



Quantitative Structure-Activity Relationship Analysis of Antimalarial Compound of Mangostin Derivatives Using Regression Linear Approach

RUSLIN HADANU^{1,*} and SYAMSUDIN^{2,*}

¹Department of Chemistry, Faculty of Teachership and Educational Science, Pattimura University, Poka, Ambon, Indonesia

²Department of Pharmacology, Faculty of Pharmacy, Pancasila University, Jakarta, Indonesia

*Corresponding authors: E-mail: ruslin_hadanu@yahoo.com; syamsudin.abdillah@yahoo.co.id

(Received: 11 July 2012;

Accepted: 30 April 2013)

AJC-13420

Quantitative electronic structure-activity relationship (QSAR) analysis of a series of mangostin derivatives as antiplasmodial compounds have been conducted using atomic net charges (q), dipole moment (μ) E_{LUMO} , E_{HOMO} , polarizability (α), $\log P$ and massa molecular as the descriptors. The descriptors were obtained from computational chemistry method using semi-empirical PM3. Antiplasmodial activities were taken as the activity of the drugs against chloroquine-resistant plasmodiumfalciparum FCR3 strain and are presented as the value of $\ln(1/IC_{50})$ where IC_{50} is an effective concentration inhibiting 50 % of the parasite growth. The best model of QSAR model was determined by multiple linear regression method and giving equation of QSAR: $\ln 1/IC_{50} = -189.029 - (752.054) qC_5 + (1249.672) qO_6 + (3134.359) qC_9 + (1174.323) qC_{10} + (329.121) qC_{13} - (1.409) \mu + (75.214) E_{LUMO} - (23.683) E_{LUMO} - (2.156) \alpha - (3.372) \log P + (0.309) MM$. The equation was significant on the 95 % level with statistical parameters: $n = 15$; $r = 0.951$; $r^2 = 0.905$; $SE = 1.326498$; $F_{cal}/F_{table} = 1.029$ and gave the PRESS = 5.650972. Its means that there were only a relatively few deviations between the experimental and theoretical data of antimalarial activity.

Key Words: QSAR analysis, Antimalarial activity, Semi-empirical method, Mangostin.

INTRODUCTION

Malaria remains one of the most devastating diseases, causing as many as 2.7 million deaths annually with an estimated 400 to 900 million new cases each year¹. Malaria endemic areas include Africa, South east asia, India and South America. However, the disease is spreading to new areas, such as Central Asia and Eastern Europe. Local transmission of malaria in the United States, unheard of in the era between World War II and 1980, now accounts for an increasing number of cases². Clinical cases in the US now average 1,300 per year³. Worldwide, the majority of deaths occur in children; other high risk groups include pregnant women, refugees, migrant workers and non immune travelers-over 20 million Western tourists at risk annually (fact sheets from Malaria Foundation International). Although four species of the genus *Plasmodium* cause human malaria, *Plasmodium falciparum* is the deadliest and will be the subject of this review.

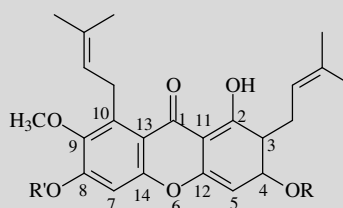
During the past two decades an increasing number of quantitative structure-activity/property relationship (QSAR/QSPR) models have been studied using theoretical molecular descriptors for predicting biomedical, activity, toxicological and technological properties of chemicals. QSAR/QSPR are mathematical models that seek to predict complicated physico-

chemical/biological properties of chemicals from their simpler experimental or calculated properties. The main problem with the use of experimental data as independent variables in QSAR is that they are not available for the majority of chemical structures, real or hypothetical.

The traditional remedies are no longer effective and the incidence of malarial by *P. falciparum*, the most dangerous species of parasite, continues to grow, while some traditional drugs such as chloroquine and its congeners are losing their activity due to the increasing multi drug resistance⁴. Therefore, it is essential to find new drugs of antimalarial having a pharmacological activity higher than that of currently available drugs of antimalarial. In this connection, quantitative structure-activity relationship (QSAR) analysis plays an important role to minimize trial and error in designing new antimalarial drugs.

QSAR studies of antimalarial activity represent an emerging and exceptionally important topic in the area of computed-aided drug design. Although the demand for 'in silico' discovery is clear in all areas of human therapeutics, the field of anti-infective drugs has a particular need for computational solutions enabling rapid identification of novel therapeutic leads.

TABLE-1
CHEMICAL STRUCTURE AND ACTIVITY DATA OF ANTIMALARIAL COMPOUNDS
OF MANGOSTIN DERIVATIVES AGAINST FCR3 STRAIN



No	R'	R	IC ₅₀ (μM)
1.	H	H	17 ± 1,0
2.	-(CH ₂) ₃ OH	H	5.5 ± 0.3
3.	H	-CH ₂ CH(OH)CH ₂ OH	7.4 ± 0.3
4.	-CH ₂ CH(OH)CH ₂ OH	H	16.3 ± 1.0
5.	H	CH ₂ CONH ₂	4.5 ± 0.3
6.	CH ₂ CONH ₂	H	8.3 ± 0.4
7.	CH ₂ CH ₂ CH ₂ CN	CH ₂ CH ₂ CH ₂ CN	18.0 ± 1.0
8.	CH ₂ CH ₂ CH ₂ CN	H	21.0 ± 2.0
9.	CH ₂ CH ₂ N(CH ₂ CH ₃) ₂	H	0.3 ± 0.02
10.	CH ₂ CH ₂ N(CH ₂ CH ₃) ₂	CH ₂ CH ₂ N(CH ₂ CH ₃) ₂	1.5 ± 0.1
11.	CH ₂ CH ₂ N(CH ₃) ₂	H	0.6 ± 0.3
12.	CH ₂ CH ₂ CH ₂ N(CH ₃) ₂	H	0.1 ± 0.1
13.	CH ₂ CH(OH)CH ₂ N(CH ₂ CH ₃) ₂	H	0.05 ± 0.005
14.	CH ₂ CH(OH)CH ₂ N(CH ₃) ₂	H	0.6 ± 0.03
15.	CH ₂ CH(OH)CH ₂ N(CH ₃) ₂	CH ₂ CH(OH)CH ₂ N(CH ₃) ₂	0.6 ± 0.03

In this research semiempirical methods were used to calculate a number of properties/descriptors. In current practice, semiempirical methods serve as an efficient computational tools which can yield fast quantitative estimates for a number of properties. This may be particularly useful for correlating large sets of experimental and theoretical data, for establishing trends in classes of related molecules and for scanning a computational problem before proceeding with higher-level treatments. Compared with *ab initio* or density functional methods, semiempirical calculations are much faster, typically by several orders of magnitude⁵, but they are also less accurate, with errors that are less systematic and thus harder to correct.

Among the classes of drugs that are effective in the treatment of the *P. falciparum* malarial, there is the mangostin and its derivatives. The mangostinderivatives was synthesized from mangostin as a starting material. The first models of QSAR were modeled, described and evaluated for their antiplasmodial activity. From that evaluation, the mangostin ring system appeared as a new class of potential antimalarial compounds⁶. Previous QSAR^{7,8} analysis have been reported using electronic descriptor producing by PM3 calculation using 15 mangostin derivatives compound.

EXPERIMENTAL

Geometry optimization and calculation of molecular descriptors: QSAR models are evaluated using sets of mangostin derivatives compounds whose molecular structure and antiplasmodial activity are known (Table-1). Antiplasmodial activity of these compound were taken as the activity against chloroquine-resistant *P. falciparum* (FCR3) strain and is presented as the value of $\ln(1/IC_{50})$ where IC_{50} is an effective concentration inhibiting 50 % growth of the parasite⁶. All the compounds (Table-1) were calculated using package HyperChem® Program Version 7.0 and complete geometry

optimization with the semi-empirical Parameterized Model 3 (PM3) method was performed. Quantum-chemical descriptors were calculated, as for example: atomic net charges, dipole moment, E_{HOMO} , E_{LUMO} , polarizability and log P. From all the descriptors above mentioned, it can be considered that some of them give valuable information about the influence of electronic and coefficient partition features upon the biological activity of drug molecules. In this work, the molecular descriptors were selected so that they represent the features necessary to quantify the activity.

QSAR evaluation: The steps of quantitative structure-activity relationship analysis were conducted at this research: to determinate a series mangostin derivatives compound to be analyzed along with value of IC_{50} yielded through laboratory experiment; to look for elementary ring structure of mangostin derivatives compound, which is the most stable using optimization process; to calculate descriptor through optimized structure; statistical analysis process using SPSS 16 for Windows to get equation of QSAR; selected of the best equation model of QSAR; to design the new compound of mangostin derivatives according to the best model of QSAR and consider of eligibility of synthesis; and making of mangostin derivatives resulted from molecular design which has the highest predicted value of IC_{50} to be synthesized in laboratory. The net charge atoms for three dimension of molecule structure which have experienced of converged at process of optimization can be recorded through menu file; start log; compute; single point and then stop log. The net charge of atoms on its structural compound can be seen through display menu, labels and charge.

RESULTS AND DISCUSSION

Descriptor is a parameter or property of molecule used as independent variables in calculation of predicted activity (theoretical IC_{50}). The descriptors used in this research are atomic

TABLE-2
DESCRIPTORS/INDEPENDENT VARIABLES USED FOR QSAR ANALYSIS OF ANTIMALARIAL COMPOUNDS
OF MANGOSTIN DERIVATIVES CALCULATED BY SEMI-EMPIRICAL PM3 METHOD

Comp. number	Atomic net charges (Coulomb)									
	qC1	qC2	qC3	qC4	qC5	qO6	qC7	qC8	qC9	qC10
1	0.340193	0.202540	-0.221100	0.172785	-0.289721	-0.135351	-0.245510	0.110548	-0.017694	0.051572
2	-0.002760	-0.021200	0.003040	0.003170	-0.016670	-0.000350	0.067420	-0.038470	-0.031330	0.072850
3	0.340124	0.199445	-0.222872	0.176900	-0.278653	-0.134735	-0.245696	0.110721	-0.018135	0.051455
4	0.340364	0.202452	-0.221340	0.172927	-0.289477	-0.134794	-0.235419	0.116544	-0.020908	0.049194
5	0.339661	0.200832	-0.221759	0.174548	-0.276875	-0.133801	-0.244890	0.111651	-0.016952	0.051691
6	0.339745	0.203139	-0.220479	0.174005	-0.289014	-0.133868	-0.231524	0.111648	-0.019202	0.050554
7	0.418395	0.237821	-0.259518	0.197694	-0.290147	-0.110965	-0.239326	0.137691	-0.022515	0.047937
8	0.419589	0.240357	-0.265139	0.205581	-0.297887	-0.110075	-0.240966	0.142752	-0.026964	0.046791
9	0.419092	0.240474	-0.264711	0.205624	-0.297884	-0.110583	-0.242968	0.131383	-0.023244	0.046326
10	0.419281	0.236726	-0.264048	0.202662	-0.291077	-0.110174	-0.243298	0.131122	-0.023719	0.046211
11	0.419475	0.240432	-0.265113	0.205636	-0.297875	-0.109883	-0.244999	0.141905	-0.025907	0.047275
12	0.419354	0.240517	-0.265238	0.205591	-0.298265	-0.109830	-0.244181	0.140539	-0.024781	0.046967
13	0.419107	0.240617	-0.264460	0.206044	-0.29731	-0.109371	-0.241865	0.138166	-0.026188	0.048572
14	0.419100	0.240623	-0.264467	0.206049	-0.297316	-0.109364	-0.241866	0.138134	-0.026138	0.048565
15	0.419782	0.231581	-0.262973	0.199012	-0.269500	-0.105925	-0.225019	0.136019	-0.025124	0.041539
Comp. number	Atomic net charges (Coulomb)				α (Debyes)	E_{LUMO} (eV)	E_{HOMO} (eV)	α (\AA^3)	Log P	MM
	qC11	qC12	qC13	qC14						
1	-0.261254	0.172412	-0.199533	0.127437	3.732	-0.592052	-8.980277	44.230	-1.020	410.470
2	0.022670	0.012590	-0.045230	-0.028310	5.177	-0.564386	-8.943869	50.370	-1.380	468.550
3	-0.259296	0.168967	-0.199960	0.127806	3.876	-0.580620	-8.957765	51.010	-1.800	484.550
4	-0.261491	0.172518	-0.199317	0.125598	2.741	-0.580620	-8.957765	51.010	-1.800	484.550
5	-0.256465	0.170649	-0.199976	0.126830	2.071	-0.706499	-9.067646	49.330	-2.310	467.520
6	-0.262182	0.171874	-0.195605	0.126734	2.565	-0.742620	-9.113591	49.330	-2.310	467.520
7	-0.292034	0.187065	-0.231605	0.139147	3.396	-0.714321	-9.109834	58.940	0.310	544.650
8	-0.296739	0.189507	-0.235027	0.140013	3.286	-0.521657	-8.952111	52.920	-1.270	481.590
9	-0.296684	0.189240	-0.232855	0.139952	4.027	-0.557141	-8.893980	56.590	-0.320	509.640
10	-0.294489	0.187114	-0.233013	0.140369	4.448	-0.525456	-8.951884	68.950	0.370	608.820
11	-0.296835	0.189447	-0.234663	0.142133	3.612	-0.530195	-8.959011	52.920	-1.010	481.590
12	-0.296879	0.189569	-0.234145	0.141615	4.259	-0.530123	-8.957450	54.750	-0.950	495.620
13	-0.296921	0.188905	-0.233352	0.143046	3.823	-0.591736	-9.022517	59.060	-0.690	539.670
14	-0.296926	0.188911	-0.233332	0.143032	3.823	-0.591881	-9.022533	59.060	-0.690	539.670
15	-0.292390	0.184108	-0.233280	0.139121	7.199	-0.294656	-8.721663	66.550	-1.740	612.760

TABLE-3
FIVE SELECTED MODELS AND THEIR STATISTICAL PARAMETERS FOR THE CORRELATION BETWEEN
MOLECULAR PROPERTIES AND ANTIMALARIAL ACTIVITY OF MANGOSTIN DERIVATIVES

QSAR model	Variables	R	R ²	F _{calc} /F _{table (0.05)}	SE	PRESS
1.	qC5, qO6, qC9, qC10, qC13, μ , E_{LUMO} , E_{HOMO} , α , log P, MM	0.951	0.905	1.029	1.326498	5.650972
2.	qC5, qO6, qC9, qC10, qC13, μ , E_{LUMO} , α , log P, MM	0.941	0.886	1.219	1.257674	6.370408
3.	qC5, qO6, qC9, qC10, qC13, E_{LUMO} , α , log P, MM	0.938	0.879	1.557	1.157364	6.756751
4.	qC5, qO6, qC9, qC10, qC13, E_{LUMO} , log P, MM	0.904	0.818	1.272	1.297023	10.099475
5.	qC5, qO6, qC9, qC10, E_{LUMO} , log P, MM	0.874	0.763	1.189	1.367755	13.619751

net charges, dipole moment, log P, $E_{HOMO-LUMO}$ and polarizability. To obtain the structural properties of each test compound and modeling compound after process of geometry optimization, the calculation process is continued with single point at sub menu. Descriptors or structural properties yielded from calculation with single point were atomic net charges and dipole moment, whereas descriptors of log P and polarizability were obtained from menu compute of QSAR properties. E_{HOMO} , E_{LUMO} descriptors can be obtained from the menu compute, vibrations then click orbital sub menu. The $E_{HOMO-LUMO}$ descriptors would not be obtained if calculation was performed using molecular mechanic method. All of the descriptors was given in Table-2.

To calculate of structural electronic and other descriptors of a series mangostin derivatives was conducted with the semi

empiric PM3 method. PM3 method which is better method than the others semi empiric method. Method of PM3 is repair method of before all like MNDO method¹¹, which can predicts compounds having valence many with the best accuracies¹². The PM3 method can be used for the analysis of a series mangostin derivatives, because mangostin derivatives is organic compound considering atoms as C, H, Cl, Br, O and N.

Selection of the best model: According to result of calculation statistic of multilinear regression by using SPSS version 16.0 for windows were obtained 5 QSAR model as listed in Tables 3 and 4. From 5 equation models were determined 1 of the best QSAR model. Model 1 is selected as the best model among 5 models, based on: 1. The value of r and r² to look for analysis data linearity that the model 1 having r equal to 1 is 0.951 and r² is 0.905. 2. The smallest value of SE

TABLE-4
COEFFICIENT OF SELECTED INDEPENDENT VARIABLES FOR 5 QSAR MODELS
AS OBTAINED FROM MULTILINEAR REGRESSION ANALYSIS

QSAR models	Coefficient of independent variables										Constants	
	qC5	qO6	qC9	qC10	qC13	μ	E_{LUMO}	E_{HOMO}	α	Log P		MM
1.	-752.054	1249.672	3134.359	1174.323	329.121	-1.409	75.214	-23.683	-2.156	-3.372	0.309	-189.029
2.	-593.934	952.123	2256.933	1234.837	222.380	-0.533	38.741		-1.403	-3.228	0.215	-30.407
3.	-547.478	861.933	2003.267	1195.016	199.474		31.558		-1.129	-3.212	0.179	-39.487
4.	-398.327	631.892	1503.264	1003.314	125.758		22.046			-3.310	0.045	-41.615
5.	-264.147	465.266	1199.958	1092.323			18.652			-2.909	0.035	-58.038

(standard error of estimation) is model 1 having value of SE equal to 1.326498. 3. If the value of F exceed value of F_{table} or comparison of F_{calc}/F_{table} more than 1. The model 1 have value of F_{calc}/F_{table} is 1.029. 4. The 1 model QSAR have smaller value of PRESS than another QSAR model.

The value of $r = 0.951$ and $r^2 = 0.905$ to indicate that correlation between electronic structure (independent variables) with antimalarial activity. It means that the change of activity of antimalarial $\ln(1/IC_{50})$ a series mangostin derivatives compound resulted 95.1 % from the existence of change of descriptor: electronic structure, dipole moment, E_{LUMO} , E_{HOMO} , log P; polarizability and molecular massa, that all are independent variables. Comparing parameters F and SE of the six models, it is easy revealed that model 1 is the best model because it highest F and lowest SE value. According to F value indicated that model 1 is significance at trust level 95 % as shown by ratio of F_{calc}/F_{table} which the value more than 1. The value of F_{calc} larger than F_{table} to indicated that H_1 accepted and its showing correlates electronic structure (dependent variables) a series mangostin derivatives between activity of antimalarial ($\ln(1/IC_{50})$) having significance relation at trust level 95 %. Meanwhile, the smallest value of SE to indicate that the QSAR model have very small deviation of data or have highest significance data. At Table 5 indicating that model 1 have smallest value of PRESS than another models. The small value of PRESS to indicate antimalarial activity of experiment between activity of predict have very small difference value. The mentioned can be made guide that the model 1 have more good ability activity of antimalarial mangostin derivatives to design. The model 1 is presented at Tables 3 and 4, completely can be write at following.

$$\ln(1/IC_{50}) = -189.029 - (752.054)qC_5 + (1249.672)qO_6 + (3134.359)qC_9 + (1174.323)qC_{10} + (329.121)qC_{13} - (1.409)\mu + (75.214)E_{LUMO} - (23.683)E_{HOMO} - (2.156)\alpha - (3.372)\log P + (0.309)MM \quad (1)$$

$n = 15$; $r = 0.951$; $r^2 = 0.905$; $SE = 1.326498$; $F_{calc}/F_{table} = 1.029$; $PRESS = 5.650972$.

The QSAR model obtained is ideal if its has r^2 value equal to and its indicating that correlation independent variable between dependent variables is perfect and significance¹⁰ if its has r^2 value equal to 0 indicating structural electronic and properties of molecular (dependent variables) between antimalarial activity $\ln(1/IC_{50})$ have no correlation or no significance. The statistical parameters commonly using r^2 value because have more correctness level than r value. The value of r^2 have larger interval than r value so that small difference which no perceived at r value but its can perceived clearly at r^2 . The r and r^2 value as statistical parameters only showing

linearity measures of relevant model, but cannot depict measure of predicts of equation model, so that require to be paid attention by other statistical parameters.

The others statistical parameters, beside r and r^2 which need to be paid attention in this research is SE and F value. The smallest value of SE to express the model obtained is progressively and more significance. At Table-3 showing the model 1 was selected the most significance at trust level 95 %. Ones statistical parameters to look for ability of QSAR model is to be analyze of PRESS parameter. The smallest value of PRESS to indicate the QSAR model have good ability to predict antimalaria activity. The PRESS value of QSAR models was listed in Tables 3 and 4. The model 1 is the most reliable model because its has the smallest value.

The result of evaluation of antimalarial activity [predicted $\ln(1/IC_{50})$] and correlation with antimalarial activity [experiment $\ln(1/IC_{50})$] for the model 1 by using semi empiric PM3 method have linearity ($r^2 = 0.9748$) and slope value (0.974) can be seen at Fig. 1.

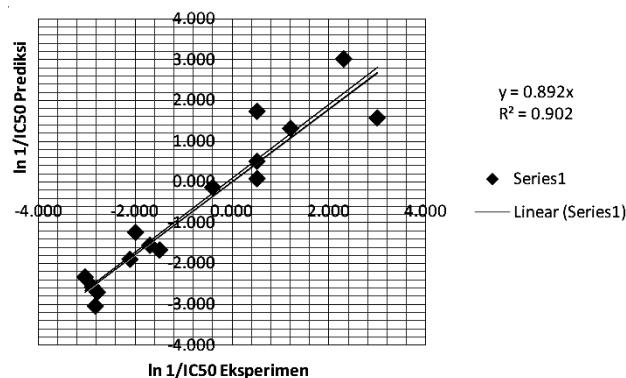


Fig. 1. Linear regression of experimentally observed antimalarial activity $\ln(1/IC_{50})$ versus calculated one based on QSAR model 5

According to the value of variable dipole moment, polarizability, molecular mass, E_{LUMO} ; E_{HOMO} and log P were obtained by the variation of atoms included in multilinear analysis (Table-4), atoms C5, O6, C9, C10 and C13 seems the most responsible for the pharmacological activity.

Conclusion:

We have used a semi-empirical molecular calculation PM3 to study the correlation of antimalarial activity of a series of mangostin derivatives drugs against chloroquine-resistant FCR3 strain. The best overall correlation is given by the computed molecular properties of atomic net charges of heterocyclic ring, dipole moment, LUMO-HOMO energies, polarizability, molecular mass and log P.

Significant regression model was obtained by multiple linear regression method for structural properties of mangostin derivatives *versus* antimalarial activity against *P. falciparum*. We reported the descriptors dipole moment, polarizability, molecular mass, E_{LUMO} ; E_{HOMO} and log P, atomic net charges: C_5 , O_6 , C_9 , C_{10} and C_{13} seems to be the most responsible for the pharmacological activity. The model is significant on the 95 % level with statistical parameters.

REFERENCES

1. J.G. Breman, *Am. J. Tropic Med. Hyg.*, **64**, 1 (2001).
2. S.B. Mundy, A.C. White, Jr., J.S. Hines, B.J. Marino and E.J. Young, *Southern Med. J.*, **89**, 616 (1996).
3. S. Filler, L.M. Causer, R.D. Newman, A.M. Barber, J.M. Roberts, J. MacArthur, M.E. Parise and R.W. Steketee, *MMWR surveillance summaries*, **52**, 1 (2003).
4. W.H. Wernsdorfer, *Parasitol. Today*, **7**, 297 (1991).
5. W. Thiel, *Adv. Chem. Phys.*, **93**, 703 (1996).
6. W. Mahabusarakam, K. Kuaha, P. Wilairat and W.C. Taylor, *Chem. Planta Med.*, **72**, 912 (2006).
7. Mustofa, I. Tahir and Jumina, *Indo. J Chem.*, **2**, 91 (2002).
8. Mustofa, A.D. Yappi, A. Valentin and I. dan Tahir, *Berkala Ilmu Kedokteran*, **35**, 67 (2003).
9. M.A. Wijayanti, E.N. Sholikhah, I. Tahir, R. Hadanu, Jumina, Supargiyono and Mustofa, *J. Health Sci.*, **52**, 794 (2006).
10. A.R. Leach, *Molecular Modelling, Principles and Applications*, Longman, Singapore (1996).
11. M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).