Quantitative Structure-Activity Relationship Analysis (QSAR) of Antimalarial Vincadifformine Analogues Compounds

Mudasir*[,1], Iqmal Tahir[1], Irza Yulistia[1], Mustofa[2] and Bernd M. Rode[3]

[1]*Austrian-Indonesian Centre for Computational Chemistry, Chemistry Department, Gadjah Mada University, Jogjakarta 55281, Indonesia.*
[2]*Faculty of Medicine, Gadjah Mada University, Jogjakarta 55281, Indonesia.*
[3]*Institute for General, Inorganic and Theoretical Chemistry, Insbruck University, Innrain 52a, A-6020 Innsbruck, Austria*

Abstract

Quantitative Structure-Activity Relationship (QSAR) analysis of vincadifformine analogs as antimalarial drug has been conducted using atomic net charges (q), moment dipole (μ), LUMO (Lowest Unoccupied Molecular Orbital) and HOMO (Highest Occupied Molecular Orbital) energies, molecular mass (m) as well as surface area (A) as predictors of their activity. Data of predictors are obtained from computational chemistry method using semi-empirical molecular orbital AM1 calculation. Antimalarial activities are taken as the activity of the drugs against chloroquine-resistant plasmodium falciparum (FcM29-Cameroun) 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 QSAR model has been determined by multiple linier regression analysis giving QSAR equation:

$$ln(1/IC_{50}) = 20.055.qC_1 - 16.109.qC_2 - 23.745.qC_5 - 10.208.qC_6 + 6.582.qC_7 - 9.758.qN + 4.076.qC_8 - 0.176.\ \mu - 1.087.E_{LUMO} + 0.664.E_{HOMO} - 0.008.m + 0.021.A - 5.168$$

$$(n=16;\ r^2 = 0.998;\ SE = 0.07;\ F_{calc}/F_{table} = 4.612;\ sig = 0.006)$$

Keywords*: QSAR analysis, Quantum chemical parameter, Antimalaria, Vincadifformine*

Introduction

Malaria has been known since ancient times and today it is one of the major causes of morbidity and death in the world. Although exact figures are not readily available, it is estimated that 2.5 billion people are at risk [1]. Every year 300-500 million new malaria infections arise, and 1-3 million deaths causes by the disease, most of them are children [2,3]. In Indonesia, the endemic area of malaria covers almost every province, especially in eastern regions of Papua and Maluku provinces. It is estimated that some 15 million peoples of Indonesia are infected and around 30 thousands of them are killed by the disease annually [4].

The prevention and cure of malaria relies on a limited number of drugs. Apart from the natural product, quinine, the main synthetic drugs currently in use are quinoline derivatives such as primaquine, chloroquine and mefloquine. Other products are simple sulfonamide and pyrimidine derivatives, which find use as antimalarial drugs by inhibiting the synthesis and reduction of dihydrofolate in the parasite [5].

The traditional remedies are no longer effective and the incidence of malaria 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 increasing multidrug resistence [6]. Therefore, it is essential to find new drugs of antimalaria having a pharmacological activity higher than that of currently available drugs of antimalaria. In this connection, quantitative structure-activity relationship (QSAR) analysis plays an important role to minimize *trial and error* in designing new antimalarial drugs.

Numerous studies on QSAR of antimalarial drugs have been performed in the past two decades based on quantum chemical parameters. Most of them deal with quinoline derivatives such as primaquine [7], chloroquine [8] and mefloquine [9] as well as 1,2,4-trioxanes [1], but no QSAR analysis, so far, has been reported on antimalarial vincadifformine. We now report a QSAR study on vincadifformine analogs based on quantum chemically calculated molecular properties with the aim of elucidating linear correlation model between antimalarial activity and molecular properties as well as obtaining information about the active center of the drug.

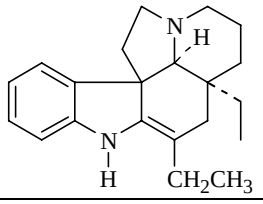
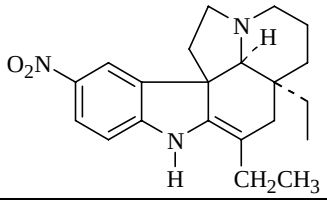
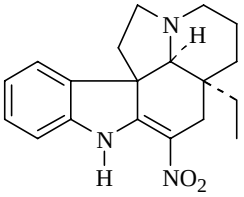
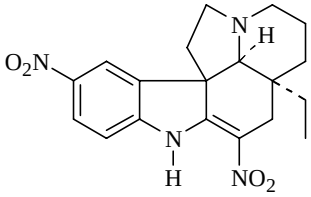
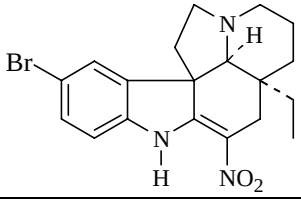
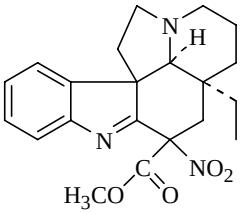
Experimental

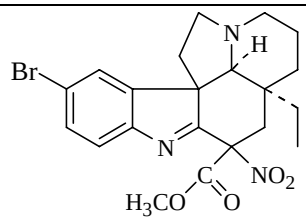
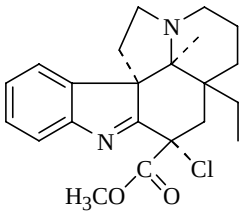
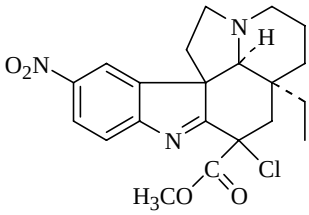
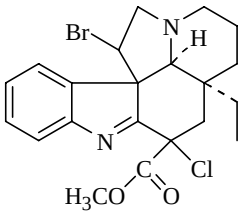
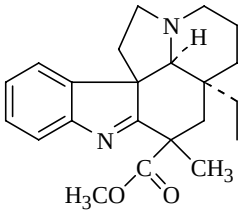
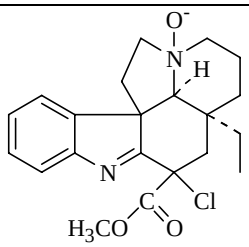
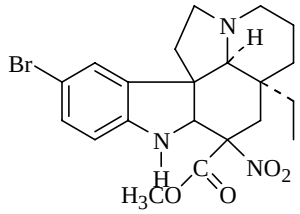
Vincadifformine compounds were extracted from *A. Pyrifolium* and *A. Megalocarpon* plants

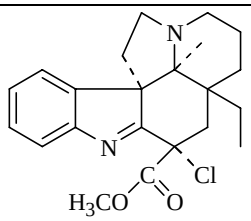
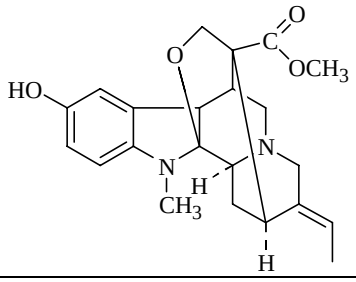
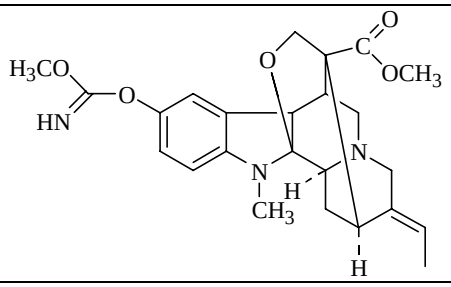
which is traditionally used as antimalaria in America. Antiplasmodial activity of these compounds were taken as the activity against chloroquine-resistant plasmodium falciparum (FcM29-Cameroun) strain and is presented as the value of $\ln(1/IC_{50})$ where IC_{50} is an effective concentration inhibiting 50% growth of the parasite [10]. The molecular backbone of vincadifformine-type drugs is illustrated in Fig. 1. Sixteen pharmacologically well-characterized compounds have been considered for our investigation, whose activity data have been previously determined [10]. Chemical derivatives and their antimalarial activities are presented in Table 1.

The variables for multilinear regression analysis consisting of atomic net-charge (q), moment dipole ( ), LUMO (E_{LUMO}) and HOMO (E_{HOMO}) energies, polarisability ( ), molecular mass (m) as well as surface area (A) were calculated by semi-empirical AM-1 MO SCF method using HyperChem Version 6.0.

Table 1 Chemical structure and activity data of antimalarial vincadifformine analogs against FcM29 – cameroun strain

Compound no	Chemical Structure	ln 1/C
1		-1. 3856
2		-1. 4548
3		-1.4362
4		-1.1072
5		-1.2148
6		-2.0187

Compound no	Chemical Structure	ln 1/C
7		-1.9518
8		-0.8921
9		-1.0828
10		-1.2788
11		-1.3096
12		-1.4062
13		-1.6839

Compound no	Chemical Structure	ln 1/C
14		-1.9731
15		-1.0253
16		-1.6232

Before calculation of predictors was done, the geometries of the drug molecules were optimized on the basis of Polar-Ribiere algorithm with convergence limit of 0.001 kcal/mol Å. The correlation models between predictors and pharmacological activity were evaluated by a multilinear regression analysis software SPSS for Windows™ using *enter* and *backward* method according to the following equation:

$$\ln (1/IC_{50}) = \sum P_{(qi)}q_{(i)} + P_{(\mu)}\mu + P_{(E\text{-LUMO})}E_{LUMO} + P_{(E\text{-HOMO})}E_{HOMO} + P_{(\alpha)}\alpha + P_{(m)}m + P_{(A)}A + D \quad (1)$$

where P is a fitting coefficient for corresponding variables and D is a constant.

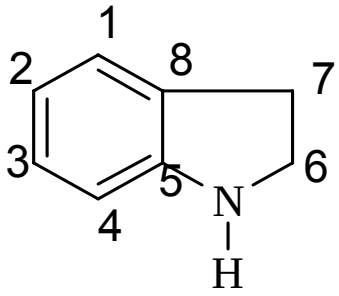


Fig.1 Molecular backbone of antimalarial vincadifformine analogs. Atomic numbering is not a standard one and it is used only for the purpose of modeling.

Results and Discussion

The calculated AM-1 descriptors of vincadifformine-type antimalarial drugs used in model fitting are listed in Table 2. In searching for suitable models according to Eq. (1), the relative importance of descriptors, i. e.: atomic net-charge and other properties can be recognized from the variable coefficient size (*P*) and from the result of inter variable correlation analysis by bivariate method. This allows the exclusion of less relevant descriptor and gradual evaluation of the structure of the active center of the drug.

Selection of the best model

To obtain the best model that correlates independence variables (descriptors) and dependence variable (biological activity), mulitilinear regression analysis using SPSS version 10 for windows has been performed. 15 independent variables consisting of 9 atomic net-charges: qC_1, qC_2, qC_3, qC_4, qC_5, qC_6, qC_7, qC_N and qC_8 as well as other properties such as dipole moment, LUMO and HOMO energies, polarisability, molecular mass and surface area were included in the model set-up. At the first step, all variables are included in the model and the less relevant variables were then eliminated gradually from the model by *enter* and *backward* method.

Table 2 Descriptors/independent variables used for QSAR analysis of antimalarial vincadifformine analogs calculated by semi-empirical AM1 method

Compound no	Atomic net charges									Dipole moment	E_{LUMO}	E_{LUMO}	Polaris-ability	Molecular mass	Molecular surface
	qC_1	qC_2	qC_3	qC_4	qC_5	qC_6	qC_7	qN	qC_8						
1	-0.161	-0.099	-0.163	-0.071	-0.122	0.015	0.121	-0.241	0.045	1.920	-0.104	-8.371	37.46	322.45	355.32
2	-0.187	-0.025	-0.180	-0.008	-0.143	0.019	0.108	-0.246	0.104	6.146	-1.108	-8.952	39.94	383.45	433.85
3	-0.154	-0.097	-0.157	-0.069	-0.123	0.016	0.153	-0.241	0.046	5.508	-0.696	-8.741	35.45	325.45	341.41
4	-0.178	-0.026	-0.173	-0.007	-0.141	0.019	0.138	-0.245	0.100	2.559	-1.487	-9.337	37.39	370.41	401.84
5	-0.159	-0.068	-0.200	-0.043	-0.126	0.017	0.148	-0.242	0.062	3.993	-0.905	-8.856	38.17	404.31	391.52
6	-0.084	-0.132	-0.122	-0.105	-0.097	-0.045	-0.013	-0.101	-0.060	7.498	-0.795	-8.120	38.01	371.44	354.08
7	-0.082	-0.105	-0.166	-0.068	-0.102	-0.015	-0.043	-0.084	0.050	5.349	-1.004	-8.200	40.64	450.33	426.51
8	-0.087	-0.131	-0.125	-0.101	-0.100	-0.037	-0.015	-0.113	-0.054	2.475	-0.580	-7.674	38.10	360.88	382.55
9	-0.099	-0.066	-0.136	-0.032	-0.115	-0.016	-0.016	-0.079	0.012	5.007	-1.535	-9.397	41.64	417.89	478.83
10	-0.078	-0.130	-0.119	-0.097	-0.073	-0.045	-0.065	-0.042	0.063	3.454	-1.065	-8.332	40.73	439.78	456.41
11	-0.092	-0.130	-0.133	-0.102	-0.108	-0.040	-0.003	-0.102	0.044	4.313	-0.119	-7.749	38.01	340.47	375.08
12	-0.091	-0.132	-0.126	-0.089	-0.106	-0.013	-0.025	-0.103	-0.064	3.974	-0.591	-7.43	37.71	376.88	394.53
13	-0.159	-0.070	-0.201	-0.058	-0.127	0.004	0.015	-0.265	0.040	2.284	-0.158	-8.132	41.06	452.35	417.59
14	-0.081	-0.133	-0.126	-0.094	-0.097	-0.104	-0.044	-0.068	-0.062	2.448	-0.578	-9.127	39.80	372.59	421.21
15	-0.099	-0.142	0.060	-0.175	-0.062	-0.116	0.147	-0.182	-0.018	3.326	0.319	-8.384	40.43	382.46	417.59
16	-0.130	-0.107	0.002	-0.114	-0.091	-0.114	0.149	-0.192	0 025	6.807	-0.134	-8.821	45.54	439.51	510.29

Table 3 Eight selected models and their statistical parameters for the correlation between molecular properties and antimalarial activity of vincadifformine

QSAR Model	Variables	R	R^2	F_{calc}/F_{table}	SE	PRESS
1	qC_1, qC_2, qC_4, qC_5, qC_6, qC_7, qN, qC_8, m, E_{HOMO}, E_{LUMO}, m, A	0.995	0.990	2.678	0.080	0.0597
2	qC_1, qC_2, qC_5, qC_6, qC_7, qN, qC_8, m, E_{HOMO}, E_{LUMO}, m, A	0.994	0.988	4.612	0.077	0.0414
3	qC_5, qC_6, qC_7, m, E_{HOMO}, E_{LUMO}, a, m, A	0.910	0.829	1.024	0.219	1.9600
4	qC_1, qC_5, qC_6, qC_7, qC_8, m, E_{HOMO}, a, A	0.937	0.878	3.677	0.184	0.2290
5	qC_1, qC_5, qC_6, qC_7, m, E_{HOMO}, m, A	0.912	0.832	3.438	0.202	0.0757
6	qC_1, qC_2, qC_3, qC_5, qC_6, qC_7, qN, qC_8, m, E_{HOMO}, E_{LUMO}, m, A	0.994	0.989	2.338	0.085	2.8400
7	qC_3, qC_6, qC_7, m, E_{HOMO}, E_{LUMO}, m, A	0.897	0.805	1.204	0.218	0.0933
8	qC_1, qC_2, qC_3, qC_5, qC_6, qC_7, qN, m, E_{HOMO}, E_{LUMO}, m, A	0.976	0.952	1.303	0.152	0.0796

Table 4 Coefficients of independent variables for 8 selected QSAR models as obtained from multilinear regression analysis

QSAR model	Coefficient of independent variables															Constants
	qC_1	qC_2	qC_3	qC_4	qC_5	qC_6	qC_7	qN	qC_8	N	E_{LUMO}	E_{HOMO}	a	m	A	
1	20.552	-18.173		1.452	-23.987	-11.342	6.486	-10.49	4.317	-0.177	-1.140	0.665		-0.007	0.022	-5.638
2	20.055	-16.109			-23.745	-10.28	6.582	-9.758	4.076	-0.176	-1.087	0.664		0.021	0.021	-5.168
3					-2.592	-4.770	3.534			-0.078	-0.263	0.478	-0122	-0.004	0.013	3.240
4	11.880				-15.752	-8.288	6.412		2.744	-0.132		0.480	-0.213		0.017	3.317
5	10.238				-14.220	-6.187	7.137			-0.089		0.419	-0.197		0.015	3.315
6	19.943	-17.387	1.097		-25.025	-9.429	6.248	-10.03	4.631	-0.186	-1.139	0..674		-0.007	0.021	-5.469
7			-2.303			-6.899	3.981			-0.086	-0.347	0.477		-0.007	0.010	0.677
8	18.428	-7.767	-4.235		-16.449	-16.449	7.688	-7.777		-0.793	-0.793	0.584		-0.088	0.0171	-2.784

This procedure finally gives 8 QSAR models as listed in Table 3 and the fitting coefficient of each variable included in the fitting is collected in Table 4. From Table 3, it is immediately emerged that all selected models show a good correlation between biological activity and descriptors selected for fitting ($r \geq 0.9$). This suggests that justification of the best model among 8 models selected in Table 3 is not adequate only by comparing the *r* size, especially for model 1, 2 and 6 because their *r* values is almost similar. Therefore ones should also take into account other statistical parameters of F (model significance) and SE (standard error). Comparing parameters F and SE of the three models, it is easily revealed that model 2 is the best model because it has highest F and lowest SE and moreover it also contains the least variables which means that the model is the simplest one; thus model 2 is selected as the best QSAR model. This model could therefore rationalize the search for new antimalarial drugs, necessary due to the rapid resistance development of plasmodium falcipare in tropical countries.

Model validation

It has been concluded that model 2 is the best model from the point of view of statistical parameters of the multilinear analysis. To see how good the model predict the activity of the drug series, calculation of the activity for each vincadifformine antimalarial drug has been performed using model 2 and the result of the calculation (predicted $\ln 1/IC_{50}$) was plotted against those obtained by experiments (observed $\ln 1/IC_{50}$) by regression method to see how well they correlate each other (Fig. 2). It is f this figure that model 2 predicts very well the activity of the series of vincadifformine drugs as can be seen from the values of the slope and correlation coefficient (*r*) of the plot which is approaching unity, i.e.: 1.004 and 0.993, respectively.

Further validation of the model can also be accessed by comparing PRESS (predictive residual sum of square) parameters of the models. This parameter defined as sum of square of the difference between observed activity and predicted activity calculated by corresponding model. The smaller the parameter the better the model because the difference between observed and predicted activity is minimized. Results of PRESS calculation for each model as listed in Table 4 again confirms that model 2 is the most reliable model because it induces the smallest PRESS value.

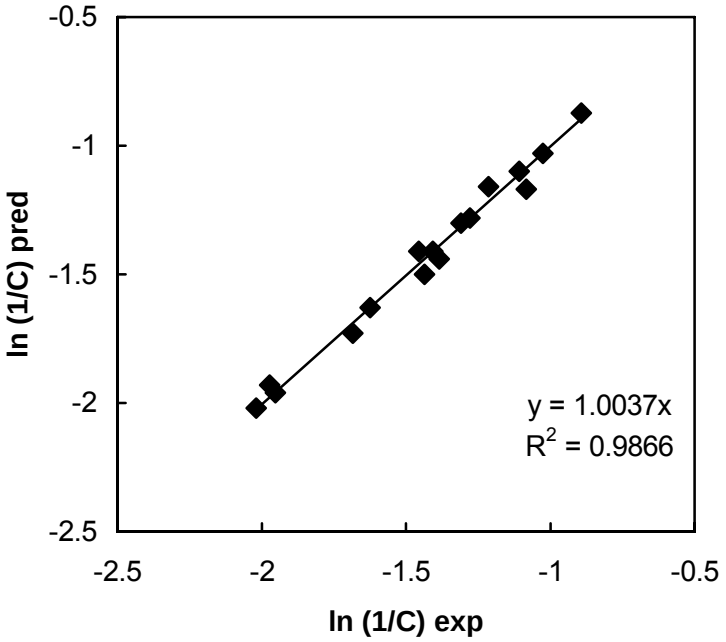


Fig. 2 Linear regression of experimentally observed antimalarial activity ($\ln 1/IC_{50}$) versus calculated one based on QSAR model 2 (n = 16, slope= 1.004 and r = 0.993).

Model evaluation and prediction of active center

According to the value of variable coefficient (*P*) obtained by the variation of atoms included in multilinear analysis (Table 4), atoms C_1, C_2, C_5, C_6, C_7, C_8 and N seems to be the most responsible for the pharmacological activity. Although some models listed in Table 3 also include C_3 (model 6-8) and C_4 (model 1), values of their coefficient are quite small indicating that these atoms are of rather less importance and therefore can be eliminated. By comparing the value of variable coefficient obtained for various models (see Table 4), ones observed that C_1 and C_5 assume prominent role in most as compared with the others, and that either C_6, C_2 or N has a significant influence, respectively. The less significant influence is observed for C_8 and C_7. This leads to the assumption that the most probable active center of vincadifformine drugs is C_6-N-C_5-C_8-C_1-C_2 or via the other side of N-containing ring, i.e: C_5-N-C_6-C_7-C_8-C_1-C_2 (see Fig. 3). This quite resembles to those reported for the active center of primaquine and mefloquine antimalarial drugs. The similarity of the active center of these type of the drugs to either primaquine and mefloquine but not to chloroquine may be one of the reason for the superiority of this drug in still being pharmacologically active against chloroquine-resistence strain, i.e: FcM29-Cameroun.

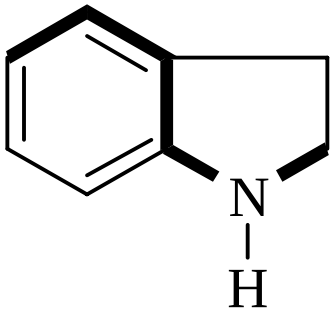


Fig. 3 Hypothetical active molecular region of antimalarial vincadifformine analogs.

For some antimalarial drugs, the search of relation between atomic net charges and pharmacological activity had shown that the most important atom are the heterocyclic planar ring containing nitrogen atom and some atoms which connect N atom and planar benzene ring containing amino substituent [7-9, 11]. The hypothetical "active molecular region" of vincadifformine (Fig. 3) suggests that the binding of the drug to nucleic bases of plasmodium's RNA most probably occurs *via* intercalation of this planar active center into the base-pairs of RNA supported by hydrogen bond between N atom of the drug and functional group of RNA bases. Such type of interaction will leads to the situation that the bulky rest of the drug molecule could block and hinder any further replication.

Conclusion

We have used a semi-empirical molecular orbital calculation AM-1 to study the correlation of antiplasmodial activity of a series of vincadifformine drugs against chloroquine-resistance FcM29 strain. The best overall correlation is given by the computed molecular properties of atomic net charges of heterocyclic ring, moment dipole, LUMO and HOMO energies, molecular mass as well as surface area. It is gratifying to observe that the hypothetical active center of the drugs corroborate nicely in terms of possible mode of binding of the drug to the plasmodium's RNA bases, i. e.: via intercalation and hydrogen bond. It has also been demonstrated from this result that AM1 or similar semi-empirical methods, although induces possible error sources, still seem to be a necessary and acceptable compromise for quantum pharmacological calculations on series of drug molecules of this size, including the search for active drug center.

References

[1] M. Grigorov, J. Weber, J. M. J. Tronchet, C.W. Jefford, W.K. Milhous, and D. Maric, J. Chem. Inf. Comput. Sci., 37 (1997) 124 130.
[2] T.D.R. News, 46 (1994, 5.
[3] S. L. Hoffman, G.M. Subramanian, F.H. Collins and J.C. Venter, Nature, 415 (2002**)** 702 709.
[4] Anonymous, Kompas Daily Newspaper, 27 January 2003, p. 10.
[5] in L.J. Bruce-Chwatt (ed.), Chemotherapy of Malaria*,* 2nd ed; W.H.O.: Geneva, 1981.
[6] W.H. Wernsdorfer, Parasitology Today, 7 (1991) 297 231.
[7] S.U Kokpol, S.V. Hannongbua,., N. Thongrit, and B. M. Rode, Anal. Sci., 4 (1988) 565 568.
[8] B. M. Rode, M. G. Schwendinger, S.U. Kokpol, S.V. Hannongbua, and S. Polman, Monatshefte fur Chemie, 120 (1989) 913 921.
[9] S. Polman, S. Kokpol, S. Hannongbua and B.M. Rode, Anal. Sci*.*, 5 (1989) 641 644.
[10] Mustofa, Ph.D. Dissertation, Laboratoire d'Immunologie et Parasitologie UFR Sciences Pharmatiques, Montpellier, France, 2001, pp. 103 107.
[11] S.N. Ethica and Mudasir, in H.D. Pranowo (Ed.), International Conference on Organic Chemistry (Indonesia) 2001. Proceeding of the Conference, 8 September 2001, Yogyakarta, Indonesia, Chemistry Department, Gadjah Mada University, Yogyakarta, 2001, pp. 313 321.