level of validation of the best equation model. The PRESS value is the sum of the squared differences between experimental activity (log IC50 experiment) and prediction activity (log IC50 prediction). If the PRESS value of an equation model is small, it means that the equation model obtained is excellent because the value is almost near to the experimental value.

Table 4. PRESS value in the MLR result equation model

Model	Descriptor				PRESS
1	qC2	qC14	Log P	SAA	0.253
2	qC2	qC14	Log P	SAG	0.257
3	qC4	qC14	Log P	SAG	0.268
4	qC5	qC14	Log P	SAG	0.263

In Table 4 the PRESS value is getting higher from model 1 to model 4. The determination of the best equation model is chosen based on the smallest PRESS value in an equation model. Equation 1 has the smallest PRESS value of 0.253.

After completing all the validation methods and statistical criteria parameters, equation 1 model was chosen as the best equation model. The model can be described as follows:

Log (IC50) = 1.3983 (± 0.7192) - 8.2044 (± 4.5362) qC2 + 3.4984 (± 2.3215) qC14 + 0.4657 (±0.1191) log P + 0.0019 (±0.0012) SAA with n = 20; R = 0.914; R2 = 0.8354; PRESS = 0.253. The research that has been done shows the results of equation 1 model as the best equation model. Furthermore, this equation model was used to calculate the antimalarial activity of the 1,8-naphthalmide-4-aminoquinoline derivative compound expressed by the predicted IC50 log.

In Table 5, it can be seen that the predicted log IC50 value data was calculated using the equation 1 model and compared with the IC50 log value of the experimental research results Shalini et al., (2020). The predicted log IC50 value close to the experimental log IC50 value. The predicted log IC50 will produce a good equation model. The distribution of the predicted IC50 log value and the experimental IC50 log value can be seen in the graph in Figure 1.

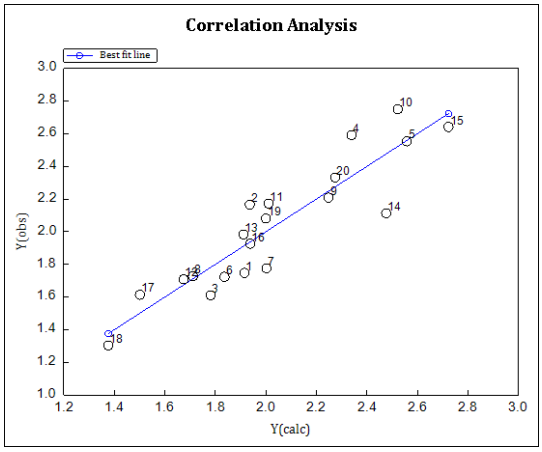

Figure 1. Correlation analysis between log IC50 prediction (Ycalc) and log IC50 experiment (Y(abs))

The value of a good prediction activity is equal to or close to the value of the experimental activity indicated by the points that are close to the x=y line. From the graph in Figure 1, the 20 compounds have points that are close to the x=y line which indicates the value of predictive activity and experimental activity is almost the same.

In this study the equation 1 model has several parameters consisting of electronic parameters represented by qC2 and qC14; steric parameters are represented by SAA; hydrophobic parameters are represented by log P. These parameters were used to predict the antimalarial activity of 1,8-naphthalmide-4-aminoquinoline derivatives.

In the structure of the derivative compounds of 1,8-naphthalimide-4-aminoquinoline has a quinoline flat ring which can be intercalated with the base pairs of the parasitic DNA double helix. In this process, a charge transfer complex occurs between the flat ring and the base pairs of guanine and cytosine, thus forming a drug-DNA complex. This complex is strengthened by the presence of side chain bonds with phosphate groups (bonds to the N atom at qC14) and their hydrogen bonds with DNA purine base molecules (adenine). N atom in the antimalarial compound which has a quinoline ring affects the mechanism action of the antimalarial compound. The N atom is one of the atoms that is predicted to affect the antimalarial activity of 1,8-naphthalimide-4-aminoquinoline derivatives (Block & Beale, 2011). Thus, the process of transcription and translation of the parasite's DNA is disrupted which causes a reduction in the synthesis of DNA and RNA of the malaria parasite. The net charge of the qC2 atom is estimated to affect antimalarial activity because it is close to the changing net charge of the qC1 atom where one of the H atoms acts as R.

Steric parameters also have an influence on the formation of bonds between drug compounds and receptors. In the formation of binding to the receptor, the drug compound must have a size and conformation to form an optimum drug-receptor compound bond. The selected steric parameter in the best equation model is SAA. SAA describes the approximate surface area of the molecule that interacts with the receptor. The wider the molecule, the better the interaction with the receptor.

In the best equation model, hydrophobic parameters such as log P are needed to predict the antimalarial activity of 1,8-napthalimide-4-aminoquinoline derivatives. In testing the antimalarial activity of the derivative compound 1,8-napthalimide-4-aminoquinoline. These compounds must be able to penetrate P. falciparum cells. Absorption into cells or penetration of cell membranes determines what percentage of 1,8-napthalimide-4-aminoquinoline derivatives enter and interact with receptors to kill *P. falciparum* cells. The absorption process requires an optimal log P value because log P is the partition coefficient of the compound in the aqueous phase and the lipid phase.

Conclusion and Recommendation

Conclusion

The best equation in the QSAR analysis of 1,8-naphthalimide-4-aminoquinoline derivatives using the AM1 semiempirical method is Log (IC50) = 1.3983 (± 0.7192) - 8.2044 (± 4.5362) qC2 + 3, 4984 (± 2.3215) qC14 + 0.4657 (± 0.1191) log P + 0.0019 (± 0.0012) SAA

The best descriptors of antimalarial activity were atomic net charges (qC2 and qC14) representing electronic parameters. Log P represents the hydrophobic parameter, and SAA represents the steric parameter.

Recommendation

It is necessary to model new antimalarial compounds based on the best equation model for the derivative compound 1,8-naphthalimide-4-aminoquinoline that has been obtained as a response to the increasingly worrying antimalarial drug resistance.

References

Armunanto, R. and Sudiono, S. (2010) 'Relation of Electronic Structures With Their Antimalarial Activities on Artemisinin Derivatives', *Indonesian Journal of Chemistry*, 4(3), pp. 212–217. doi: 10.22146/ijc.21856.

Block, J. H. and Beale, J. M. (2011) *Organic Medicinal and Pharmaceutical Chemistry*. 12th edn. Philadelphia: Lippincott Williams & Wilkins.

FDA (2020) *FDA cautions against use of hydroxychloroquine or chloroquine for COVID-19 outside of the hospital setting or a clinical trial due to risk of heart rhythm problems*.

Garcia, M. L. *et al.* (2022) 'QSAR studies on benzothiophene derivatives as Plasmodium falciparum N-myristoyltransferase inhibitors: Molecular insights into affinity and selectivity', *Drug Development Research*, 83(2), pp. 264–284. doi: 10.1002/ddr.21646.

Hamzah, N., Rauf, A. and Anam, A. (2014) 'Studi Hubungan Kuantitatif Struktur-Aktivitas (HKSA) dan Docking Molekuler Senyawa Turunan Oxabicycloheptene Sulfonamide (OBHS) sebagai Antagonis Reseptor Estrogen-α pada Terapi Kanker Leher Rahim (Serviks)', 2, pp. 39–45.

Kesar, S. *et al.* (2019) 'Quantitative Structure-Activity Relationship Analysis of Selective Rho Kinase Inhibitors as Neuro-regenerator Agents', *Turkish Journal of Pharmaceutical Sciences*, 16(2), pp. 141–154. doi: 10.4274/tjps.galenos.2018.70288.

MAG (2018) *World Malaria Day 2018*.

De Oliveira, D. B. and Gaudio, A. C. (2001) 'BuildQSAR: A new computer program for QSAR analysis', *Quantitative Structure-Activity Relationships*, 19(6), pp. 599–601. doi: 10.1002/1521-3838(200012)19:6<599::AID-QSAR599>3.0.CO;2-B.

Piekuś-Słomka, N., Zapadka, M. and Kupcewicz, B. (2022) 'Methoxy and methylthio-substituted trans-stilbene derivatives as CYP1B1 inhibitors – QSAR study with detailed interpretation of molecular descriptors', *Arabian Journal of Chemistry*, 15(11). doi: 10.1016/j.arabjc.2022.104204.

Rakhman, K. A. *et al.* (2019) 'Kajian Senyawa Turunan Benzopirazin sebagai Antimalaria Menggunakan Metode HKSA dan MLR', *EduChemia (Jurnal Kimia dan Pendidikan)*, 4(2), p. 112. doi: 10.30870/educhemia.v4i2.4989.

Shalini *et al.* (2020) 'Functionalized Naphthalimide-4-aminoquinoline Conjugates as Promising Antiplasmodials, with Mechanistic Insights', *ACS Medicinal Chemistry Letters*, 11(2), pp. 154–161. doi: 10.1021/acsmedchemlett.9b00521.

Siswanta, D. and Nugraha, G. (2017) 'Pemodelan dan Analisis QSAR Turunan Metilkarbamat sebagai Insektisida menggunakan Metode Semiempirik Austin Model 1', 1, pp. 43–49.

 | ISSN 2621-5985 (online) | ISSN 2549-385X (print)
Volume 5, Issue 2, 2022

Swinney, D. C. (2011) *Molecular Mechanism of Action (MMoA) in Drug Discovery*. 1st edn, *Annual Reports in Medicinal Chemistry*. 1st edn. Elsevier Inc. doi: 10.1016/B978-0-12-386009-5.00009-6.

Tilley, L., Loria, P. and Foley, M. (2003) 'Chloroquine and Other Quinoline Antimalarials', *Antimalarial Chemotherapy*, pp. 87–121. doi: 10.1385/1-59259-111-6:87.

WHO (2021) *World Malaria Report 2021*, *Angewandte Chemie International Edition, 6(11), 951–952.*

Yulianto, T. (2014) 'Hubungan Kuantitatif Struktur dan Aktivitas Antikanker Senyawa Turunan Bakteriofeoforbid A', *thesis*.