QSAR STUDY OF A SERIES OF CHOLESTERYL ESTER TRANSFER PROTEIN INHIBITORS

Mahesh T. CHHABRIA[a1], Bhanubhai N. SUHAGIA[a2], Appaji B. MANDHARE[b] and Pathik S. BRAHMKSHATRIYA[a3,*]

[a] *Department of Pharmaceutical Chemistry, L M College of Pharmacy, Navrangpura, Ahmedabad-380 009, Gujarat, India; e-mail: [1] mahesh.chhabria@rediffmail.com, [2] bnsuhagia@rediffmail.com, [3] pathik26@yahoo.com*

[b] *Torrent Research Centre, Torrent Pharmaceuticals Ltd., Medicinal Chemistry-Drug Discovery, Gandhinagar-382 428, Gujarat, India; e-mail: appajimandhare@torrentpharma.com*

Received January 25, 2011
Accepted April 11, 2011
Published online June 15, 2011

Cholesteryl ester transfer protein (CETP), an enzyme which catalyses the transfer of cholesteryl ester from HDL to VLDL, is a promising target for discovery of novel anti-hyperlipidemic agents due to its pivotal role in HDL metabolism and reverse cholesterol transport. Quantitative structure activity relationship study of a series of CETP inhibitors was carried out using genetic function approximation to study various structural requirements for CETP inhibition. Various lipophilic, electronic, geometric and spatial descriptors were correlated with CETP inhibitory activity. Developed models were found predictive as indicated by their good r^2_{pred} values and satisfactory internal and external cross-validation results. Study reveals that lipophilicity (ClogP), with parabolic relationship, contributed significantly to the activity along with some electronic, geometric and quantum mechanical descriptors. The present study can be applied to future lead optimization of CETP inhibitors.
Keywords: QSAR; Genetic function approximation; Molecular descriptors; Predictive models, Lack-of-fit; Medicinal chemistry; Drug design; Structure-activity relationships.

Atherosclerosis is the underlying condition of coronary artery disease (CAD). Elevated blood cholesterol levels are one of the most common risk factors of CAD apart from other risk factors[1]. Lipid lowering therapies include the blockbuster statins, cholesterol absorption inhibitor ezetimibe and fibrates. The routine use of statins has been questioned due to their varied chronic or acute toxicities[2]. Moreover, withdrawals of some blockbuster statin drugs also question their routine use in the management of hyperlipidemia[3]. Plasma cholesteryl ester transfer protein (CETP) is a hydrophobic glycoprotein that is secreted mainly from the liver and that circulates in plasma and is bound mainly to high density lipoprotein (HDL). It

© 2011 Institute of Organic Chemistry and Biochemistry
doi:10.1135/cccc2011019

promotes the redistribution of cholesteryl esters, triglycerides and to a lesser extent, phospholipids between plasma lipoproteins. CETP catalyses transfer of cholesteryl ester from HDL to low density lipoprotein (LDL)/very low density lipoprotein (VLDL) in exchange for triglyceride[4,5]. Although various dyslipidemias have been linked with increased CETP concentrations[6–8], it is possible that elevated CETP is the result of dyslipidemia rather than its cause[9]. By transferring cholesteryl esters from HDL to VLDL and LDL, CETP reduces the concentration of antiatherogenic HDL cholesterol while increasing the concentration of LDL cholesterol[10]. CETP inhibitors are thus thought to increase the concentration of HDL-cholesterol levels and thus are expected to aid in controlling dyslipidemia. Varieties of small molecule CETP inhibitors are under experimental studies and few are in clinical trials[11]. Recently, dalcetrapib (JTT-705) has been reported to be safe and well-tolerable in a 48-week double-blind study[12] in contrast with the withdrawal of torcetrapib in the ILLUMINATE study[13].

Quantitative structure activity relationship (QSAR) study is a very useful tool in the era of modern drug discovery to get better insights into structure activity relationships[14–17]. During the process, the behavior of QSAR models developed is examined with a variety of statistical parameters and the contribution of various descriptors is analyzed. Various methods have been developed for statistical analysis of generated QSAR models. Multiple Linear Regression (MLR) and Partial Least Squares (PLS) are one of the most popular statistical methods. PLS regression technique is especially useful in the cases where the number of descriptors (independent variables) is comparable to or greater than the number of compounds (data points) and/or there exist other factors leading to correlations between variables[18]. A comprehensive review of comparison of MLR, PLS and GA-MLR has been published[19]. In this communication, we describe results of QSAR studies carried out on a series of CETP inhibitors using genetic function approximation (GFA) technique[20].

GFA algorithm offers a new approach to build structure-activity models. It automates the search for QSAR models by combining a genetic algorithm with statistical modeling tools. Thousands of candidate models are created and tested during evolution; only the superior models survive, and are then used as "parents" for the creation of the next generation of candidate models. GFA has been successfully applied for the generation of variety of QSAR models[21–23]. Such model provides structure-activity insights, which can be used for designing of new compounds and activity prediction prior to synthesis.

EXPERIMENTAL

Data Set

In present studies, a series of substituted 1,2,4-triazole derivatives, reported by Sikorski et al.[24] as potential CETP inhibitors, was selected. Thirty seven compounds were randomly divided into training and test sets, the former set consisting of 28 compounds and the remaining 9 compounds were taken in the test set. Structures of all the compounds used for QSAR analysis and their CETP enzyme inhibition activity (IC_{50}, micromolar concentrations, μM) are given in Table I. For every compound of the series, the experimental values of biological activity (pIC_{50}) are used in the negative logarithmic scale. The structures of all compounds used in this study were sketched by using Visualizer module of Discovery studio 2.1 software (Accelrys Inc., USA)[25]. CHARMM force field was used for the calculation of potential energy. An energy minimization of all the compounds was done by using Smart Minimizer method until the root mean square gradient value becomes smaller than 0.001 kcal/mol Å. Further, optimized structures for all compounds were aligned with compound **1** and these structures were used for calculation of various descriptors.

Descriptor Calculation

Various two-dimensional (structural, thermodynamic and quantum mechanical) and three-dimensional (steric, electronic and geometric) physicochemical descriptors were calculated using calculate molecular properties protocol of the Discovery Studio 2.1. Theoretical ClogP calculations were carried out using WINDOWS based ClogP software program (version 4.0, BioByte Corp, Claremont, CA). A correlation matrix of the molecular descriptors was prepared and highly correlated descriptors with a correlation value of 0.9 or above were removed from the study. Remaining descriptors were used to develop QSAR models. Descriptors included to develop QSAR models are listed and described in Table II.

Regression Analysis

The data set was modeled using the genetic function approximation (GFA) technique to generate a population of equations rather than one single equation for correlation between biological activity and descriptors. GFA is genetics based method of variable selection, which combines Holland's genetic algorithm with Friedman's multivariate adaptive regression splines to evolve the population of equations that best fit the training set data.

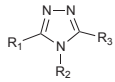
The GFA method works in the following way: A particular number of equations (set at 100 by default in the Discovery studio 2.1 software) are generated randomly. Then pairs of "parent" equations are chosen randomly from this set of 100 equations and "crossover" operations were performed at random. The number of crossing over was set by default at 5000. The goodness of each progeny equation is assessed by Friedman's lack of fit (LOF) score, which is given by following formula

$$\mathrm{LOF} = \mathrm{LSE}/[1 - (c + dp)/m]^2$$

where LSE is the least-squares error, c is the number of basis functions in the model, d is smoothing parameter, p is the number of descriptors and m is the number of observations in the training set. The smoothing parameter, which controls the scoring bias between equations of different sizes, was set at default value of 0.5 and GFA crossover of 5000 were set to

TABLE I

Chemical structures and biological activity of training set (**1–28**) and test set (**29–37**) compounds



Compound	R_1	R_2	R_3	IC_{50}, μM
1	n-$C_{13}H_{27}$–	3-CH_3O-C_6H_4–	–SH	2
2	n-$C_{13}H_{27}$–	2-CH_3-C_6H_4–	–SH	4
3	n-$C_{13}H_{27}$–	2-CH_3O-C_6H_4–	–SH	7
4	n-$C_{13}H_{27}$–	3-CH_3-C_6H_4–	–SH	7
5	n-$C_{13}H_{27}$–	cyclohexyl–	–SH	7
6	n-$C_{13}H_{27}$–	4-F-C_6H_4–	–SH	8
7	n-$C_{13}H_{27}$–	4-C_6H_5O-C_6H_4–	–SH	8
8	n-$C_6H_{13}CC(CH_2)_5$–	3-CH_3O-C_6H_4–	–SH	8
9	n-$C_{13}H_{27}$–	3-F-C_6H_4–	–SH	9
10	n-$C_{13}H_{27}$–	3,4-$(OCH_2O)C_6H_3$–	–SH	10
11	n-$C_{13}H_{27}$–	4-CH_3-C_6H_4–	–SH	10
12	n-$C_{13}H_{27}$–	3-CF_3-C_6H_4–	–SH	10
13	n-$C_{13}H_{27}$–	4-Cl-2-CH_3-C_6H_3–	–SH	10
14	n-$C_{13}H_{27}$–	2-CH_3S-C_6H_4–	–SH	10
15	n-$C_{13}H_{27}$–	4-$C_6H_5CH_2O$-C_6H_4–	–SH	15
16	n-$C_{13}H_{27}$–	2-naphthyl–	–SH	30
17	$CH_3(CH_2)_6S(CH_2)_5$–	3-CH_3O-C_6H_4–	–SH	15
18	$HCC(CH_2)_{11}$–	3-CH_3O-C_6H_4–	–SH	15
19	n-$C_{13}H_{27}$–	4-CH_3-3-Cl-C_6H_3–	–SH	20
20	$CH_3(CH_2)_{10}OCH_2$–	3-CH_3O-C_6H_4–	–SH	20
21	4-(n-$C_{10}H_{21}$benzyl)–	3-CH_3O-C_6H_4–	–SH	4.5
22	4-(n-C_9H_{19}benzyl)–	3-CH_3O-C_6H_4–	–SH	6
23	4-(n-C_7H_{15}benzyl)–	3-CH_3O-C_6H_4–	–SH	6
24	4-(n-C_8H_{17})$C_6H_4OCH_2$–	3-CH_3O-C_6H_4–	–SH	15
25	4-(n-C_7H_{15})$C_6H_4OCH_2$–	3-CH_3O-C_6H_4–	–SH	20
26	4-(n-C_8H_{17})$C_6H_4(CH_2)_3$–	3-CH_3O-C_6H_4–	–SH	4
27	4-(n-C_7H_{15})$C_6H_4(CH_2)_3$–	3-CH_3O-C_6H_4–	–SH	7.5
28	4-(n-C_6H_{13})$C_6H_4(CH_2)_3$–	3-CH_3O-C_6H_4–	–SH	20
29	n-$C_{13}H_{27}$–	3-Cl-C_6H_4–	–SH	5
30	n-$C_{13}H_{27}$–	2-Cl-C_6H_4–	–SH	10
31	n-$C_{13}H_{27}$–	4-Cl-C_6H_4–	–SH	15
32	n-$C_{13}H_{27}$–	4-CF_3-C_6H_4–	–SH	20
33	n-$C_{13}H_{27}$–	C_6H_5–	–SH	30
34	n-$C_{13}H_{27}$–	3-$C_6H_5CH_2O$-C_6H_4–	–SH	40
35	$CH_3(CH_2)_{10}SCH_2$–	3-CH_3O-C_6H_4–	–SH	40
36	n-$C_{13}H_{27}$–	2-CH_3CH_2O-C_6H_4–	–SH	50
37	4-(n-C_6H_{13})$C_6H_4OCH_2$–	3-CH_3O-C_6H_4–	–SH	25

give reasonable convergence. The length of equation was fixed to seven terms, the population size was established as 100, the equation term was set to linear polynomial functionality and the mutation probability was specified as 0.1. The best equations, out of the 100 equations, were chosen based on the statistical parameters such as regression coefficient (r), adjusted regression coefficient (r_{adj}), regression coefficient cross validation (r_{cv}) and F-test values.

TABLE II
List of descriptors used in the study

No.	Descriptor	Definition
1	CLogP	Log of the octanol–water partition coefficient
2	Total_Energy	Total energy by VAMP AM1 calculation
3	Dipole_mag	Magnitude of dipole moment
4	Jurs_RNCG	Charge of most negative atom divided by the total negative charge
5	Jurs_RPCS	Solvent-accessible surface area of the most positive atom divided by relative positive charge
6	Jurs_WNSA_3	Surface weighted atomic partial negative surface area
7	Chi_V_1	Valence first order connectivity index

Validation Test

To check the intercorrelation of descriptors, variance inflation factor (VIF) analysis was performed. VIF value is calculated from $1/(1 - r^2)$, where r^2 is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. VIF value larger than 10 signals towards chance-correlation and hide the information of descriptors by intercorrelation of descriptors[26].

It has been shown that optimal statistical characteristics (high values of r and F, low values of s and LOF) need not be the proof of a highly predictive model[27]. Hence, in order to evaluate the predictive ability of the QSAR model, we used the method described by Roy et al.[27]. External predictability of the model was determined by calculating the value of predictive r^2 (r^2_{pred}) using the following equation

$$r^2_{pred} = 1 - \frac{\sum (Y_{Pred(test)} - Y_{Obs(test)})^2}{\sum (Y_{Obs(test)} - \overline{Y}_{training})^2}$$

RESULTS AND DISCUSSION

In the present study, we have screened 16 preselected descriptors for 28 CETP inhibitors using GFA method. The choice of descriptors was based on the type of interaction between the inhibitors and the enzyme (principally in-

volving hydrophobic, electronic and steric interactions). The present series was selected for the structure-activity relationship study due to its remarkable structural similarity with dalcetrapib (JTT-705). Moreover, the most potent compound showed IC_{50} value of 2 μM against the reference standard of 6 μM for JTT-705 [28]. Thus, the present study would definitely be of great value for a plausible optimization of lead using QSAR approach to find more potent analogs. Initially, 100 QSAR equations were generated that consist of six descriptors among QSAR random models. However, finally the results of the best five models are given in Table III along with their regression statistics.

For a statistically significant model, it is necessary that the descriptors evolved in the equation should not be inter-correlated with each other. The intercorrelation of the descriptors used in the selected models (Table IV) was very low. The correlation matrix for the used descriptors is shown in Table IV. To further check the intercorrelation of descriptors, variance inflation factor (VIF) analysis was performed. The VIF values of these descriptors were found to be 7.588 (ClogP), 9.185 (Total_Energy), 1.895 (Dipol_mag),

TABLE III
Selected QSAR equations and their regression statistics

Eq.	Description	r^2	r^2_{adj}	r^2_{cv}	r^2_{pred}	LOF	F
(*1*)	pIC_{50} = –2.45069 + 2.07188 * ClogP – 0.662325 * Jurs_RPCS + 0.00042879 * Total_Energy_VAMP – 0.11094 * $(ClogP)^2$ – 0.0174428 * $Dipole\_mag^2$ + 20.0114 * $(Jurs\_RNCG)^2$	0.655	0.566	0.548	0.707	0.125	6.646
(*2*)	pIC_{50} = –2.69336 + 2.05384 * ClogP + 6.31237 * Jurs_RNCG – 0.723074 * Jurs_RPCS + 0.000451291 * Total_Energy_VAMP – 0.109915 * $(ClogP)^2$ – 0.0182247 * $(Dipole\_mag)^2$	0.653	0.554	0.546	0.720	0.126	6.588
(*3*)	pIC_{50} = –4.13841 + 2.50117 * ClogP + 7.03533 * Jurs_RNCG – 0.729423 * Jurs_RPCS – 0.0411908 * Jurs_WNSA_3 + 0.000778378 * Total_Energy_VAMP – 0.13285 * $(ClogP)^2$	0.651	0.550	0.543	0.712	0.127	6.512
(*4*)	pIC_{50} = –2.16309 + 1.9469 * ClogP – 0.0750192 * Dipole_mag + 5.9353 * Jurs_RNCG – 0.709883 * Jurs_RPCS + 0.000440991 * Total_Energy_VAMP – 0.103836 * $(ClogP)^2$	0.650	0.550	0.541	0.697	0.127	6.498
(*5*)	pIC_{50} = –2.2679 – 0.0977765 * CHI_V_1 + 2.10787 * ClogP – 0.599758 * Jurs_RPCS + 0.000331998 * Total_Energy_VAMP – 0.110668 * $(ClogP)^2$ + 21.9536 * $(Jurs\_RNCG)^2$	0.634	0.542	0.523	0.737	0.133	6.053

4.019 (Jurs_RNCG), 2.259 (Jurs_RPCS), 9.624 (Jurs_WNSA_3) and 2.962 (Chi_V_1). All the VIF values were found to be less than 10. Thus, from the VIF analysis, it is clear that the descriptors used in the final models have low intercorrelation.

The models were also evaluated for their capacity to predict the activity of training set and test set compounds, i.e., internal and external cross-validation, respectively. The results for the Eq. (*1*) are summarized in Tables V and VI. The models displayed satisfactory values of predicted r^2 (r^2_{pred}). For all the models, r^2_{pred} was found to be in the accepted range[27].

As expected, lipophilicity of compounds emerged as an indispensable descriptor for CETP inhibition along with other structural, spatial and electronic descriptors. Positive contribution of lipophilic parameter (ClogP) in the equations is consistent with the fact that for inhibition of CETP, a lipophilic compound is required to compete with cholesteryl esters, which are highly lipophilic natural substrates for the enzyme. Also, compounds with very high ClogP values, such as compound **15**, **16** and **19** showed less activity as evidenced from the parabolic relationship observed in the QSAR equations. Looking at the parabolic correlation of ClogP with the activity, a ClogP value of 7.00 was found to be optimum to show good bioavailability.

Jurs descriptors are a group of molecular descriptors which combine shape and electronic information to characterize molecules[29]. These descriptors are calculated by mapping atomic partial charges on solvent-accessible surface areas of individual atoms. Jurs_RPCS is positive charge surface area, i.e., solvent-accessible surface area of the most positive atom divided by relative positive charge. Jurs_WNSA_3 is surface weighted

TABLE IV
Correlation matrix of the descriptors used in the equations

	pIC_{50}	ClogP	Total_ Energy	Dipole_ mag	Jurs_ RNCG	Jurs_ RPCS	Jurs_ WNSA_3	Chi_V_1
pIC_{50}	1							
ClogP	0.198	1						
Total_Energy	0.138	–0.219	1					
Dipole_mag	–0.256	0.246	0.129	1				
Jurs_RNCG	0.101	–0.628	–0.155	–0.328	1			
Jurs_RPCS	0.015	–0.087	0.566	–0.074	0.214	1		
Jurs_WNSA_3	0.244	0.254	0.795	0.439	–0.358	0.440	1	
Chi_V_1	0.018	0.648	–0.486	0.230	–0.097	–0.155	–0.075	1

TABLE V
Observed and predicted pIC_{50} values for training set compounds (as per Eq. (*1*))

Compound	pIC_{50} (observed)	pIC_{50} (predicted)	Residual
1	5.699	5.489	0.210
2	5.398	5.214	0.184
3	5.155	5.202	–0.047
4	5.155	5.013	0.142
5	5.155	5.327	–0.172
6	5.097	4.990	0.107
7	5.097	4.881	0.216
8	5.097	4.993	0.104
9	5.046	4.958	0.088
10	5.000	4.977	0.023
11	5.000	5.011	–0.011
12	5.000	4.979	0.021
13	5.000	5.098	–0.098
14	5.000	4.994	0.006
15	4.824	4.983	–0.159
16	4.824	4.918	–0.094
17	4.824	4.840	–0.016
18	4.824	4.934	–0.010
19	4.699	4.942	–0.243
20	4.699	4.636	0.063
21	5.347	5.285	0.062
22	5.222	5.294	–0.072
23	5.222	5.121	0.101
24	4.824	4.780	0.044
25	4.699	4.669	0.030
26	5.398	5.255	0.143
27	5.125	5.227	–0.102
28	4.699	5.120	–0.421

charged partial surface area. Jurs_RPCS and Jurs_WNSA_3 showed a negative contribution towards the biological activity. This means that the charge distribution within the molecules acts as the driving force for intermolecular interactions and the lesser the relative charge the larger the interactions. The above fact is exemplified from compounds **16** and **19** (R_2 = 2-naphthyl and 4-CH_3-2-Cl-C_6H_3, respectively), where higher values of Jurs_RPCS and Jurs_WNSA_3 resulted in decrease in activity. Notably, most of the molecules which had a smaller R_2 (3-OCH_3-C_6H_4) stood out as potent molecules due to low values of Jurs_RPCS and Jurs_WNSA_3 proving importance of bulkiness at R_2. Jurs_RNCG is relative negative charge, i.e., charge of most negative atom divided by the total negative charge. An increase in the value of Jurs_RNCG resulted in increase in the CETP inhibitory activity. Thus, compounds with a lower value of total negative charge are more likely to be active. Compound **19**, with 4-CH_3-2-Cl group, had a higher value of total negative charge (lower Jurs_RNCG) and hence possibly is less active. Moreover, the coefficient for Jurs_RNCG descriptor was found to be the highest among all descriptors' coefficient values. This indicates that Jurs_RNCG is the most significant descriptor that shows correlation to CETP inhibitory activity followed by ClogP, which also showed a significantly high correlation coefficient. Another important observation was drawn from the effect of Chi_V_1 descriptor on biological activity. Chi_V_1, valence first order connectivity index, is a topological descriptor, which also showed negative contribution towards biological activity which

TABLE VI
Observed and predicted pIC_{50} values for test set compounds (as per Eq. (*1*))

Compound	pIC_{50} (observed)	pIC_{50} (predicted)	Residual
29	5.301	4.996	0.305
30	5.000	5.002	–0.002
31	4.824	5.017	–0.193
32	4.699	5.009	–0.310
33	4.523	4.808	–0.285
34	4.398	4.537	–0.139
35	4.398	4.807	–0.409
36	4.301	4.556	–0.255
37	4.602	4.390	0.212

indicates that molecules with bulkier substituents are less likely to show activity. On the other hand, total energy of the molecules, with a low value of correlation coefficient, contributed positively to the biological activity.

Authors greatly acknowledge University Grants Commission, Govt. of India, for the financial support. One of the authors (P.S.B.) thanks CSIR for financial support through award of Senior Research Fellowship.

REFERENCES

1. Chhabria M. T., Brahmkshatriya P. S.: *Curr. Opin. Invest. Drugs* **2006**, *7*, 820.
2. Brahmkshatriya P. S., Chhabria M. T., Jani M. H.: *J. Enzyme Inhib. Med. Chem.* **2006**, *21*, 1.
3. Omar M. A., Wilson J. P.: *Ann. Pharmacother.* **2002**, *36*, 288.
4. Tall A. R.: *J. Lipid Res.* **1993**, *34*, 1255.
5. Barter P. J., Brewer H. B., Chapman M. J., Hennekens C. H., Rader D. J., Tall A. R.: *Arterioscler. Thromb. Vasc. Biol.* **2003**, *23*, 160.
6. De Grooth G. J., Smilde T. J., Van Wissen V., Klerkx A. H. E. M., Zwinderman A. H., Fruchart J. C., Kastelein J. J. P., Stalenhoef A. F. H., Kuivenhoven J. A.: *Atherosclerosis* **2004**, *173*, 261.
7. McPherson R., Mann C. J., Tall A. R., Hogue M., Martin L., Milne R. W., Marcel Y. L.: *Arterioscler. Thromb. Vasc. Biol.* **1991**, *11*, 797.
8. Moulin P., Appel G. B., Ginsberg H. N., Tall A. R.: *J. Lipid Res.* **1992**, *33*, 1817.
9. Barter P. J., Kastelein J. J.: *J. Am. Coll. Cardiol.* **2006**, *47*, 492.
10. de Grooth G. J., Klerkx A. H., Stroes E. S., Stalenhoef A. F. H., Kastelein J. J. P., Kuivenhoven J. A.: *J. Lipid Res.* **2004**, *45*, 1967.
11. Chhabria M. T., Suhagia B. N., Brahmkshatriya P. S.: *Recent Patents Cardiovasc. Drug Discov.* **2007**, *2*, 214.
12. Stein E. A., Roth E. M., Rhyne J. M., Burgess T., Kallend D., Robinson J. G.: *Eur. Heart J.* **2010**, *31*, 480.
13. Tall A. R., Yvan-Charvet L., Wang N.: *Arterioscler. Thromb. Vasc. Biol.* **2007**, *27*, 257.
14. Lin L., Lin W. O., Jiang J. H., Shen G. L., Yu R. Q.: *Eur. J. Pharm. Sci.* **2005**, *25*, 245.
15. Si H. Z. , Wang T., Zhang K. J., Hua Z. D., Fan B. T.: *Bioorg. Med. Chem.* **2006**, *14*, 4834.
16. Sivakumar P. M., Geethababu S. K., Doble M.: *Chem. Biol. Drug Res.* **2008**, *71*, 447.
17. Lü J. X., Shen Q., Jiang J. H., Shen G. L., Yu R. Q.: *J. Pharm. Biomed. Anal.* **2004**, *35*, 679.
18. Höskuldsson A.: *J. Chemometrics* **1988**, *2*, 211.
19. Saxena A. K., Prathipati P.: *SAR QSAR Environ. Res.* **2003**, *14*, 433.
20. Rogers D., Hopfinger A. J.: *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854.
21. Kulkarni V. M., Karki R. J.: *Bioorg. Med. Chem.* **2001**, *9*, 3153.
22. Cho D. H., Lee S. K., Kim B. T., No K. T.: *Bull. Korean Chem. Soc.* **2001**, *22*, 388.
23. Roy K., Bhattacharya P.: *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3737.
24. Sikorski J. A.: (Searle and Co.) WO1999/14204.
25. *Discovery Studio 2.1.* Accelrys Inc., 6985 Scranton Road, San Diego, CA, USA.
26. Jaiswal M., Khadikar P. V., Scozzafava A., Supuran A.: *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3283.

27. Roy P. P., Roy K.: *QSAR Comb. Sci.* **2007**, *27*, 302.
28. Shinkai H., Maeda K., Yamasaki T., Okamoto H., Uchida I.: *J. Med. Chem.* **2000**, *43*, 3566.
29. Stanton D. T., Jurs P. C.: *Anal. Chem.* **1990**, *62*, 2323.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.