Original Article

Mahidol University Journal of Pharmaceutical Sciences 2013; 40 (3), 8-15

QSAR Analysis of Chromone Derivatives as HIV-1 Protease Inhibitors

W. Samee[a] *and J. Ungwitayatorn*[b]*

[a] *Faculty of Pharmacy, Srinakharinwirot University, Nakhonayok, 26120, Thailand.*

[b] *Faculty of Pharmacy, Mahidol University, 447 Sri-Ayudhya Road, Bangkok 10400, Thailand*

Abstact

Quantitative structure-activity relationship (QSAR) of HIV-1 protease inhibitory activity of chromone derivatives was developed using multiple linear regression (MLR) analysis. The QSAR model generated using binding energies of chromones to the enzyme as single descriptor was of high statistical quality and predictive potential (cross-validated r^2 or $q^2 = 0.941$, non cross-validated $r^2 = 0.943$, SPRESS = 6.256). The predictive abilities was comparable to those of previously studied 3D-QSAR (comparative molecular field analysis, CoMFA) model. The predicted (calculated) and experimental inhibitory activities were very well correlated. The resulting QSAR model can be used as a preliminary tool for screening purpose before costly and time-consuming synthesis.

Key words: QSAR, MLR, HIV-1 protease, Chromone derivatives, Docking

INTRODUCTION

Human immunodeficiency virus type 1 protease (HIV-1 PR), a homodimeric aspartyl protease that cleaves the gag and gag-pol viral polyproteins, plays a vital role in the replication cycle of the virus[1-3]. The relatively small size of this enzyme and the availability of good crystal structures have made HIV-1 PR an attractive target for AIDS (acquired immunodeficiency syndrome) treatment. Several HIV-1 PR inhibitors, e.g., amprenavir, atazanavir, darunavir, indinavir, fosamprenavir, lopinavir, nelfinavir, ritonavir, saquinavir and tipranavir) have been approved by the United States Food and Drug Adminis tration (US FDA) and are currently in use in combination with reverse transcriptase inhibitors[4-5]. Despite combination chemotherapy, resistant variants have developed with reduced sensitivity to these inhibitors[6]. The bioavailability and toxicity profiles of HIV-1 PR inhibitors are also of importance. Therefore, the need to discover a new generation of inhibitors which possess low toxicity, high bioavailability and more potency against the mutant forms of the virus is still continuing[7-12].

Previous investigations in our research group have designed and synthesized a series of chromone derivatives as a new class of non-peptide HIV-1 PR inhibitors[13]. In this study, quantitative structure-activity relationship (QSAR) was investigated based on the fact that the biological activity of a compound is a function of its physicochemical properties. The multiple linear regression (MLR) analysis was used for study the correlation between HIV-1 PR inhibitory activity and binding energy obtained form docking study and other molecular properties such as highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), logarithm of partition coefficient (log P), and molar refractivity (MR).

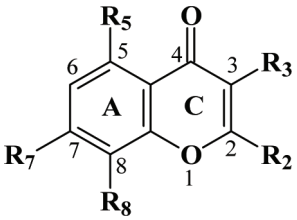
***Corresponding author:** Tel: +66-2-644-8680; Fax: +66-2-644-8695; E-mail: jiraporn.ung@mahidol.ac.th

MATERIALS AND METHODS

2.1 Biological data

A series of chromone derivatives listed in Table 1 were tested in vitro for HIV-1 PR inhibitory activity by stop time HPLC analysis of enzyme-substrate interaction using His-Lys-Ala-Arg-Val-Leu-(p-NO2-Phe)-Glu-Ala-Nle-Ser-Amide as a substrate. The activity was measured corresponding to the degree of inhibition of the cleavage of substrate by the synthesized compounds. The results were reported as % inhibition. More details concerning HIV-1 PR assay used in this study was reported in reference[13].

Table 1. Structure and HIV-1 PR inhibitory activity of studied chromone derivatives.



Compd	R_2	R_3	R_3	R_7	R_8	% inhibition
1	Phenyl	H	H	H	OH	64.34±1.01
2	CH_3	H	H	OH	H	56.00±1.74
3	Phenyl	H	H	OH	H	75.04±1.72
4	Benzyl	H	H	OH	H	50.26±4.73
5	CH_3	H	H	OH	OH	92.02±0.24
6	Phenyl	H	H	OH	OH	88.17±1.09
7	Phenyl	CH3	H	OH	H	57.29±5.06
8	Benzyl	CH3	H	OH	H	18.97±5.79
9	4-(NO_2)-phenyl	H	H	OH	H	63.52±1.63
10	3-(CF_3)-phenyl	H	H	OH	H	29.27±1.10
11	4-(F)-phenyl	H	H	OH	H	37.52±7.03
12	3,5-(diNO_2)-phenyl	H	H	OH	H	26.32±4.49
13	3-(Cl)-phenyl	H	H	OH	H	27.36±3.84
14	3,4-(diCl)-phenyl	H	H	OH	H	47.63±2.11
15	3-(CF_3)-phenyl	H	OH	OH	H	74.80±0.26
16	4-(F)-phenyl	H	OH	OH	H	88.13±2.33
17	3,4-(diF)-phenyl	H	OH	OH	H	80.25±3.25
18	4-(t-butyl)-phenyl	H	OH	OH	H	89.29±3.47
19	3-(Cl)-phenyl	H	OH	OH	H	73.62±0.58
20	3,4-(diCl)-phenyl	H	OH	OH	H	85.26±1.20
21	4-(OCH_3)-phenyl	H	OH	OH	H	88.68±2.27
22	4-(NO_2)-phenyl	4-(NO2)-benzoyl	H	OH	OH	92.24±1.70
23	4-(NO_2)-phenyl	4-(NO2)-benzoyl	H	OH	H	74.47±2.76
24	3-(CF_3)-phenyl	3-(CF3)-benzoyl	H	OH	OH	93.16±1.74
25	3-(CF_3)-phenyl	3-(CF3)-benzoyl	H	OH	H	74.98±3.86
26	4-(F)-phenyl	4-(F)-benzoyl	H	OH	OH	84.94±1.54
27	4-(F)-phenyl	4-(F)-benzoyl	H	OH	H	50.79±0.97
28	4-(NO_2)-phenyl	4-(NO2)-benzoyl	OH	OH	H	88.39±0.11
29	4-(OCH_3)-phenyl	4-(OCH3)-benzoyl	H	OH	H	34.06±9.89
30	3-(OCH_3)-phenyl	3-(OCH3)-benzoyl	H	OH	H	24.67±4.31
31	Benzyl	H	H	OH	OH	93.30±0.07
32	4-(*t*-butyl)-phenyl	H	H	OH	H	78.89±5.71
33	4-(NO_2)-phenyl	H	OH	OH	H	27.20±0.82
34	3,5-(diNO_2)-phenyl	H	OH	OH	H	10.48±2.52
35	3,4-(diF)-phenyl	3,4-(diF)-benzoyl	H	OH	H	88.98±1.44

2.2 Generation of the molecular structures and docking

The molecular structures of chromone derivatives were modeled with SYBYL version 7.0 molecular modeling program (Tripos Associates, Saint Louis, MO) on an Indigo Elan workstation (Silicon Graphics Inc., Mountain View, CA) using the sketch approach. The fragment libraries in SYBYL database were used as building blocks for the construction of larger ones. Each structure was energy minimized using the standard Tripos force field (Powell method and 0.05 kcal/mol.Å energy gradient convergence criteria) and electrostatic charge was assigned by the Gasteiger-Hückel method. These conformations were used as starting conformations to perform docking. The docking was performed using FlexiDock option in SYBYL/Biopolymer program. The crystal structure of the HIV-1 PR complexed with inhibitor (pdb 1AJX) was obtained from the Brookhaven Protein Data Bank (http://www.rcsb.org/pdb). The inhibitor structure was first removed from the complex structure, then chromone inhibitor was placed into the binding site. Water molecules and ions were removed and hydrogen atoms were added at appropriate geometry. The charges were assigned by Kollman force field for protein and Gasteiger Hückel for ligands. The amino acids involved in the binding pocket were chain A: Arg8, Leu23, Asp25, Thr26, Gly27, Ala28, Asp29, Val32, Ile47, Gly48, Gly49, Ile50, Pro81, Val82, Ile84, and chain B: Arg8′, Leu23′, Asp25′, Thr26′, Gly27′, Ala28′, Asp29′, Val32′, Ile47′, Gly48′, Gly49′, Ile50′, Pro81′, Val82′, Ile84′. The binding energy calculated from docking was used as one of the independent variables (descriptors).

2.3 Calculation of physicochemical properties

The physicochemical properties used in this study were electronic, lipophilic and steric properties. The representatives of electronic property were frontier orbital, i.e, HOMO and LUMO. Log P was used as representative of lipophilic property and MR was used for steric property. MR and logP were calculated by Chem Office version 10.0. HOMO and LUMO energies were calculated by MOPAC 6.0-PM3 option in SYBYL version 8.0.

2.4 Multiple linear regression

MLR was performed using the SPSS for Windows Release 11.0 package by the stepwise method. The statistical values, multiple correlation coefficient (r), standard errors (s), cross-validation r^2 (q^2), non cross-validated r^2 and standard error of prediction (S_{PRESS}) were used to evaluate the obtained QSAR models. A total of 35 chromone compounds were used as data set, of which 5 compounds were chosen as a test set while the remaining 30 compounds were treated as a training set. The selected test set represented a range of activity similar to that of the training set and was used to evaluate the predictive power of the QSAR model.

RESULTS AND DISCUSSION

QSAR has been performed concerning the HIV-1 PR inhibitory activity of a series of chromone derivatives by MLR approach. The binding energy obtained from docking study, HOMO, LUMO, logP, and MR were used as molecular descriptors (Table 2).

In order to explore contribution of each descriptor separately, all possible combinations of the descriptors were investigated (Table 3). The MLR equation used for the QSAR model developed was as followed:

$y = a1x1 + a2x2 + a3x3 + \ldots. + anxn + b$

y = dependent variable (% inhibition)

a1, a2, a3, …. an = the regression coefficients of independent variables

x1, x2, x3, …. xn = independent variables

b = the regression constant obtained from the fit

To avoid self correlation among the variables, the correlation matrix was calculated and result as shown in Table 4.

Table 2. Independent variables (descriptors) used in QSAR models.

Compd	Binding energy (kcal/mol).	HOMO	LUMO	logP	MR
Training set					
1	-36.05	-9.56	-0.77	2.68	68.44
2	-33.99	-9.74	-0.51	0.61	48.80
3	-36.05	-9.54	-0.80	2.68	68.44
4	-33.17	-9.70	-0.49	2.28	73.42
5	-38.27	-9.52	-0.61	0.22	50.49
6	-39.61	-9.49	-0.89	2.29	70.14
7	-33.66	-9.51	-0.68	3.03	72.80
8	-29.57	-9.60	-0.47	2.63	77.78
9	-34.48	-9.99	-1.73	1.98	74.92
10	-31.38	-9.81	-1.15	3.60	74.42
11	-32.38	-9.57	-0.94	2.84	68.66
12	-31.08	-10.20	-2.11	1.27	81.39
13	-31.64	-9.61	-0.93	3.24	73.25
14	-32.65	-9.65	-1.11	3.80	78.05
15	-36.29	-9.80	-1.26	3.21	76.11
16	-37.55	-9.70	-1.11	2.45	70.35
17	-38.57	-9.79	-1.30	2.61	70.57
18	-39.53	-9.56	-0.88	4.00	88.80
19	-36.07	-9.74	-1.04	2.85	74.94
20	-37.74	-9.76	-1.21	3.41	79.75
21	-39.54	-9.39	-0.85	2.17	76.60
22	-39.75	-10.22	-1.97	1.95	112.97
23	-36.08	-10.27	-1.88	2.34	111.28
24	-39.95	-9.95	-1.36	5.21	111.98
25	-36.00	-10.12	-1.27	5.60	110.28
26	-37.87	-9.78	-1.12	3.68	100.46
27	-34.90	-9.77	-1.04	4.07	98.77
28	-39.21	-10.28	-1.96	1.95	112.97
29	-32.45	-9.20	-0.77	3.50	111.26
30	-31.04	-9.27	-0.80	3.50	111.26
Test set					
31	-39.82	-9.49	-0.57	1.89	75.11
32	-36.52	-9.44	-0.76	4.39	87.11
33	-31.90	-9.93	-1.82	1.59	76.61
34	-30.29	-10.05	-2.17	0.88	83.08
35	-37.54	-9.90	-1.29	4.39	99.20

Table 3. Statistical results for the QSAR models.

QSAR models	q^2	*PRESS*	S_{PRESS}	*r*	r^2	*F*	*S*
B.E.	0.941	1095.959	6.256	0.971	0.943	464.837	5.879
HOMO	-0.083	18456.410	25.674	0.256	0.066	1.963	23.844
LUMO	-0.132	19287.180	26.246	0.233	0.054	1.608	23.987
LogP	-0.155	19673.820	26.507	0.041	0.002	0.048	24.645
MR	-0.147	19548.640	26.423	0.074	0.005	0.152	24.599
B.E. + HOMO	0.933	1148.604	6.522	0.972	0.945	232.479	5.885
B.E. + LUMO	0.932	1165.646	6.571	0.971	0.943	224.695	5.980
B.E. + logP	0.931	1176.509	6.601	0.972	0.944	228.989	5.927
B.E. + MR	0.937	1076.388	6.314	0.973	0.948	244.344	5.750
B.E. + HOMO + LUMO	0.946	1006.313	6.221	0.976	0.952	171.961	5.607
B.E. + HOMO + logP	0.926	1266.321	6.979	0.973	0.946	152.165	5.942
B.E. + HOMO + MR	0.939	1042.341	6.332	0.976	0.953	174.350	5.570
B.E. + LUMO + logP	0.926	1263.322	6.971	0.972	0.944	147.347	6.033
B.E. + LUMO + MR	0.931	1169.298	6.706	0.974	0.948	158.815	5.823
B.E. + logP + MR	0.932	1167.377	6.701	0.973	0.948	156.891	5.857
B.E. + HOMO + LUMO + logP	0.936	1094.229	6.616	0.976	0.953	126.404	5.666
B.E. + HOMO + LUMO + MR	0.939	1035.125	6.435	0.978	0.956	134.584	5.499
B.E. + HOMO + logP + MR	0.934	1124.667	6.707	0.976	0.953	127.396	5.645
B.E. + LUMO + logP + MR	0.926	1258.572	7.059	0.974	0.948	114.939	5.928
All descriptors	0.935	1109.389	6.799	0.978	0.956	103.644	5.605

B.E. = binding energy obtained from docking

Table 4. The correlation matrix of all descriptors used in QSAR study.

	B.E.	HOMO	LUMO	logP	MR
B.E.	1	0.073	-0.005	-0.111	-0.141
HOMO		1	0.838	0.109	-0.347
LUMO			1	0.151	-0.424
logP				1	0.525
MR					1

From Table 3, it was found that only binding energy (B.E.) was sufficient for getting a good statistical results ($q^2 = 0.941$, $r^2 = 0.943$, SPRESS = 6.256). A plot of % inhibition and binding energy shown in Figure 1 indicated that the molecular docking results were closely related to the experimental inhibitory activity results. The top five most potent compounds, i.e., chromones **5, 18, 22, 24** and **31** exhibited binding energy -38.27 to -39.95 kcal/mol. The increase in the number of descriptors was not much improving the statistical quality of the model (with all descriptors, $q^2 = 0.935$). The results also indicated that using each of HOMO, LUMO, logP and MR as descriptor gave poor statistical outcomes for developing the model. Because HOMO and LUMO were correlated descriptors (correlation = 0.838, Table 4), therefore, even if model with 3 descriptors, i.e., binding energy, HOMO, and LUMO exhibited the highest q^2 (0.946) and the lowest SPRESS (6.221), the best QSAR model was built from binding energy as the only descriptor. The final QSAR equation was shown as followed:

% inhibition = -7.559 (±0.351) B.E. -204.355 (±12.511)

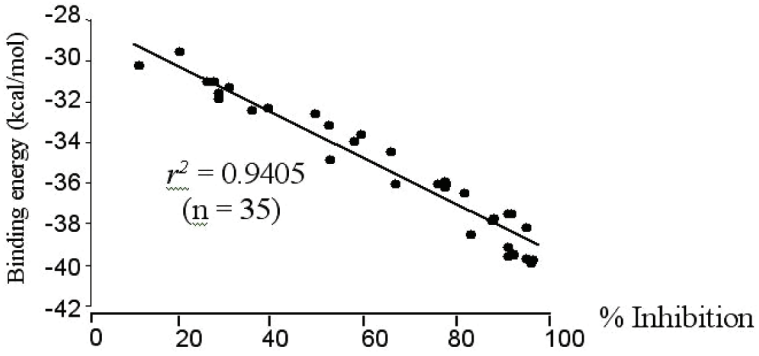


Figure 1. Plot of linear relationship between % inhibition and binding energy of 35 chromone derivatives.

This model was used to calculate the % inhibition of all compounds in training set and test set. Table 5 summarized the predicted activity of compounds in the training set and test set using QSAR equation. The scattered plots of the experimental and predicted activity of compounds in the training set and test set were shown in Figure 2. The statistical outcomes and the linearity of the scattered plots indicated the high fitting and predictive ability of the derived QSAR model. In our previous study, 3-dimensional QSAR approach using comparative molecular field analysis (CoMFA) was also applied to the same data set[14]. The CoMFA model was obtained with q^2 = 0.646, SPRESS = 28.474 and number of the optimum components = 3, including the LUMO in addition to CoMFA (steric and electrostatic) fields. As seen from the results in Table 5, both classical QSAR and CoMFA models gave comparable correlation between experimental and calculated values.

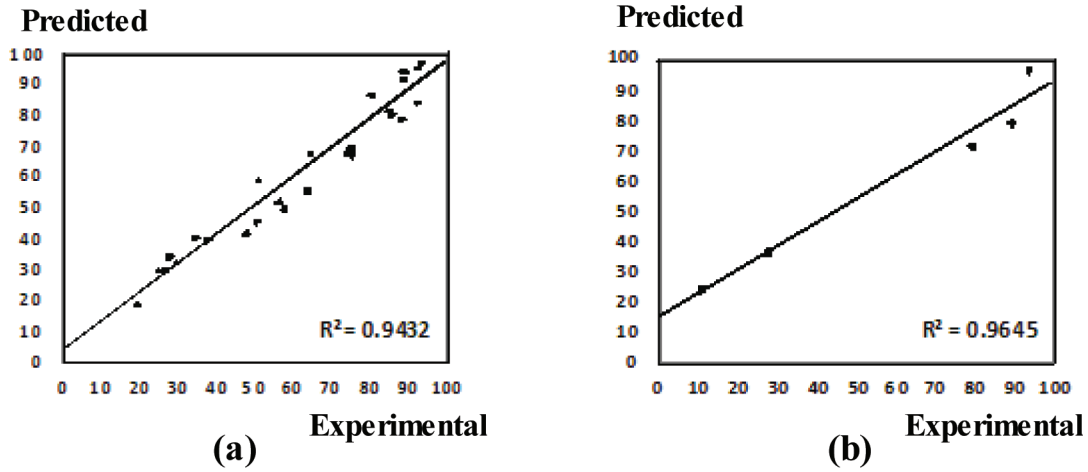


Figure 2. Scattered plots between experimental and predicted % inhibition for training set (a), and test set (b) obtained from QSAR model.

Table 5. Predicted, experimental activities and the residuals obtained from QSAR model comparing to CoMFA model.

Compd	HIV-1 PR inhibitory activity (% inhibition)				
		QSAR		CoMFA	
	Experimental	Predicted	Residual	Predicted	Residual
Training set					
1	64.34	68.15	-3.81	66.36	-2.02
2	56.00	52.58	3.42	53.96	2.04
3	75.04	68.15	6.89	73.79	1.25
4	50.26	46.38	3.88	50.69	-0.43
5	92.02	84.93	7.09	87.86	4.16
6	88.17	95.06	-6.89	92.58	-4.41
7	57.29	50.08	7.21	58.04	-0.75
8	18.97	19.16	-0.19	19.64	-0.67
9	63.52	56.28	7.24	56.20	7.32
10	29.27	32.85	-3.58	28.86	0.41
11	37.52	40.41	-2.89	46.67	-9.15
12	26.32	30.58	-4.26	25.45	0.87
13	27.36	34.81	-7.45	35.35	-7.99
14	47.63	42.45	5.18	49.10	-1.47
15	74.80	69.96	4.84	73.11	1.69
16	88.13	79.49	8.64	80.90	7.23
17	80.25	87.20	-6.95	81.17	-0.92
18	89.29	94.45	-5.16	89.99	-0.70
19	73.62	68.30	5.32	70.88	2.74
20	85.26	80.92	4.34	82.45	2.81
21	88.68	94.53	-5.85	90.73	-2.05
22	92.24	96.12	-3.88	92.73	-0.47
23	74.47	68.37	6.10	65.64	8.83
24	93.16	97.63	-4.47	97.94	-4.78
25	74.98	67.77	7.21	70.52	4.46
26	84.94	81.90	3.04	81.91	3.03
27	50.79	59.45	-8.69	54.32	-3.53
28	88.39	92.03	-3.64	97.88	-9.49
29	34.06	40.93	-6.87	31.09	2.97
30	24.67	30.28	-5.61	25.67	-1.00
Test set					
31	93.30	96.64	-3.34	98.44	-5.14
32	78.89	71.70	7.19	85.69	-6.80
33	27.20	36.78	-9.58	37.15	-9.95
34	10.48	24.61	-14.13	5.43	5.05
35	88.98	79.41	9.57	84.47	4.51

CONCLUSION

In this study, a simple but powerful QSAR model was developed based on 35 chromone derivatives as data set. The binding energy of chromone molecules against HIV-1 PR was the single best descriptor and showed strong correlation with the activity. The obtained result indicated that the proposed QSAR model provided a feasible and practical tool for the rapid screening of the HIV- PR inhibitory activity of compounds in the chromone series.

ACKNOWLEDGEMENT

This research is supported by the Office of the Higher Education Commission and Mahidol University under the National Research Universities Initiative.

REFERENCES

1. Kohl NE, Emini EA, Schleif WA. Active human immunodeficiency virus protease is required for viral infectivity. *Proc Natl Acad Sci USA* 1988; 85: 4686-4690.
2. McQuadre TJ, Tomasselli AG, Liu L, Karacostas V, Moss B, Sawyer TK, Heinrikson RL, Tarpley WG. A preliminary study of ritonavir, an inhibitor of HIV protease, to treat HIV-1 infection. *Science* 1990; 247: 454-456.
3. Nagy K, Young M, Baboonian C, Merson J, Whittle P, Oroszlan S. Antiviral activity of human immunodeficiency virus type 1 protease inhibitors in a single cycle of infection: evidence for a role of protease in the early phase. *J Virol* 1994; 64: 757-765
4. Patick AK , Boritzki TJ, Bloom LA. Activities of the human immunodeficiency virus type 1 (HIV-1) protease inhibitor nelfinavir mesylate in combination with reverse transcriptase and protease inhibitors against acute HIV-1 infection in vitro. *Antimicrob Agents Chemother* 1997; 1: 2159-2164.
5. Temesgen Z, Warnke D, Kasten MJ. Current status of antiretroviral therapy. *Expert Opin Pharmacother* 2006; 7: 1541-1554.
6. Boden D, Markowitz M. Resistance to human immunodeficiency virus type 1 protease inhibitors. *Antimicrob Agent Chemother* 1998; 42: 2775-2783.
7. Dorsy BD, Levin RB, McDanial DL, Vacca JP, Guare JP, Darke PL, Zugay JA, Emini EA, Schleif WA, Quintero JC, Lin JH, Chen W, Holloway MK, Fitzgerald PMD, Axel MG, Ostovic D, Anderson PS, Hufl JR. The design of a potent and orally bioavailable HIV protease inhibitor. *J Med Chem* 1994; 37: 3443-3451.
8. Danner SA, Carr A, Leonard JM. A short-term study of the safety pharmacokinetics, and efficacy of ritonavir, an inhibitor of HIV-1 protease. *N Engl J Med* 1995; 333: 1528-1533.
9. Markowitz M, Saag M, Powderly WG. A preliminary study of ritonavir, an inhibitor of HIV protease, to treat HIV-1 infection. *N Engl J Med* 1995; 333: 1534-1539.
10. Lehr P, Billich A, Charpiot B, Ettmayer P, Scholz D, Rosenwirth B, Gstach H. Inhibitors of human immunodeficiency virus type 1 protease containing 2-aminobenzyl-substituted 4-amino-3-hydroxy-5-phenylpentanoic acid: Synthesis, activity, and oral bioavailability. *J Med Chem* 1996; 39: 2060-2067.
11. Vacca JP, Dorsey BD, Schleif WA. An orally bioavailable human immunodeficiency virus type-1 protease inhibitors with oral bioavailability. *J Med Chem* 1996; 39: 3203-3216.
12. De Lucca GV, Liang J, I. De Lucca I. Stereospecificity synthesis, structure activity relationship, and oral bioavailability of tetrahydropyrimidine-2-one HIV protease inhibitors. *J Med Chem* 1999; 42: 135-152.
13. Ungwitayatorn J, Wiwat C, Samee W, Nunthanavanit P, Phosrithong N. Synthesis, *in vitro* evaluation, and docking studies of novel chromone derivatives as HIV-1 protease inhibitor. *J Mol Struct* 2011; 1001: 152-161.
14. Ungwitayatorn J, Samee W, Pimthon J. 3D-QSAR studies on chromone derivatives as HIV-1 protease inhibitors. *J Mol Struct* 2004; 689: 99-106.