Aplikasi deskriptor kimia kuantum dalam analisis QSAR derivat kurkumin sebagai penghambat o-dealkilasi ethoxyresorufin

Application of quantum chemical descriptors in QSAR analysis of curcumin derivatives as ethoxyresorufin o-dealkylation inhibitor

B.S. Ari Sudarmanto*) dan R.A. Oetari
Faculty of Pharmacy, Gadjah Mada University, Yogyakarta, Indonesia

Abstrak

Telah dilakukan studi Hubungan Kuantitatif Struktur-Aktivitas (HKSA) terhadap 26 senyawa turunan kurkumin yang digunakan untuk memprediksi aktivitas penghambatan ethoxyresorufin O-deethylation (EROD).

Struktur turunan kurkumin dioptimasi geometrinya menggunakan metode mekanika kuantum semiempirik AM1, kemudian deskriptor kimia kuantum dihitung dari struktur teroptimasi. Sedangkan tehnik Genetic Algorithm (GA) yang digabungkan dengan analisis regresi multilinier digunakan untuk memilih dan membangun persamaan HKSA, yang menghubungkan sifat struktural dengan aktivitas biologis.

Hasil analisis menunjukkan bahwa secara statistik deskriptor kimia kuantum memberikan hubungan yang bagus dengan aktivitas penghambatan EROD. Tingkat energi orbital terendah yang tak terisi elektron (E_{lumo}), koefisien partisi (logP), momen dipol (μ) dan muatan atom pada C1, C6 dan C9 memainkan peran penting dalam aktivitas penghambatan EROD.
Kata kunci : HKSA, deskriptor kimia kuantum, turunan kurkumin, EROD.

Abstract

Quantitative Structure-Activity Relationship (QSAR) have established 26 curcumin derivatives to correlate and predict ethoxyresorufin O-deethylation (EROD) inhibitory activity.

The AM1 semiempirical quantum mechanic method was applied in geometry optimization and descriptor calculation. Genetic Algoritm combined with Multiple Linear Regression Analysis (GA-MLRA) technique was applied to select the descriptors and to generate the equation that relate the structural features to the biological activity.

The result of GA-MLRA showed that quantum chemical descriptors QSAR had good statistical fits. Energy level of the lowest unoccupied molecular orbital, logP, momen dipole, and nett charge of C1, C6 and C9 atoms play an important role in EROD inhibition.
Key words : QSAR, quantum-chemical descriptor, curcumin derivatives, EROD.

Introduction

Human cytochrome P450 1A1 is a well known aryl hydrocarbon hydroxylase and is involeved in the metabolic activation of procarcinogens of the polycyclic aromatic hydrocarbons. Cytochrome P450 1A1 expressed in lung, placenta, lymphocyte and liver (Chun *et al*, 2001).

Curcumin had a potent inhibitory effect on the EROD activity in β-naphtoflavone

(βNF)-induced isoenzyme cytochrome P450 1A1, with Ki value of 0.14 μM. To find more potent and selective P450 1A1 inhibitors, several curcumin derivatives compounds were evaluated for selective inhibition of P450 1A1 activity (Oetari, 1998).

QSAR (Kubinyi, 1993; Seydel, 1990) is a powelful lead-compound optimization technique, which quantitatively relates variations in biological activity to changes in molecular properties (descriptors). Usually there are two major approaches to analyze QSAR data: first, the activity of a series of compounds is expressed as multiple linear regression of descriptors, and second, the non-linear regression method represents the activity.

This study was used Genetic Algorithm (GA) method, combined with Multiple Linear Regression Analysis (MLRA) for deriving and validating QSAR equation. GA is an optimization algorithm based on the mechanisms of Darwinian evolution that uses random mutation, crossover and selection procedures to generate better models or solutions from an originally random starting population or sample (Melanie, 1999; Devillers, 1996; Hasegawa *et al,* 1999). GA can not only automatically select the optimum number of descriptors in regression analysis, but also construct multiple linear regression models through the use of linear and higher order polynomials. The GA method was used to select the optimum number of descriptors for use in regression analysis (Cho *et al*, 2001) .

The purpose of this research was to determine predictive QSAR models by analysis of data set containing 26 curcumin derivatives compounds. If the models are reasonable, it is possible to predict biological activity of non-tested molecules. Finally, the successful models of QSAR certainly decrease the number of compounds to be synthesized, by making it possible to select the most promising compounds.

Methodology

Data set

The data set contains 26 curcumin derivatives with ethoxyresorufin O-dealkylation inhibitory antivity. The compounds with CUR-, PAL- and PAR-code are curcumin derivatives modified at the two aromatic rings, C_4 of the heptadiene by aliphatic substitutions, and C_4 of the heptadiene by aromatic substitutions, respectively. (Oetari, 1998).

The activity was expressed in terms of 50% inhibition concentration (IC_{50}) of ethoxyresorufin O-deethylation. The IC_{50} value was converted by – log function to fit scale.

Molecular modelling

The molecular modelling studies were carried out using HyperChem 7.5 software package (HyperChem, 2002) All structures were drawn as enol form and *E* configuration of heptadiene moety. The Molecular Mechanics (MM+) force field was applied for preliminary structure optimization and study of the conformational behaviour of each compound. The next step was a reoptimization of the MM+ optimized structures by applying AM1 semiempirical method.

Descriptors

Table II list the descriptors employed in this work. All descriptors were calculated from AM1 optimized structure. The qCi, homo, lumo, logP, α, MR, Etot, HF and μ descriptor were calculated by HyperChem software package. The MTI, PSAr, ShpA and WIndx descriptors were calculated by ChemOffice software package.

Statistical methods

Preliminary model selection was performed by means of GA-MLRA technique as implemented in BuildQSAR program (de Oliviera and Gaudio, 2000). This approach allows selection of the models with the following characteristic: high correlation coefficient r, low standard deviation s, and high Fisher coefficient F. The selected models were identified by applying the leave-one-out technique which its predicting ability being evaluated and confirmed by cross validation coefficient Q^2 based on predictive error sum of squares (S_{PRESS}).

Result And Discussion

The QSAR study was established on 26 curcumin derivatives compounds for their EROD inhibition activity, which expressed in terms of IC_{50} (Table I). In this study, we screened 21 preselected descriptors, consisting 12 molecular descriptors and 9 atomic descriptors (nett atomic charge). The descriptors used in this study were listed in Table II.

Several runs of GA-MLRA variable selection technique implemented in Build QSAR program have resulted in models

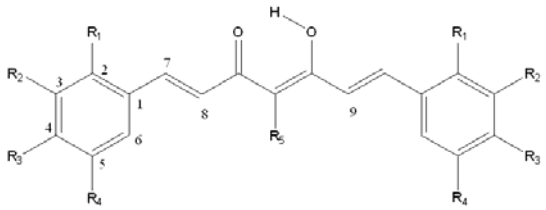

Table I. Structures of curcumin derivatives in training set and their EROD inhibitory activities

No.	Compound	R1	R2	R3	R4	R5	IC_{50} (μM)
1	CUR1	H	H	H	H	H	0.150
2	CUR2	H	H	OH	H	H	2.480
3	CUR3	H	OCH_3	OH	H	H	2.010
4	CUR4	H	OCH_3	H	H	H	0.020
5	CUR5	H	OCH_3	OCH_3	OCH_3	H	20.930
6	CUR6	H	CH_3	OH	CH_3	H	3.490
7	CUR7	H	C_2H_5	OH	C_2H_5	H	7.710
8	CUR8	H	i-C_3H_7	OH	i-C_3H_7	H	167.600
9	CUR9	H	t-C_4H_9	OH	t-C_4H_9	H	2624.000
10	CUR10	H	H	Cl	H	H	9.450
11	CUR11	H	H	OCH_3	H	H	2.750
12	CUR12	H	H	CH3	H	H	5.700
13	CUR14	H	OCH_3	OH	OCH_3	H	13.390
14	CUR15	H	O-CH_2-C_6H_5	H	O-CH_2-C_6H_5	H	470.900
15	CUR16	OCH_3	H	H	H	H	0.003
16	PAL1	H	OCH_3	OH	OCH_3	CH_3	1.470
17	PAL2	H	OCH_3	OH	OCH_3	C_2H_5	2.880
18	PAL3	H	OCH_3	OH	OCH_3	n-C_3H_7	0.310
19	PAL4	H	OCH_3	OH	OCH_3	i-C_3H_7	3.830
20	PAL5	H	OCH_3	OH	OCH_3	n-C_4H_9	2.030
21	PAR1	H	OCH_3	OH	OCH_3	C_6H_5	0.380
22	PAR2	H	OCH_3	OH	OCH_3	m-C_6H_4-CF_3	0.110
23	PAR3	H	OCH_3	OH	OCH_3	p-C_6H_4-OCH_3	0.220
24	PAR4	H	OCH_3	OH	OCH_3	p-C_6H_4-CH_3	0.430
25	PAR6	H	OCH_3	OH	OCH_3	p-C_6H_4-F	0.090
26	PAR7	H	OCH_3	OH	OCH_3	m,p-C_6H_4-$(NO_2)_2$	4.200

containing mainly lumo, logP and qC6 descriptors. The best models of QSAR equation and their regression statistics listed in Table III. Model 6 is the best QSAR model, which have highest r, F and Q^2, lowest s and s_{PRESS}, indicating that is the most powerful QSAR equation.

The best equation and cross validation results observed in this QSAR study are:

$$
\begin{aligned}
-\log IC_{50} = {} & 5.825\ (\pm 1.829)\ \text{lumo} \\
& -0.403\ (\pm 0.106)\ \text{logP} \\
& -0.236\ (\pm 0.156)\ \mu \\
& +155.876\ (\pm 36.980)\ \text{qC1} \\
& +80.053\ (\pm 20.885)\ \text{qC6} \\
& -161.083\ (\pm 46.308)\ \text{qC9} \\
& +4.995\ (\pm 5.820)
\end{aligned}
$$

$n = 26$; $r = 0.949$; $s = 0.465$; $F = 28.656$; $Q^2 = 0.843$; $s_{PRESS} = 0.585$

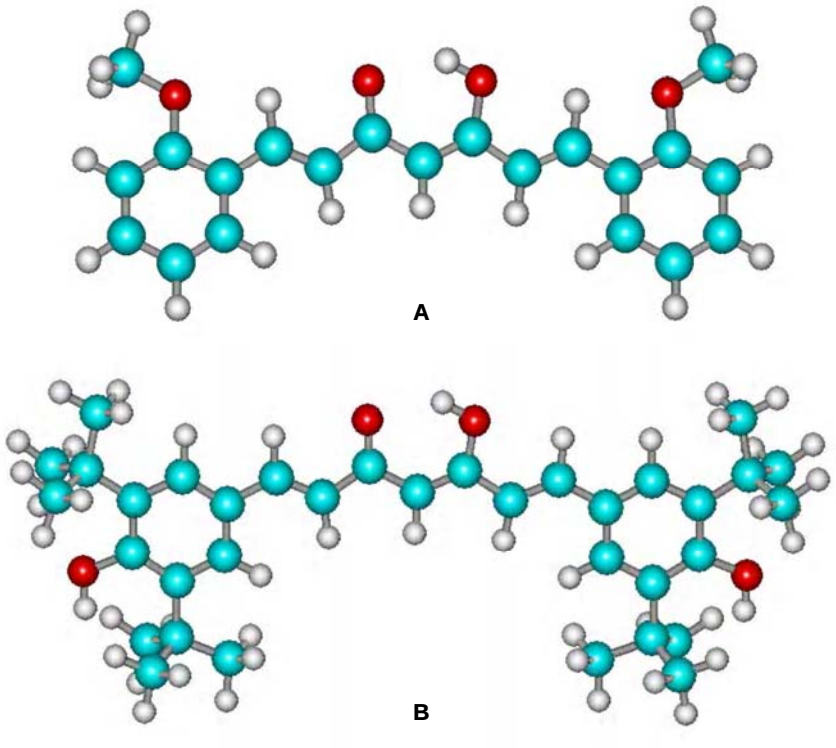

Figure 1. Ball and stick representations of AM1-optimized curcumin derivatives, CUR16 – the most potent (A) and CUR9 – the less potent (B) O-ethoxyresorufin dealkylation inhibitor. Carbon atoms are shown in cyan; Hydrogen atoms are shown in white; and Oxigen atoms are shown in red. The figure was constructed and rendered with HyperChem.

Table II. List of desciptor used in this work

Symbol	Descriptors
qCi *)	Nett atomic charge of i-th C atom
homo	Energy of the highest occupied molecular orbital
lumo	Energy of the lowest unoccupied molecular orbital
logP	Partition coefficient (octano-water)
α	Polarizability
MR	Molar Refractivity
Etot	Energy total
HF	Heat of formation
μ	Dipole momen
MTI	Molecular topological index
PSAr	Polar surface area
ShpA	Shape attribute
WIndx	Wiener index

*) contain 9 descriptors, $i = 1, 2, 3, \ldots, 9$

Calculated Q^2 and s_{PRESS} values show that the predictive power of this QSAR equation is significant. Predictions for the lead optimization in this set of compound can be summarized as follows:

The activity contributions obtained in this QSAR model show that the hydrophobic parameter, logP, energy level of unoccupied molecular orbital, lumo, and electronic parameter at C-6, qC6, play more important role than others.

Hydrophobic properties play an important role in molecule transport and penetration between membrane. This QSAR model suggest that logP value should be small enaouh. It can be understood, if logP value is too large molecule have worse distribution.

Table III. The best models and their regression statistics

Model	n	m	r	s	F	Q^2	s_{PRESS}
3	26	5	0.924	0.549	23.376	0.782	0.670
5	26	5	0.918	0.571	21.302	0.768	0.692
6	**26**	**6**	**0.949**	**0.465**	**28.656**	**0.843**	**0.585**
7	26	6	0.944	0.485	26.094	0.831	0.605
14	26	6	0.941	0.497	24.641	0.810	0.643

n = number of compounds; m = number of descriptors; r = correlation coefficient; s = standard deviation; F = Fisher coefficient; Q^2 = squared cross validation regression coefficient; s_{PRESS} = standard deviation of cross validation prediction

Table IV. Experimental and predicted EROD inhibitory activity for training set (-log IC_{50} values were calculated from IC_{50} in mM)

No.	Experimental -logIC_{50}	Predicted -logIC_{50}	Residual
1	4.824	4.562	0.262
2	3.606	3.306	0.300
3	3.697	4.497	-0.800
4	5.699	5.646	0.053
5	2.679	2.656	0.023
6	3.457	3.653	-0.196
7	3.113	2.738	0.375
8	1.776	1.993	-0.217
9	0.581	0.908	-0.327
10	3.025	3.094	-0.069
11	3.561	3.369	0.192
12	3.244	4.175	-0.931
13	2.873	3.004	-0.131
14	1.327	1.164	0.163
15	6.523	6.275	0.248
16	3.833	3.792	0.041
17	3.541	4.440	-0.899
18	4.509	4.245	0.264
19	3.417	2.756	0.661
20	3.693	4.098	-0.405
21	4.420	4.381	0.039
22	4.959	4.669	0.290
23	4.658	4.222	0.436
24	4.367	4.046	0.321
25	5.046	4.734	0.312
26	3.377	3.381	-0.004

Energy level of homo and lumo play an important role in chemical reaction. In drug – receptor interaction, lumo corresponds to ability of electron affinity. Curcumin derivatives molecule play as electron donor from cytochrome P450. This QSAR model suggest that higher lumo energy level give better activity.

The substitutions at aromatic rings that be able to give more positif C6 charge through inductive effect improve EROD activity.

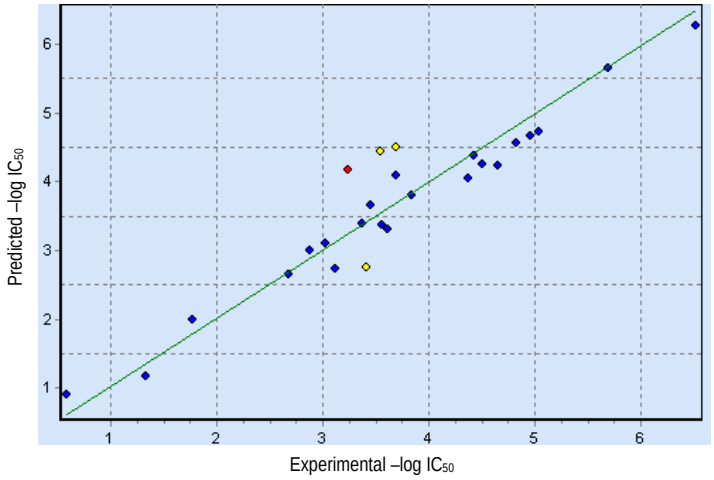

Figure 2. Plots of Predicted and experimental EROD inhibitory activity

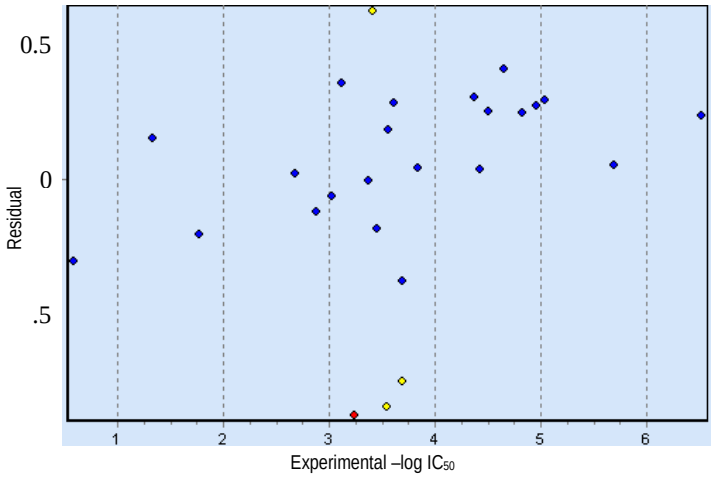

Figure 3. Plots of Residual and experimental EROD inhibitory activity

Conclusion

1. A linear model was obtained with a relatively good predictive ability to guide the synthesis of other curcumin derivatives EROD inhibitor.
2. Energy level of lowest unoccupied molecular orbital, logP, momen dipole, and nett charge of C1, C6 and C9 atoms play an important role in EROD inhibition.

References

Cho, D.H., Lee, S.K, Kim, B.T., No, K.T., 2001, Quantitative Structure–Activity Relationship (QSAR) Study of New Fluorovinyloxyacetamides, *Bull. Korean Chem. Soc.*, *22*, 388 – 394.

Chun, Y.J., Ryu, S.Y., Jeong, .T.C., and Kim, M.Y., 2001, Mechanism-Based Inhibition of Human Cytochrome P450 1A1 by Rhapontigenin, *Drug Metabolism and Disposition*, *29*, 4, 389-393.

De Oliveira, D.B., Gaudio, A.C, 2000, *Quant. Struct. Act. Relat.*, 19, 599.

Devillers, J., *Genetic Algorithm in Molecular Modelling*, Academic Press, London.

Franke, R., 1984, *Theoretical Drug Design Methods*, Elsevier, Amsterdam.

Hasegawa, K., Kimura, T., Funatsu, K., 1999, GA Strategy for Variable Selection in QSAR Studies: Application of GA–Based Region Selection to a 3D–QSAR Study of Acetylcholinesterase Inhibitors, *J. Chem. Inf. Comput. Sci.*, *39*, 112 – 120.

HyperChem v7.5, 2002, HyperCube.

Karelson, M., Lobanov, V.S., Katritzky, A. R., 1996, Quantum–Chemical Descriptors in QSAR/QSPR Studies, *Chem. Rev., 96,* 1027 – 1043.

Kubinyi, H., 1993, *QSAR : Hansch Analysis and Related Approaches*, VCH Verlagsgesellschaft, Weinheim.

Lee, K.W., Kwon, S.Y., Hwang, S., Lee, J.U., Kim, H., 1996, Quantitative Structure– Activity Relationships (QSAR) Study on C– 7 Substituted Quinolone, *Bull. Korean Chem. Soc.*, *17*, 147 – 152.

Melanie, M., 1999, *An Introduction to Genetic Algorithms*, 5th edition, MIT Press, London.

Oetari, R.A., 1998, *The Interaction Between Curcumin an Curcumin Analogues and Cytochrome P450*, Dissertation, UGM, Yogyakarta.

Seydel, J.K., 1990, *Summary Lecture Course QSAR, Mid Career Training in Pharmacochemistry*, Fakultas Farmasi UGM, Yogyakarta.

* Korespondensi : B.S. Ari Sudarmanto, S.Si., M.Si.
Fakultas Farmasi Universitas Gadjah Mada Yogyakarta
Sekip Utara Yogyakarta, 55281, Telp. 0274-6492565
E-mail: arie_sudarmanto@ugm.ac.id