

## Various Partial Charge Schemes on 3D-QSAR Models for P-gp Inhibiting Adamantyl Derivatives

Changdev G. Gadhe, Thirumurthy Madhavan, Gugan Kothandan, Tae-Bum Lee,<sup>†</sup> Kyeong Lee,<sup>‡,\*</sup> and Seung Joo Cho<sup>§,\*</sup>

*Department of Bio-New Drug Development, College of Medicine, Chosun University, Gwangju 501-759, Korea*

<sup>†</sup>*Research Center for Resistant Cells, College of Medicine, Chosun University, Gwangju 501-759, Korea*

<sup>‡</sup>*College of Pharmacy, Dongguk University-Seoul, Seoul 100-715, Korea. \*E-mail: kaylee@dongguk.edu*

<sup>§</sup>*Department of Cellular · Molecular Medicine and Research Center for Resistant Cells, College of Medicine, Chosun University, Gwangju 501-759, Korea. \*E-mail: chosj@chosun.ac.kr*

*Received January 13, 2011, Accepted March 22, 2011*

We developed three-dimensional quantitative structure activity relationship (3D-QSAR) models for 17 adamantyl derivatives as P-glycoprotein (P-gp) inhibitors. Eighteen different partial charge calculation methods were tested to check the feasibility of the 3D-QSAR models. Best predictive comparative molecular field analysis (CoMFA) model was obtained with the Austin Model 1-Bond Charge Correction (AM1-BCC) atomic charge. The 3D-QSAR models were derived with CoMFA and comparative molecular similarity indices analysis (CoMSIA). The final CoMFA model ( $q^2 = 0.764$ ,  $r^2 = 0.988$ ) was calculated with an AM1-BCC charge and electrostatic parameter, whereas the CoMSIA model ( $q^2 = 0.655$ ,  $r^2 = 0.964$ ) was derived with an AM1-BCC charge and combined steric, electrostatic, hydrophobic and HB-acceptor parameters. Leave-five-out (LFO) cross-validation was also performed, which yielded good correlation coefficient for both CoMFA (0.801) and CoMSIA (0.656) models. Robustness of the developed models was checked further with 1000 run bootstrapping analyses, which gave an acceptable correlation coefficient for CoMFA (BS- $r^2 = 0.997$ , BS-SD = 0.003) and CoMSIA (BS- $r^2 = 0.996$ , BS-SD = 0.018).

**Key Words :** 3D-QSAR, P-gp Inhibitors, Partial Charges, CoMFA, CoMSIA

### Introduction

Multidrug resistance (MDR) is a phenomenon whereby tumor cells exposed to one cytotoxic agent develop cross resistance to a range of structurally and functionally unrelated compounds. The drug resistance that develops in cancer cells often results from over-expression of particular proteins, such as cell membrane transporters, which can result in an increased efflux of the cytotoxic drugs from the cancer cells, thus lowering their intracellular concentrations.<sup>1,2</sup> The cytotoxic drugs that are most frequently associated with MDR are hydrophobic, amphipathic natural products, such as the taxanes, vinca alkaloids, anthracyclines, epipodophyllotoxins, antimetabolites topotecan, dactinomycin, and mitomycin C.<sup>3-5</sup> Development of resistance to a wide spectrum of drugs occurs through a variety of mechanisms. A number of mechanisms have been proposed to explain the development of MDR. These include modulation of genes, alteration in DNA repair capacity, altered target enzyme levels, detoxification involving glutathione conjugation and activity of efflux pumps such as P-glycoprotein (P-gp).<sup>6,7</sup>

Although several mechanisms have been proposed for drug resistance, the best-studied mechanism of MDR is related to the over-expression of P-gp, a 170 kDa ATP dependent membrane transporter that acts as a drug efflux pump.<sup>8</sup> In MDR tumor cells, ATP binding cassette (ABC) proteins are simultaneously over-expressed: examples include P-gp (ABCB1), breast cancer resistance protein (BCRP,

ABCG2) and multidrug resistance protein (MRP, ABCC family).<sup>9</sup> These transporters utilize energy from ATP hydrolysis in the efflux of various substances out of cells against a concentration gradient. This active efflux decreases the intracellular concentration of the transported compounds, leading to failure of the cancer chemotherapy. P-gp is among the 49 human ABC transporters identified to date. P-gp belongs to human MDR/TAP (transporter associated protein) subfamily, and is the most intensively studied of these proteins.<sup>2</sup> P-gp is extensively distributed and expressed in the intestinal epithelium, hepatocytes, renal proximal tubular cells, adrenal gland and capillary endothelial cells comprising the blood-brain and blood-testis barriers. P-gp transports structurally diverse substances; most of them are anticancer drugs.

Since the discovery of first P-gp inhibitor (Verapamil),<sup>10</sup> many studies have sought to understand the efflux mechanism and to create selective, specific and effective inhibitors, also called MDR modulators. MDR modulators are characterized into three categories (first, second and third generation). The first generation compounds, which include cyclosporine-A, quinidine and verapamil, are already used clinically as therapeutic agents for other diseases. The high dose that is typically required is associated with problems of toxicity. The intensive search for more specific and less toxic compounds led to the development of the second and third generation of MDR inhibitors. Nowadays, it is the third generation of MDR modulators that are the focus of interest.<sup>4</sup>

In contrast to the second-generation MDR modulators, these inhibitors are not CYP3A4 substrates, and do not significantly influence the pharmacokinetic profile of co-administered drugs.<sup>11</sup> Representative third generation modulators are tariquidar, elacridar and zosuquidar.

In this paper, we compare the different partial charges in comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) prediction accuracy on adamantyl derivatives.<sup>12</sup> The present approach was ligand based; this approach is especially useful when the X-ray crystal structure of the target protein is unavailable, as is the case for human P-gp. The X-ray crystal structures of mouse P-gp proteins (PDB codes: 3G5U, 3G60 and 3G61) has been recently elucidated.<sup>13</sup> But, the low resolution structures are limited in their value in structure based designs such as docking guided quantitative structure-activity relationship (QSAR) study. Hence, the present ligand based (CoMFA and CoMSIA) approach.

CoMFA<sup>14</sup> and CoMSIA<sup>15</sup> are popular QSAR techniques, which correlate biological activity with molecular features. CoMFA and CoMSIA results are highly sensitive to settings of the steric field, charge calculation, molecular alignment, grid spacing and dimensions.<sup>16,17</sup> Studies have sought to enhance QSAR quality by adjusting these factors optimally.<sup>18-22</sup> In CoMFA and CoMSIA studies, molecular mechanics (MM) are widely used in calculating interactions energies because of speed. In contrast, quantum mechanics (QM) methods demand time and computational resource. In MM calculations, assigning proper atomic charges to the ligand atoms is a crucial step that eventually affects the quality of a QSAR model.<sup>21</sup> Many efforts have been made to develop accurate charge assignment methods. Semi-empirical and empirical methods including AM1,<sup>23</sup> AM1-BCC,<sup>24</sup> Del-Re,<sup>25,26</sup> Gasteiger-Hückel, Gasteiger-Marsilli,<sup>27,28</sup> Hückel, MMFF<sup>29</sup> and Pullman<sup>30</sup> are widely used in QSAR model development because of their convenience and speed in calculation. Hartree-Fock (HF) charges have also been considered in calculating the partial atomic charges. We did not try density function theory (DFT), believing that the 18 methods used were suitably comprehensive. However, inclusion of DFT would be good idea in future studies. Presently, we studied the effects of various charge calculation methods on the predictive accuracy of a CoMFA model with adamantyl derivatives as P-gp inhibitors.

### Computational Details

**Dataset.** The inhibitor dataset was comprised of 17 adamantyl derivatives reported as potent P-gp inhibitors.<sup>12</sup> This dataset consists of three compounds with arbitrary activity (> 15  $\mu$ M). For calculations, 15  $\mu$ M was considered as a representative value. The given EC<sub>50</sub> values for all compounds were converted to molar (M) concentration and then converted to logarithmic scale for further analysis. The biological activities and structures of 17 compounds are summarized in Table 1.

### Electrostatic Potential Assignment and Bioactive Con-

**Table 1.** Structures and biological activities of adamantyl derivatives

Compound	R	R <sup>1</sup>	EC <sub>50</sub>	pEC <sub>50</sub>
1			> 15	4.82
2			8.48	5.07
3			7.78	5.11
4			> 15	4.82
5			2.74	5.56
6			1.13	5.95
7			1.58	5.80
8			0.58	6.24
9			0.74	6.13
10			1.10	5.96
11			1.10	5.96
12			1.24	5.91
13			> 15	4.82
14			10.76	4.97
15			0.84	6.08
16			0.71	6.15
17			0.83	6.08

**former Determination.** SYBYL 8.1<sup>31</sup> software was used to calculate empirical partial charges such as Del-Re, Gasteiger-Hückel, Gasteiger-Marsilli, Hückel, MMFF94 and Pullman. Other quantum mechanical partial charges were calculated using Gaussian 03 suite of the software program.<sup>32</sup> Semi-

empirical charge calculations were performed with Hamiltonian such as AM1, AM1-BCC, INDO, MINDO, MINDO3, MNDO and PM3. We also tried various charge calculation schemes such as MPA, NPA and ESP charges. For ESP charges, the Merz-Singh-Kollman method,<sup>33</sup> RESP charges<sup>34</sup> and AM1-BCC charges<sup>35</sup> were tested. HF charges were calculated with 6-31G\* basis set. In every charge calculation scheme, atomic charges were evaluated at the fully optimized geometries at the given level of theory. The most active molecule (**8**) was considered as a standard molecule and sketched by SYBYL 8.1. Minimization was performed by Tripos force field with conjugate gradient minimizer. We optimized the geometry without consideration of any salvation. Taking the salvation issue into account, while very important, is also very onerous and beyond our present capability. A systematic search was carried out with default parameters to obtain global minimum energy conformations. This global minimum energy conformation was assumed as a bioactive conformation. For the remaining molecules, bioactive conformers were determined based on this compound. While determining bioactive conformations for other inhibitors, all the possible single bond conformations were systematically searched for variable part while keeping the scaffold moiety unchanged. These conformations for all molecules were used for CoMFA analysis. The most active compound (**8**) was considered as a reference to align remaining compounds of the dataset using atom-by-atom matching method. Figure 1 shows the alignments for Gasteiger-Hückel, MNDO, AM1 and HF charge schemes.

**CoMFA and CoMSIA Calculation Methods.** The CoMFA method was introduced by Cramer *et al.* in 1988.<sup>14</sup> For the alignment set, the steric and electrostatic CoMFA fields were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. The grid box dimensions were determined automatically in such a way that the region boundaries were extended beyond 4 Å in each direction from the co-ordinates of each molecule. The steric and electrostatic potentials are represented by the van der Waals potential and Columbic terms, respectively, and calculated by using the standard Tripos force field. The distance dependent dielectric constant of 1.00 was used. A sp<sup>3</sup> carbon with +1 charge served as a probe for calculating steric and electrostatic fields. The cutoff range for the steric and electrostatic field energies was +30 kcal/mol. The CoMSIA method was introduced by Gerard Klebe.<sup>15</sup> The CoMSIA method is an extension of the CoMFA method. It calculates steric, electrostatic, hydrophobic, hydrogen bond acceptor and hydrogen bond donor fields. The lattice dimensions were selected with a sufficiently large margin (> 4 Å) to enclose all aligned molecules. Singularities were avoided at atomic positions in CoMSIA fields. The attenuation factor ( $\alpha$ ) was set to 0.3 for CoMSIA analysis. Statistical analysis was performed in a similar manner as CoMFA. The regression analysis was carried out by using the partial least square (PLS) method.<sup>36-38</sup> To derive 3D-QSAR models, CoMFA and CoMSIA descriptors were used as an independent variable and pEC<sub>50</sub> activity value as a dependent variable. The cross-validation

analysis was performed using the leave-one-out (LOO) method. Finally, non-cross-validated analysis was performed and Pearson's correlation coefficient ( $r^2$ ) was calculated. To further analyze the robustness and statistical confidence of the models, bootstrapping analysis for 1000 runs was performed and corresponding  $r^2$  and standard error was reported.

## Results and Discussion

**Comparison of Correlation Coefficients of Different Charge Models.** We used 18 different types of charges for QSAR model development. The statistical results are summarized in Table 2. In this table, three blocks show charges calculated using different methods. The first block, empirical charges, were calculated using SYBYL software. The semi-empirical and HF charges were calculated using the Gaussian 03 quantum mechanical package. For each charge method, QSAR models were developed using steric (S), electrostatic (E) and a combination of both factors (S+E). An accurate charge calculation method would produce better prediction accuracy of a QSAR model. In the first set of charges, Hückel charge displayed better internal predictive ability ( $q^2 = 0.760$ ,  $N = 5$ ) for electrostatic charge. The non-cross-validated ( $r^2$ ) correlation coefficient for Hückel charge was 0.988. Following the Hückel charge, the MMFF94 charge showed good predictive ability ( $q^2 = 0.558$ ,  $r^2 = 0.984$ ,  $N = 5$ ) for the combined (S+E) model. Among the different charges from this set (Del-Re, Gasteiger-Hückel, Gasteiger-Marsilli, Hückel, MMFF94 and Pullman), only the Hückel and MMFF94 performed well in predicting model. However, the remaining four charges predicted  $q^2 < 0.5$  for all models (S, E and S+E). In QSAR modeling studies, Gasteiger-Hückel and Gasteiger-Marsilli charges are widely-used in drug design. However, presently, the Gasteiger-Hückel and Gasteiger-Marsilli charges did not yield a better performance. Our results are in agreement with a previous study,<sup>21</sup> in the sense that MMFF94 charges showed statistically better predictivity than the Gasteiger-Hückel and Gasteiger-Marsilli charges.

In the next set, i.e., semi-empirical charges (AM1, AM1-BCC, MINDO, MINDO3, MNDO and PM3), predictive models were obtained with AM1, AM1-BCC and MNDO charges. The best model was obtained with the AM1-BCC charge scheme. For electrostatic parameters, it produced higher  $q^2$  (0.764) and  $r^2$  (0.988) values. We used this model for final CoMFA analysis. For the combined model, good predictions ( $q^2 = 0.724$ ,  $r^2 = 0.980$ ) were obtained. Following AM1-BCC, AM1 charges performed well in prediction statistics, it also produced a higher statistical result for electrostatic and combined parameters. Our finding is in agreement with a previous study,<sup>35</sup> which suggested that an additional BCC term to the AM1 method can improve the charge assignment accuracy. Followed by AM1-BCC and AM1, MNDO charges showed good prediction ability. For the electrostatic model ( $q^2 = 0.650$ ,  $r^2 = 0.980$ ) and combined model ( $q^2 = 0.724$ ,  $r^2 = 0.980$ ), MNDO charges displayed

**Table 2.** PLS summary of CoMFA models using different electrostatic potential calculation methods

Charge Scheme	No	(S)			(E)			(S+E)		
		$q^2$	$r^2$	SEE	$q^2$	$r^2$	SEE	$q^2$	$r^2$	SEE
<i>Empirical</i>										
Gasteiger-Hückel	1	0.297(3)	0.822	0.253	0.349(6)	0.998	0.027	0.453(4)	0.988	0.067
Gasteiger-Marsilli	2	0.326(4)	0.878	0.218	-0.106(1)	0.176	0.507	0.231(4)	0.956	0.131
Huckel	3	0.379(3)	0.890	0.199	0.760(5)	0.988	0.070	0.491(3)	0.926	0.164
MMFF94	4	0.389(4)	0.850	0.242	0.419(2)	0.813	0.250	0.558(4)	0.984	0.079
Del-Re	5	0.320(4)	0.873	0.223	-0.143(1)	0.123	0.523	0.268(4)	0.963	0.121
Pullman	6	0.345(4)	0.886	0.211	0.289(4)	0.869	0.226	0.461(4)	0.987	0.071
<i>Semi-Empirical</i>										
MNDO(MPA)	7	0.258(3)	0.747	0.302	0.650(5)	0.979	0.094	0.507(3)	0.893	0.197
MINDO(MPA)	8	0.088(4)	0.905	0.192	-0.148(4)	0.833	0.256	-0.029(5)	0.987	0.076
MINDO3(MPA)	9	0.088(4)	0.905	0.192	-0.150(1)	0.464	0.409	0.001(6)	0.994	0.052
PM3(MPA)	10	0.113(2)	0.720	0.306	0.229(3)	0.919	0.171	0.262(6)	0.990	0.068
AM1(MPA)	11	0.379(3)	0.890	0.199	0.688(4)	0.977	0.095	0.663(4)	0.975	0.099
<b>AM1-BCC</b>	12	0.379(3)	0.890	0.199	<b>0.764(5)</b>	<b>0.988</b>	<b>0.070</b>	0.724(4)	0.980	0.088
<i>Hartree-Fock</i>										
HF(MPA)	13	0.467(4)	0.899	0.199	0.495(4)	0.977	0.094	0.530(4)	0.965	0.116
HF-Chelp	14	0.467(4)	0.899	0.199	0.382(5)	0.984	0.083	0.591(4)	0.974	0.101
HF-ChelpG	15	0.467(4)	0.899	0.199	0.405(4)	0.976	0.097	0.578(4)	0.971	0.107
HF(MK)	16	0.467(4)	0.899	0.199	0.462(6)	0.986	0.080	0.567(4)	0.969	0.110
HF(NPA)	17	0.467(4)	0.899	0.199	0.530(5)	0.987	0.075	0.540(6)	0.995	0.048
HF(RESP)	18	0.467(4)	0.899	0.199	0.519(6)	0.987	0.078	0.572(4)	0.969	0.109

S = steric, E = electrostatics, SE = combination of steric and electrostatic field,  $q^2$  = cross validated correlation coefficient by LOO,  $r^2$  = non-cross-validated correlation coefficient, SEE = standard error of estimate, parenthesis contains optimum number of components. MPA: Mulliken population analysis; MK: Merz-Singh-Kollman electrostatic potential derived charge, NPA: Natural population analysis. AM1-BCC: bond charge correction from AM1 Hamiltonian. RESP: Restricted electrostatic potential, STDEV: Standard deviation. For HF, 6-31G\* basis set was used. Model used for CoMFA analysis is marked in **bold font**.

good performance. A previous study on MX-compound using CoMFA indicated that MNDO charges have lower predictive ability than AM1 charges,<sup>39</sup> which is consistent with our results. The MINDO, MINDO3 and PM3 charges performed very poorly in prediction statistics, this might happen because of inaccurate assignment of charges.

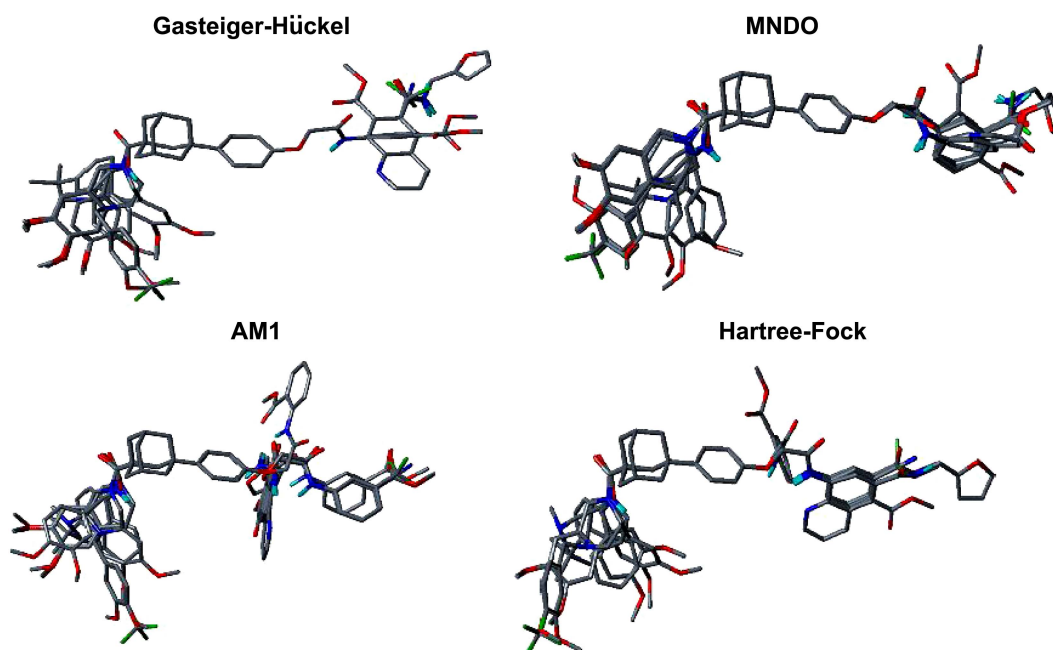
The third set of charge assignment scheme consisted of HF charges. Charges were calculated using the 6-31G\* basis set, which includes HF-Chelp, HF-ChelpG, HF-MK, HF-NPA and HF-RESP. In this charge scheme Chelp, ChelpG and MK performed slightly poor in statistical prediction than HF, HF-NPA and HF-RESP for electrostatics field. For the combined model, all charges showed good predictive ability ( $q^2 > 0.5$ ). Statistical analyses of CoMFA models indicated that all the charges are consistent.

In all superimposed structures some geometry difference was observed among the different charges scheme. The different geometry for variable parts could result because of different level of theories used for charge calculations. The AM1 method displayed quite different orientations for all aligned molecules than others (Fig. 1). These orientation differences in three dimensional spaces could be also important for variable  $q^2$  values.

**CoMFA Model.** The final CoMFA model was derived with the electrostatic field from AM1-BCC charges. A

statistical result for model is shown in Table 3. The model was selected based on lowest SDEP (0.317) and higher LOO cross-validation correlation coefficient ( $q^2 = 0.764$ ) values. Additional LFO cross-validation was performed, wherein five molecules were moved out from dataset and their activity was predicted using the remaining molecules. This produced a correlation coefficient value (0.801) indicative of the stability and predictive capacity of the model. SEE (0.070) was estimated along with non-cross-validated correlation coefficient ( $r^2 = 0.988$ ) and ANOVA test value (187.4). Further robustness of developed CoMFA model was checked with bootstrapping analysis. For the derived CoMFA model, BS- $r^2$  was 0.997 with SD = 0.003. The graph of predicted versus actual activity for CoMFA model is presented in Figure 2.

**CoMSIA Model.** Different CoMSIA models were generated using AM1-BCC charges with the combination of five descriptors and listed in Table 4. All the generated models demonstrated reasonable  $q^2$  and  $r^2$  values. The final model for CoMSIA analysis was selected based on higher  $q^2$  value and lower SDEP. The final CoMSIA model obtained with S, E, H and A parameters. The final model for CoMSIA yielded  $q^2 = 0.655$  and  $r^2 = 0.964$  with N=3. LFO=0.656; a higher value indicates a stable and predictive model. The percentage field contributions value showed that S (16.4%),

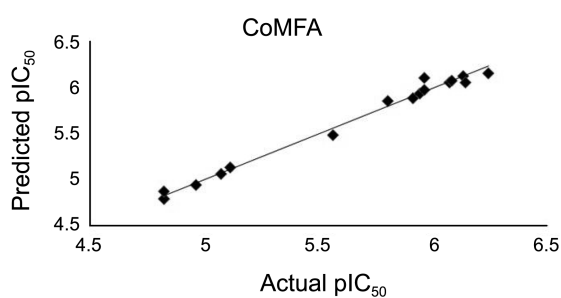


**Figure 1.** Molecular superposition of all adamantyl derivatives over template molecule (**8**). Alignments are shown for different charges scheme (Gasteiger-Hückel, MNDO, AM1 and Hartree-Fock).

**Table 3.** PLS summary of final electrostatic CoMFA model using AM1-BCC charge assignment scheme

CoMFA	N	$q^2$	SDEP	LFO	$r^2$	SEE	F	BS- $r^2$	BS-SD
E	5	0.764	0.317	0.801	0.988	0.070	187.4	0.997	0.003

E = electrostatics field, N = optimum number of components,  $q^2$  = leave one out cross-validated correlation coefficient, SDEP = leave one out standard error of prediction, LFO = leave five out cross validation correlation coefficient,  $r^2$  = non-cross-validated correlation coefficient, SEE = non-cross-validated standard error of estimate, F = ANOVA test value, BS- $r^2$  = correlation coefficient after 1000 runs of bootstrapping analysis, BS-SD = standard deviation after 1000 runs of bootstrapping analysis.



**Figure 2.** Scatter plot diagram of predicted *versus* actual activity of CoMFA (E) model.

E (32.5%), H (21.9%) and A (29.2%) fields contributing more or less for the model development. The SDEP and SEE for final CoMSIA model was found to be 0.352 and 0.114, respectively. The robustness and quality of the developed model was checked using bootstrapping analysis (BS- $r^2$  = 0.964 with SD = 0.018). The graph of predicted versus actual activity for CoMSIA model is depicted in Figure 3. The actual and predicted activities for all compounds using CoMFA and CoMSIA are listed in Table 5.

**Contour Map Analysis.** The different density maps of the steric, electrostatic, hydrophobic and HB-acceptor parameters of CoMFA and CoMSIA models are represented as 3D

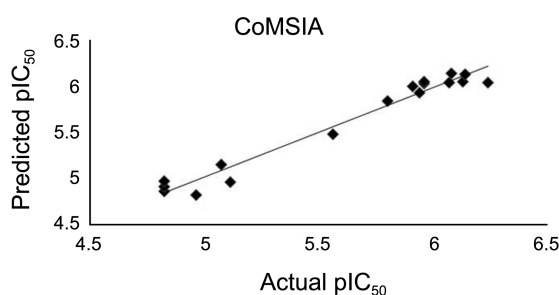
coefficient contour maps. The contour maps surround all lattice points where the QSAR is found to be strongly associated with changes in the molecular field values. According to the standard SYBYL settings, steric interactions are represented by green and yellow contours, while electrostatic interactions are represented by red and blue contours. Green contours denote points where the Lennard-Jones potential has to be increased by appropriate groups to increase the biological activity. The yellow contours are used to underline the points where such a potential has to be decreased by suitable substituent's to correlate with increased binding affinity. The electrostatic red color shows the presence of a negative charge that is expected to enhance the activity, whereas the blue color indicates positive charge that is expected to correlate with increased binding affinity. In case of the HB-acceptor field, the magenta color denotes the presence of acceptor atom that is favorable for increasing potency, whereas the red color decreases the activity. The HB-donor field is represented by cyan and purple colors, with the cyan color denotes HB-donor favorable property, while the purple color indicates disfavor for HB-donor property.

The electrostatics AM1-BCC charge model was used for CoMFA contour map analysis. CoMFA electrostatic contour map is depicted in Figure 4. Template molecule **8** (pEC<sub>50</sub>

**Table 4.** PLS regression summary of different CoMSIA models using AM1-BCC charge scheme

CoMSIA	N	$q^2$	SDEP	LFO	$r^2$	SEE	F	BS- $r^2$	BS-SD	S	E	H	A	D
S	4	0.366	0.478	0.434	0.794	0.273	16.68	0.868	0.075	100	-	-	-	-
E	6	0.583	0.442	0.612	0.993	0.057	239.5	0.997	0.003	-	100	-	-	-
H	2	0.554	0.386	0.589	0.922	0.161	82.79	0.914	0.037	-	-	100	-	-
D	2	0.510	0.405	0.508	0.686	0.324	15.32	0.742	0.183	-	-	-	-	100
A	6	0.551	0.459	0.624	0.943	0.163	27.62	0.972	0.023	-	-	-	100	-
SE	3	0.583	0.388	0.586	0.940	0.147	68.04	0.933	0.035	42.4	57.6	-	-	-
EH	5	0.620	0.402	0.638	0.987	0.074	168.9	0.996	0.004	-	46.4	53.6	-	-
HD	2	0.567	0.381	0.600	0.849	0.225	39.38	0.887	0.058	-	-	47	-	53
DA	3	0.456	0.443	0.472	0.792	0.274	16.50	0.884	0.065	-	-	-	48.6	51.4
SH	3	0.575	0.391	0.590	0.891	0.198	35.55	0.946	0.031	42.5	-	57.5	-	-
SD	3	0.576	0.391	0.564	0.863	0.222	27.23	0.923	0.041	46.9	-	-	-	53.1
ED	2	0.514	0.403	0.541	0.811	0.252	29.96	0.872	0.062	-	45.7	-	-	54.3
EA	6	0.538	0.465	0.592	0.989	0.073	145.0	0.994	0.006	-	53.1	-	46.9	-
SEH	5	0.581	0.422	0.646	0.984	0.082	138.4	0.994	0.005	23.7	35.5	40.8	-	-
SEA	5	0.638	0.393	0.635	0.984	0.083	135.2	0.991	0.007	29.3	34.6	-	36.1	-
SED	3	0.572	0.393	0.587	0.912	0.178	44.65	0.955	0.022	28.5	32.7	-	-	38.8
EHA	6	0.651	0.404	0.622	0.991	0.064	191.3	0.997	0.004	-	38.5	30.3	31.2	-
EHD	3	0.574	0.392	0.590	0.945	0.141	74.29	0.967	0.020	-	32.9	35.7	-	31.4
SHA	3	0.631	0.432	0.645	0.956	0.122	100.9	0.962	0.022	26.5	-	37.2	36.3	-
SHD	3	0.604	0.377	0.602	0.931	0.158	58.13	0.957	0.027	26.9	-	36.8	-	36.3
EDA	3	0.488	0.429	0.482	0.860	0.225	26.62	0.934	0.048	-	36.7	-	31.4	31.9
SDA	3	0.592	0.384	0.627	0.873	0.214	29.87	0.931	0.038	28.5	-	-	33.7	37.8
SEHD	3	0.599	0.380	0.622	0.949	0.136	79.90	0.966	0.021	19.2	24.3	28.3	-	28.2
<b>SEHA</b>	<b>3</b>	<b>0.655</b>	<b>0.352</b>	<b>0.656</b>	<b>0.964</b>	<b>0.114</b>	<b>116.5</b>	<b>0.964</b>	<b>0.018</b>	<b>16.4</b>	<b>32.5</b>	<b>21.9</b>	<b>29.2</b>	-
SEDA	3	0.572	0.392	0.560	0.897	0.192	37.93	0.943	0.033	21.7	24.4	-	25.4	28.5
EHAD	3	0.574	0.392	0.607	0.935	0.153	62.73	0.956	0.019	-	21.9	29.7	23.8	24.6
SHDA	3	0.619	0.371	0.626	0.931	0.158	58.39	0.960	0.021	18.7	-	27.4	27.8	26.1
SEHAD	3	0.605	0.377	0.634	0.943	0.144	71.39	0.963	0.022	15.2	18.3	23.1	23.1	20.3

S = steric, E = electrostatics, H = hydrophobic, A = HB-acceptor, D = HB-donor, N = optimum number of components,  $q^2$  = leave one out cross-validated correlation coefficient, SDEP = leave one out standard error of prediction, LFO = leave five out cross validation correlation coefficient,  $r^2$  = non-cross-validated correlation coefficient, SEE = non-cross-validated standard error of estimate, Fvalue = ANOVA test value, BS- $r^2$  = correlation coefficient after 1000 runs of bootstrapping analysis, BS-SD = standard deviation after 1000 runs of bootstrapping analysis. Final model chosen for analysis is highlighted in bold font.

**Figure 3.** Scatter plot diagram of predicted *versus* actual activity of CoMSIA (SEHA) model.

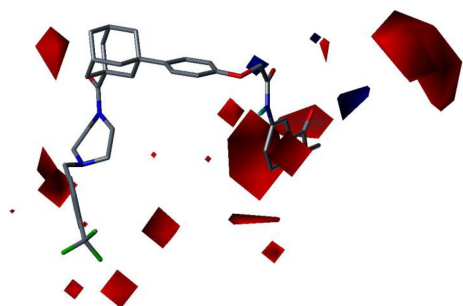
= 6.24) was displayed in the background of contour maps. A red contour was observed near the R<sup>1</sup> substituents, indicating that electronegative substituents are favorable for inhibitory potency. In case of compound **12** (pEC<sub>50</sub> = 5.91), the para substituted ester group was favorable for inhibitory activity and exhibited the corresponding big red contour. A small blue contour was observed near R<sup>1</sup> substituents indicating the favorability of the electropositive substituents for

inhibitory potency. In compound **15** (pEC<sub>50</sub> = 6.08) and **16** (pEC<sub>50</sub> = 6.15) side chain at meta position (R<sup>1</sup> substituents) showed improved binding affinity. The meta substituents, i.e., -CN and -CF<sub>3</sub> are slightly electropositive in nature and favorable for improved inhibitory potency. A small red polyhedra observed around the -CF<sub>3</sub> group of R substitution indicated that the electron withdrawing nature of the -CF<sub>3</sub> group was favorable for the inhibitory effect.

The CoMSIA steric contour map is shown in Figure 5(a). Green and yellow contours were observed around the ligand molecule. The big green contour map was observed near the R substituents over piperazine ring, indicating the favorability of the sterically bulky substituents for inhibitory potency. In compound **5** (pEC<sub>50</sub> = 5.56) and **6** (pEC<sub>50</sub> = 5.95), the substituents 3,4-dimethoxybenzylamine and 6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline showed increased inhibitory potency. This indicated that conformationally rigid and bulky substituents were more advantageous in terms of higher inhibitory activity, indicated by the green contour in the vicinity. A big yellow contour was observed around the R<sup>1</sup> substituents, indicating the disfavor of the sterically bulky

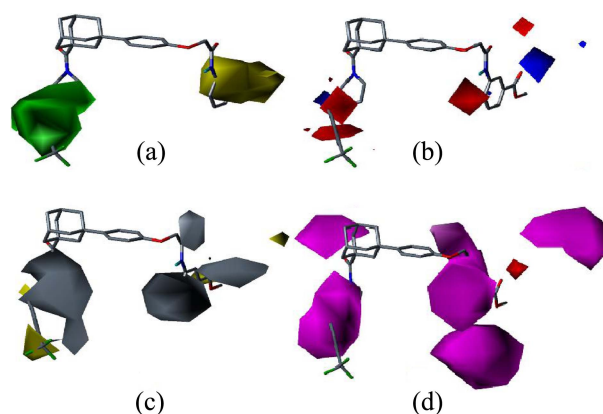
**Table 5.** Actual and predicted activities for CoMFA and CoMSIA model along with residual values

Compound	pEC <sub>50</sub>	Predicted Activity		Residual Activity	
		CoMFA	CoMSIA	CoMFA	CoMSIA
1	4.82	4.88	4.97	-0.06	-0.15
2	5.07	5.07	5.15	0.00	-0.08
3	5.11	5.14	5.14	-0.03	0.15
4	4.82	4.80	4.86	0.02	-0.04
5	5.56	5.49	5.48	0.07	0.08
6	5.95	5.94	5.93	0.00	0.01
7	5.80	5.86	5.84	-0.06	-0.04
8	6.24	6.16	6.04	0.08	0.20
9	6.13	6.13	6.05	0.00	0.08
10	5.96	6.11	6.03	-0.15	-0.07
11	5.96	5.98	6.05	-0.02	-0.09
12	5.91	5.89	6.00	0.02	-0.09
13	4.82	4.80	4.91	0.02	-0.09
14	4.97	4.95	4.82	0.02	0.15
15	6.08	6.06	6.04	0.02	0.04
16	6.15	6.06	6.13	0.09	0.02
17	6.08	6.08	6.14	0.00	-0.06

**Figure 4.** Electrostatics stdev\* coefficient contour map for highly active compound (8) generated by CoMFA analysis, where blue/red polyhedra depicts the favorable site for positively/negatively charged groups.

groups for inhibitory effect. In compound 14 (pEC<sub>50</sub> = 4.97), the bulky substituent at R<sup>1</sup> position was unfavorable for inhibitory activity, resulting in a lower potency of compound. This may occur because of the steric interactions of ligand at the receptor pocket.

The CoMSIA electrostatics contour map is shown in Figure 5(b). A blue contour map was observed over the R substituents, indicating that electropositive groups were favorable for inhibitory activity. In compound 2 (pEC<sub>50</sub> = 5.07) and 3 (pEC<sub>50</sub> = 5.11), the electropositive R substituents showed moderate potency for these inhibitors. Three small-sized red contours were observed around the R substituents, showed the importance of electronegative property for the inhibitory potency. In case of compounds 5 (pEC<sub>50</sub> = 5.56) and 6 (pEC<sub>50</sub> = 5.95), the R substituents containing dimethoxy group is favorable for inhibitory potency and corresponding favorable positions are identified by red contours. Another small-sized contour was found near an amide carbonyl group, indicating its importance for inhibitory

**Figure 5.** Stdev\* coefficient contour maps for highly active compound (8) generated by CoMSIA analysis. (a) Steric favored/disfavored areas are shown in green/yellow color. (b) Electro-positive and electro-negative favored/disfavored areas are shown in blue/red color. (c) Hydrophobic favored/disfavored areas are shown in yellow/white color. (d) HB-acceptor favored/disfavored areas are shown magenta/red color.

activity. A medium-sized blue contour map was identified near the R<sup>1</sup> substituents, indicating the favorability of the electropositive groups for P-gp inhibitory activity. In compound 17 (pEC<sub>50</sub> = 6.08), the electropositive 8-amino-quinolinyl (R<sup>1</sup>) substituent showed increased inhibitory potency, indicating the contribution of electropositive substituents towards inhibitory potency.

Another important parameter in the CoMSIA model was hydrophobicity; the contour map is depicted in Figure 5(c). A small yellow contour map was observed on the -CF<sub>3</sub> group (R substituents) of 8, indicating the importance of the hydrophobic nature of the -CF<sub>3</sub> group for inhibitory activity. Replacement of the small -CF<sub>3</sub> group by -OCH<sub>3</sub> group in compound 9 (pEC<sub>50</sub> = 6.13) preserved the inhibitory activity, as shown by the corresponding yellow contour. A large white contour was observed over the piperazine ring of R substituents of 8, indicating that hydrophilic groups were unfavorable for inhibitory effect. For example, compound 2 (pEC<sub>50</sub> = 5.07) and 3 (pEC<sub>50</sub> = 5.11) containing hydrophilic substituents were unfavorable for inhibitory potency, which was evident by the inhibitory activity values. Another white contour was observed near the R<sup>1</sup> substituents, indicating that replacement of ester with amide group was unfavorable for inhibitory activity. The R<sup>1</sup> amide substituent containing compound 13 (pEC<sub>50</sub> = 4.82) showed the least inhibitory potency. In case of compound 14 (pEC<sub>50</sub> = 4.97), the amide group extended with furfuryl displayed a weaker inhibitory potency, which corresponded to a white contour.

CoMSIA HB-acceptor contour map is displayed in Figure 5(d). A large magenta contour was observed in the vicinity to the -CO group near the adamantyl ring system, indicating that H-bond acceptor nature of -CO group was favorable for P-gp inhibitory activity. Another magenta contour was observed over the R substituents covering the piperazinyl substructure indicating that the protonable tertiary nitrogen was favorable for inhibitory activity. Tertiary nitrogen con-

taining compounds **6**, **7**, **8**, **9** and **10** displayed higher inhibitory potency, which corresponded to the magenta contour. Compounds containing a dimethoxy group at the R position were favorable as a H-bond acceptor and showed a corresponding magenta contour. Compounds with dimethoxy substituents (**5**, **6**, **15**, **16** and **17**) at R position displayed higher inhibitory potency. A small red contour was observed near the R<sup>1</sup> substituents indicating that a HB-donating group was not favorable for inhibitory activity at this position. The 3-amido substituted compounds **13** and **14** displayed lower inhibitory potency, because both contained a H-bond donating group, corresponding to the red contour. The big magenta contour at right corner indicated that R1 substituents penetrating this direction were favorable for inhibitory potency. In compound **16**, the H-bond accepting nature of -CF<sub>3</sub> group was favorable for inhibitory potency, which was indicated by the corresponding magenta contour.

Generally, in drug design and 3D-QSAR studies, the most widely used charge scheme is Gasteiger-Hückel. It should be noted that in our study on single dataset (**17** compounds), Gasteiger-Hückel did not yield the best performance on predictions. In contrast, a semi-empirical AM1-BCC charge scheme showed a better performance over the other charges used in this study. Our results demonstrate that semi-empirical charge assigning methods such as AM1-BCC and AM1 yielded accurate models than the routinely used Gasteiger-Hückel charge method. This result is consistent with previously published work of Mittal *et al.*,<sup>21</sup> and Choo *et al.*<sup>40</sup> AM1-BCC charges showed slightly better statistics than AM1 charges, which is consistent with the previously published work of Tsai *et al.*<sup>41</sup> Our results demonstrate significance of partial charge calculation methods on prediction accuracy.

### Conclusions

In this study, we performed a comparison of 18 charges scheme on prediction accuracy of 3D-QSAR models and the best CoMFA results were obtained with semi-empirical AM1-BCC and AM1 methods. In contrast, the Gasteiger-Hückel method, which is commonly used in drug design and 3D-QSAR modeling, displayed lower prediction ability than the semi-empirical methods. The additional BCC correction to AM1 method gave superior results than the AM1 method. HF charges displayed consistency in prediction accuracy for combined models. If there is no facility to calculate semi-empirical charges, MMFF94 charge could be a better alternative. MMFF94 charge showed better predictive ability than the traditional Gasteiger-Hückel charges. For our developed QSAR model, steric, electrostatics, hydrophobic and HB-acceptor parameters displayed importance. Contour map analyses indicated that, at the R position sterically bulky substitutions are favorable for inhibitory activity. The positively ionizable tertiary nitrogen at R position is important for activity and showed corresponding favorable magenta contour in the vicinity. At the R<sup>1</sup> position, the ester group was well-tolerated for inhibitory activity and HB-

acceptor substituents had better effect on inhibitory activity. The information obtained from contour map analysis could be used to design more potent inhibitor based on an adamantyl scaffold.

**Acknowledgments.** Computational resources were supported by PLSI Supercomputing Resources of KISTI (Korea Institute of Science and Technology Information). This study was supported by research funds from Chosun University 2010.

### References

1. Gottesman, M. M.; Fojo, T.; Bates, S. E. *Nat. Rev. Cancer* **2002**, *2*, 48.
2. Ambudkar, S. V.; Dey, S.; Hrycyna, C. A.; Ramachandra, M.; Pastan, I.; Gottesman, M. M. *Annu. Rev. Pharmacol. Toxicol.* **1999**, *39*, 361.
3. Avendano, C.; Carlos, M. J. *Curr. Med. Chem.* **2002**, *9*, 159.
4. Thomas, H.; Coley, H. M. *Cancer Control* **2003**, *10*, 159.
5. Stavrovskaya, A. A. *Biochemistry* **2000**, *65*, 95.
6. Advani, S. H. *Ranbaxy Science Foundation Fourth Annual Symposium*; November, 1997; p 47.
7. Choi, C. H. *Cancer Cell Int.* **2005**, *5*, 30.
8. Gottesman, M. M.; Pastan, I. *Annu. Rev. Biochem.* **1993**, *62*, 385.
9. Sarkadi, B.; Homolya, L.; Szakacs, G.; Varadi, A. *Physiol. Rev.* **2006**, *86*, 1179.
10. Tsuruo, T.; Iida, H.; Yamashiro, M.; Tsukagoshi, S.; Sakurai, Y. *Biochem. Pharmacol.* **1982**, *31*, 3138.
11. Krishna, R.; Mayer, L. D. *Eur. J. Pharm. Sci.* **2000**, *11*, 265.
12. Min, K. H.; Xia, Y.; Kim, E. K.; Jin, Y.; Kaur, N.; Kim, E. S.; Kim, D. K.; Jung, H. Y.; Choi, Y.; Park, M. K. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5376.
13. Aller, S. G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P. M.; Trinh, Y. T.; Zhang, Q.; Urbatsch, I. L. *Science* **2009**, *323*, 1718.
14. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
15. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
16. Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 457.
17. Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 538.
18. Tervo, A. J.; Nyroenen, T. H.; Roenkko, T.; Poso, A. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 807.
19. Peterson, S. D.; Schaal, W.; Karlén, A. *J. Chem. Inf. Model.* **2006**, *46*, 355.
20. Mittal, R. R.; McKinnon, R. A.; Sorich, M. J. *J. Mol. Model.* **2008**, *14*, 59.
21. Mittal, R. R.; Harris, L.; McKinnon, R. A.; Sorich, M. J. *J. Chem. Inf. Model.* **2009**, *49*, 704.
22. Tsai, K. C.; Chen, Y. C.; Hsiao, N. W.; Wang, C. L.; Lin, C. L.; Lee, Y. C.; Li, M.; Wang, B. *Eur. J. Med. Chem.* **2010**, *45*, 1544.
23. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
24. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.
25. Del Re, G.; Pullman, B.; Yonezawa, T. *Biochim. Biophys. Acta* **1963**, *75*, 153.
26. Del Re, G. *J. Chem. Soc.* **1958**, 4031.
27. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
28. Gasteiger, J.; Marsili, M. *Tetrahedron Lett.* **1980**, *19*, 3181.
29. Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520.
30. Berthod, H.; Giessner-Prettre, C.; Pullman, A. *Theor. Chim. Acta* **1967**, *8*, 212.
31. S. H. R. SYBYL 8.1; Tripos Inc., St. Louis, MO 63144 USA.
32. Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel,



- H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Wallingford CT, 2004.
33. Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comp. Chem.* **1990**, *11*, 431.
34. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.
35. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comp. Chem.* **2002**, *23*, 1623.
36. Cramer, R. D. *Perspect. Drug Discovery Des.* **1993**, *1*, 269.
37. Wold, S.; Sjöström, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109.
38. Dunn, W. J.; Wold, S.; Edlund, V.; Hellberg, S.; Gasteiger, J. *Quant. Struct.-Act. Relat.* **1984**, *3*, 131.
39. Cecil, N.; Antti, P.; Kari, T.; Jukka, G. *Quant. Struct.-Act. Relat.* **1996**, *15*, 189.
40. Park Choo, H. Y.; Choi, S.; Jung, S. H.; Koh, H. Y.; Pae, A. N. *Bioorg. Med. Chem.* **2003**, *11*, 4585.
41. Tsai, K. C.; Wang, S. H.; Hsiao, N. W.; Li, M.; Wang, B. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3509.
-