Bioorganic & Medicinal Chemistry Letters 20 (2010) 3089–3093

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

3D-QSAR study of benzene sulfonamide analogs as carbonic anhydrase II inhibitors

Kalyan K. Sethi [a,*], Saurabh M. Verma [b], Naru Prasanthi [b], Suvendu K. Sahoo [a], Rabi N. Parhi [a], P. Suresh [a]

[a] GITAM Institute of Pharmacy, GITAM University, Rushikonda, Visakhapatnam 530 045, AP, India
[b] Department of Pharmaceutical Sciences, Birla Institute of Technology, Mesra, Ranchi 835 215, India

ARTICLE INFO

Article history:
Received 17 December 2009
Revised 11 March 2010
Accepted 27 March 2010
Available online 2 April 2010

Keywords:
Benzene sulfonamides
Carbonic anhydrase inhibitors
3D-QSAR
CoMFA
CoMSIA

ABSTRACT

Three-dimensional quantitative structure–activity relationship (3D-QSAR) studies were performed for a series of carbonic anhydrase inhibitors using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) techniques. The large set of 37 different aromatic/heterocyclic sulfonamides carbonic anhydrase (CA, EC 4.2.1.1) inhibitors, such as CA II chosen for the present study. The conventional ligand-based 3D-QSAR studies were performed based on the low energy conformations employing database alignment rule. The ligand-based model gave q^2 values 0.538 and 0.527 and r^2 values 0.974 and 0.971 for CoMFA and CoMSIA, respectively, and the predictive ability of the model was validated. The predicted r^2 values are 0.565 and 0.502 for CoMFA and CoMSIA, respectively. Results indicate that the CoMFA and CoMSIA models could be reliable model which may be used in the design of novel carbonic anhydrase inhibitors as leads.

Several thousand different aromatic and heterocyclic sulfonamides carbonic anhydrase (CA, EC 4.2.1.1) inhibitors were synthesized in the last 50 years in the search of diverse pharmacological agents but the number of amino acid/oligopeptide derivatives among them is unexpectedly small.[1] Consequently, a large series of sulfonamides compounds were synthesized by Supuran and investigated for their inhibitory activity against physiologically relevant CA isozyme, such as CA II.[1] The synthesis involved the reaction of aromatic/heterocyclic sulfonamides containing amino, imino, hydrazine or hydroxyl groups with *N-tert*-butoxycarbonyl-γ-aminobutyric acid (Boc-GABA) in the presence of carbodimide derivatives.[2] Extending a QSAR[3–9] to drug other than those used to formulate it is always a new hypothesis, and although these extensions are often successful.

The large set of 37 sulfonamides, chosen for the present study, is presented in Figure 1, common template (Fig. 2) and Table 1. The total set of compounds was randomly divided into the training set of 29 compounds and the test set of 8 compounds (labeled with *). The three-dimensional structures of benzene sulfonamide derivatives were constructed by using SYBYL program version 7.1 on a Silicon Graphic workstation.[10] Energy minimizations were performed using the Tripos force field with a distance-dependent dielectric and conjugate gradient method.[11] The convergence criterion was 0.01 kcal/mol Å. The Gasteiger–Huckel charges were assigned.

The template molecule is most bioactive compound (KM4) in the training set was taken (Fig. 3a) were used as the reference molecule to align all the molecules using the 'DATABASE ALIGNMENT' option in SYBYL 7.1.[12] The aligned molecules are shown in Figure 3b.

Steric and electrostatic interactions were calculated using sp^3 carbon atom and a +1 charge as steric and electrostatic probes, respectively, with Tripos force field. The CoMFA grid spacing was

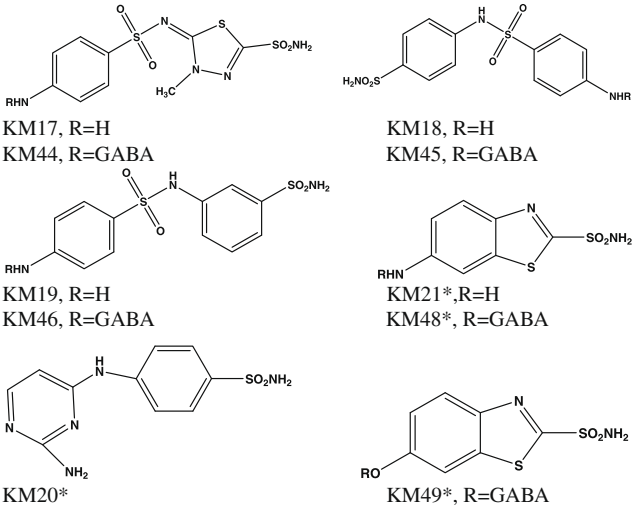

Figure 1. Structure of sulfonamides used in present data set.

* Corresponding author. Tel.: +91 9160636049/9299052717.
E-mail address: kalyansethi@gmail.com (K.K. Sethi).

0960-894X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2010.03.104

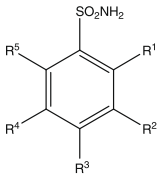

Figure 2. Common template of database molecules.

Table 1
Dataset molecules

Compd code	R^1	R^2	R^3	R^4	R^5
KM1	NH_2	H	H	H	H
KM2	H	NH_2	H	H	H
KM3	H	H	NH_2	H	H
KM4	H	H	$NHNH_2$	H	H
KM5	H	H	CH_2NH_2	H	H
KM6	H	H	$CH_2CH_2NH_2$	H	H
KM7	H	F	NH_2	H	H
KM8	H	Cl	NH_2	H	H
KM9	H	Br	NH_2	H	H
KM10	H	I	NH_2	H	H
KM11*	H	SO_2NH_2	Cl	Cl	NH_2
KM12	H	SO_2NH_2	NH_2	H	Cl
KM16*	H	H	CH_2CH_2OH	H	H
KM24	H	H	CH_2OH	H	H
KM25	H	H	CH_2CH_3	H	H
KM28	NH-GABA	H	H	H	H
KM29	H	NH-GABA	H	H	H
KM31	H	H	NHNH-GABA	H	H
KM32	H	H	CH_2NH-GABA	H	H
KM33	H	H	CH_2CH_2NH-GABA	H	H
KM34	H	F	NH-GABA	H	H
KM35	H	Cl	NH-GABA	H	H
KM36	H	Br	NH-GABA	H	H
KM37	H	I	NH-GABA	H	H
KM39	H	SO_2NH_2	NH-GABA	H	Cl
KM51*	H	H	CH_2O-GABA	H	H
KM52*	H	H	CH_2CH_2-GABA	H	H

GABA = $H_2NCH_2CH_2CH_2CO$–.
* Test set molecule.

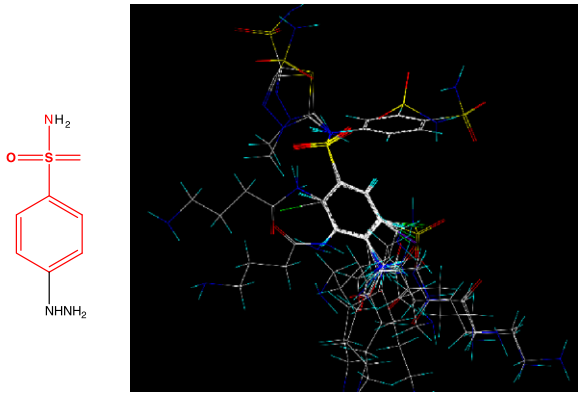

Figure 3. Alignment of the training set: a, is the common template structure (in red color) of highest bioactive molecule KM4; b, is the alignment of the training set for ligand-based model.

2.0 Å in the *x*, *y*, and *z* directions. The default value 30 Kcal/mol was set as the maximum steric and electrostatic energy cut off. Minimum-sigma (column filtering) was set to be 2.0 Kcal/mol.[13]

The CoMSIA[14] method defines five fields: steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor, which were calculated at each lattice interactions of a regularly spaced grid of 2.0 Å. A probe atom with radius 1.0 Å, +1 charge, hydrophobicity +1.0, and H-bond donor and acceptor properties of +1.0 was used to calculate steric, electrostatic, hydrophobic, and H-bond donor and acceptor fields.[15] A distance-dependent GAUSSIAN type functional form will be taken into account abrupt changes of potential energy near the molecular surface. The default value of 0.3 was used as the attenuation factor.[16]

The benzene sulfonamide ring system, found in all sulfonamide synthetic novel analogs. Such analogs have been developed that have high affinity for the human carbonic anhydrase enzyme. 3D-QSAR studies may be used for the further design of novel human carbonic anhydrase inhibitors. Lead optimization is envisaged with the help of CoMFA and CoMSIA techniques for obtaining a robust model for potent hCAII inhibitors.

CoMFA and CoMSIA are computer programs that are particularly effective in correlating the 3D structures of the molecules and their bioactivities based on statistical techniques.[17]

In order to test the predictive power of the CoMFA and CoMSIA models obtained by the training set, other eight inhibitors that were not included in the training set were used as a test set (Table 3). Table 2 summarizes the predicted results obtained from the CoMFA and CoMSIA models. The plots of predicted versus actual binding affinities for the test set inhibitors are shown in Figures 4a and b which represent models based on CoMFA and CoMSIA, respectively. By comparison of the experimentally observed and theoretically predicted pK_i values of inhibitory activity of a series of benzene sulfonamides derivatives, it can be seen that both CoMFA and CoMSIA models performed well in the prediction of the activities of the test inhibitors (Table 3). In almost all the cases, the predicted values were close to the observed pK_i values, deviating by less than small logarithm unit. Especially for KM4, both CoMFA and CoMSIA models gave ideally predictive values, both of which are less than 0.2 log units.

The predictive ability of each analysis was determined from the test set of eight compounds that were not included in the training set. These molecules were aligned in the same way as those in the training set and their activities were predicted by each PLS analysis. The predictive correlation coefficient (r^2 predicted) is defined as

$$r^2_{\text{pred}} = (\text{SD} - \text{PRESS})/\text{SD}$$

Where SD is the sum of squared deviations between the biological activities of the test set and mean activity of the training set mole-

Table 2
Summary of CoMFA and CoMSIA model results

Components	Ligand-based model	
	CoMFA	CoMSIA
		SEHDA
q^2	0.538	0.527
r^2	0.974	0.971
n	6	6
F-values	139.583	122.58
SEE	0.107	0.110
r^2 predicted	0.565	0.502
Field contribution		
Steric	2.044	0.535
Electrostatic	2.426	0.963
Hydrophobic	–	0.931
Donor	–	0.868
Accepter	–	0.652

q^2: LOO cross-validated correlation coefficient; r^2: non-cross-validated correlation coefficient; n: number of components used in the PLS analysis; SEE: standard error estimate; *F*-value: *F*-statistic for the analysis.

Table 3
Experimental and predicted pK_i, residuals by CoMFA and CoMSIA models in the training and test set (with*) with ligand-based methods

Compd code	pK_i exp	CoMFA		CoMSIA	
		PV[a]	RV[b]	PV[a]	RV[b]
KM1	2.470	2.416	0.054	2.503	−0.032
KM2	2.380	2.411	−0.031	2.398	−0.018
KM3	2.477	2.143	0.334	2.121	0.326
KM4	2.505	2.517	−0.012	2.534	−0.029
KM5	2.230	2.251	−0.021	2.231	−0.001
KM6	2.204	2.173	0.031	2.223	−0.019
KM7	1.778	1.685	0.093	1.883	−0.105
KM8	2.041	1.853	0.188	1.876	0.165
KM9	1.602	1.808	−0.206	1.741	−0.139
KM10	1.845	1.841	0.004	1.747	0.098
KM11*	1.447	0.854	0.593	1.104	0.343
KM12	1.875	1.841	0.034	1.853	0.022
KM16*	0.301	2.080	−1.779	1.959	−1.658
KM17	0.778	0.808	−0.030	0.744	0.034
KM18	0.778	0.728	0.050	0.764	0.014
KM19	0.954	0.925	0.029	0.977	−0.023
KM20*	1.079	2.060	−0.981	2.041	−0.962
KM21*	0.954	2.225	−1.271	2.239	−1.285
KM24	2.097	2.181	−0.084	2.164	−0.067
KM25	2.041	2.047	−0.003	2.058	−0.017
KM28	2.295	2.356	0.039	2.293	0.002
KM29	2.260	2.312	−0.052	2.271	−0.011
KM31	2.326	2.324	0.002	2.297	0.029
KM32	1.505	1.466	0.039	1.454	0.051
KM33	1.447	1.568	−0.121	1.554	−0.107
KM34	1.000	0.927	0.073	1.004	−0.004
KM35	1.491	1.526	−0.035	1.510	−0.019
KM36	1.477	1.468	0.009	1.473	0.004
KM37	1.431	1.470	−0.039	1.440	−0.009
KM39	0.954	0.963	−0.09	0.970	−0.016
KM44	0.602	0.586	0.014	0.582	0.020
KM45	1.146	1.186	−0.040	1.131	0.015
KM46	1.114	1.186	−0.072	1.128	−0.014
KM48*	0.699	2.318	−1.619	1.871	−1.172
KM49*	1.732	1.325	0.325	1.372	0.360
KM51*	1.863	1.573	0.290	1.466	0.397
KM52*	1.820	2.218	−0.398	2.0378	−0.2178

[a] Predictive values.
[b] Residual values.
* Test set molecule.

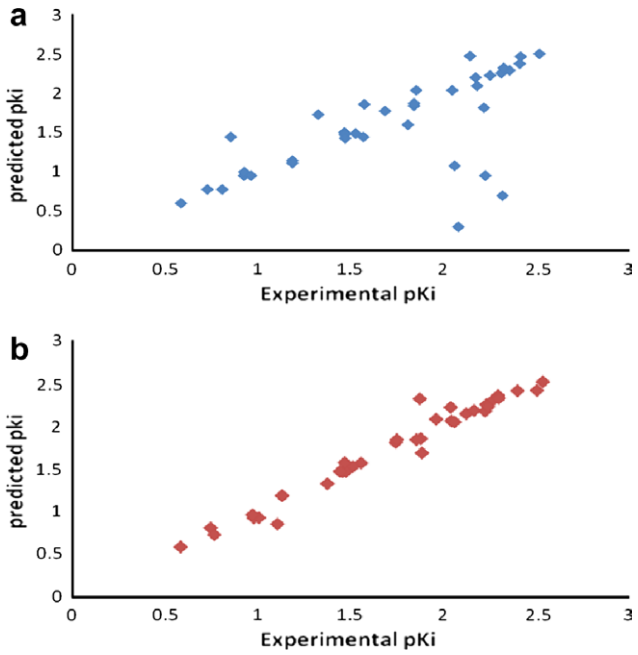

Figure 4. Predicted versus actual binding affinities for the 37 hCAII inhibitors in the set for (a) CoMFA and (b) CoMSIA models. The predicted correlation coefficients (r^2 pred) are 0.565 for the CoMFA model and 0.502 for the CoMSIA model.

cules, and PRESS is the sum of squared deviation between predicted and actual activity values for every molecule in the test set.

The predicted correlation coefficients (r^2 pred) are 0.565 for the CoMFA model and 0.502 for SEHDA (steric, electrostatic, hydrophobic, donor, acceptor) the CoMSIA model (Table 2). All results including statistical data from this predictive validation analysis further confirmed the robustness of the obtained models in terms of their predictive abilities. There was very little significant difference between the CoMFA and CoMSIA models for this predictive validation test.

Table 2 shows that the steric fields and electrostatic fields did not give the same contribution, accounting for 2.044 and 2.426, respectively, which suggests that electrostatic fields are easy in explaining the inhibition potency of these molecules and the generated CoMFA models explain well the variations between molecules having differences in steric and electrostatic interactions. More or less steric fields have nearly same contribution as electrostatic field.

Figure 5a shows the steric contour map for the CoMFA models with the highly active inhibitor KM4 (pK_i 2.505) as a reference. Three regions at benzene ring positions of KM4 have been identified with green polyhedra, which indicate that bulky substituents at these positions may improve the activities. The aromatic moiety at benzene ring positions of KM4 was surrounded by green polyhedral lying side by side of the aromatic ring, respectively. So addition of a bulky group at this position is favorable to the binding affinities.

Five CoMSIA fields (Figs. 6 and 7) steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor potentials were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. A probe atom with radius 1.0 Å, +1.0 charge, hydrophobicity of +1.0 and hydrogen-bond donor and acceptor properties of +1.0 was used to calculate steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor fields. The contribution from each one of these descriptors was truncated above 0.3 Kcal/mol.

A 3D-QSAR study using CoMFA and CoMSIA methods had been applied for the first time to a series of benzene sulfonamides as human carbonic anhydrase inhibitors (hCAII). As a result, the 3D-QSAR models with the accessible software package SYBYL 7.1 were successfully constructed. Both models from CoMFA and CoMSIA are satisfactory according to the statistical validation results as well as the contour map analysis. The CoMFA model was obtained with LOO cross-validation q^2 and non-cross validated r^2 values of 0.538 and 0.974, respectively. The best CoMSIA model with the most statistically significant correlation result has been generated with the cross-validated coefficient q^2 of 0.527, and the conventional correlation coefficient r^2 of 0.971. The standard error of estimate and the *F*-test value of this model were 0.110 and 122.58, respectively. All of the constructed models possessed good internal and external consistency and showed statistical significance and predictive abilities. Both the predictive evaluation and the contour map analysis accorded well with the experimental interaction mode. A combined application of the obtained CoMFA and CoMSIA models was further employed for the design of new hCAII inhibitors. Several structurally novel compounds were synthesized and evaluated. 3D-QSAR to drugs other than those used to formulate it is always a new hypothesis, and although these extensions are often successful, it should be no cause for surprise if they break down in particular cases. The experimental results from the biological assay indicated that the constructed models were reliable enough to be applied in both the rational design and library screening. This is the first study which evidenced by means of

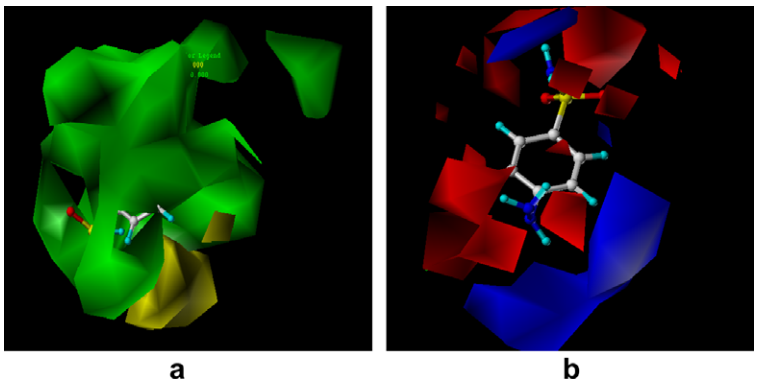

Figure 5. Standard coefficient contour maps of final CoMFA analysis with grid spacing in combination with KM4. (a) Steric contour map. Green contours (80% contribution) refer to sterically favored regions and yellow counters (20% contribution) are sterically unfavored. (b) Electrostatic contour map. Blue contours (80% contribution) refer to regions where positively charged substituents are favored; red contours (20% contribution) indicate regions where negatively charged substituents are favored.

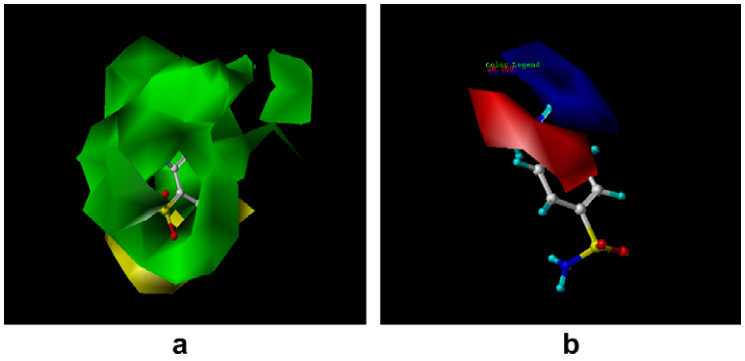

Figure 6. Standard coefficient contour maps of final CoMSIA analysis with grid spacing in combination with KM4. (a) Steric contour map. Green contours (80% contribution) refer to sterically favored regions and yellow counters (20% contribution) are sterically unfavored. (b) Electrostatic contour map. Blue contours (80% contribution) refer to regions where positively charged substituents are favored; red contours (20% contribution) indicate regions where negatively charged substituents are favored.

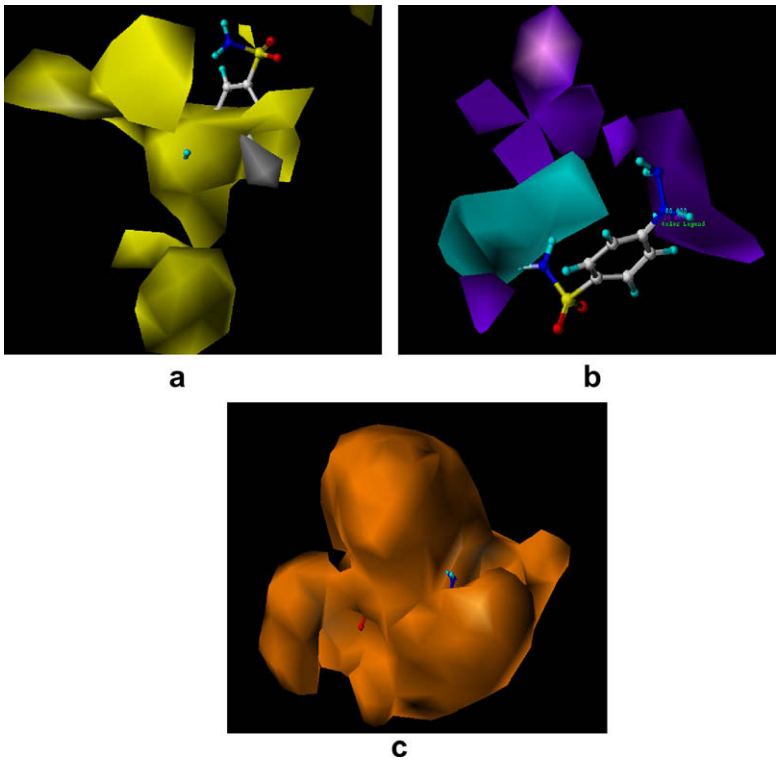

Figure 7. Standard coefficient contour maps of final CoMSIA analysis with 2 Å grid spacing in combination with KM4. (a) Hydrophobic contour map. White contours (20% contribution) refer to regions where hydrophilic substituents are favored; yellow contours (80% contribution) indicate regions where hydrophobic substituents are favored. (b) Hydrogen-bond donor contour map. Cyan contours (80% contribution) indicate regions where hydrogen-bond acceptor groups on the receptor increase activity. (c) Hydrogen-bond acceptor contour map. Orange contours majorly refer to areas where hydrogen-bond donors on the receptor promote the affinity.

3D-QSAR calculations isozyme-specific features of the isozymes CA II, in their interaction with sulfonamide inhibitors.

Acknowledgments

This work is acknowledged to our late vice-chancellor of BIT Mesra, Ranchi, Late Dr. S. K. Mukharjee, who has given all the opportunity of software and instruments. I am also greatly indebted to Central Instrumentation Facility (CIF), BIT Mesra, Ranchi, India for giving the software facility.

References and notes

1. Supuran, C. T. *Nat. Rev. Drug Disc.* **2008**, *7*, 168.
2. Mincione, G.; Menabuoni, L.; Briganti, F.; Mincione, F.; Scozzafava, A.; Supuran, C. T. *Eur. J. Pharmacol. Sci.* **1999**, *9*, 185.
3. Clare, B. W.; Supuran, C. T. *Expert Opin. Drug Metab. Toxicol.* **2006**, *2*, 113.
4. Clare, B. W.; Supuran, C. T. *Bioorg. Med. Chem.* **2005**, *14*, 2197.
5. Supuran, C. T.; Clare, B. W. *Eur. J. Med. Chem.* **1995**, *30*, 687.
6. Supuran, C. T.; Clare, B. W. *Eur. J. Med. Chem.* **1998**, *33*, 489.
7. Supuran, C. T.; Clare, B. W. *Eur. J. Med. Chem.* **1999**, *34*, 41.
8. Supuran, C. T.; Scozzafava, A. *J. Enzyme Inhib.* **1997**, *12*, 37.
9. Agrawal, V. A.; Singh, J.; Khadikar, P. V.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2044.
10. SYBYL 7.1, Tripos Associates Inc., 1699 S. Hanley Rd., St. Louis, MO 631444, USA.
11. Clark, M.; Cramer, R. D.; Opdenbosch, N. V. *J. Comput. Chem.* **1989**, *10*, 982.
12. Parra-Delgado, H.; Compadre, C. M.; Apan, T. R.; Fambuena, M. J. M.; Compadre, R. L.; Wegman, P. O.; Vazquez, M. M. *Bioorg. Med. Chem.* **2006**, *14*, 1889.
13. Zeng, J.; Liu, G.; Liu, L.; Tang, Y.; Jiang, H. *J. Mol. Model.* **2007**, *13*, 993.
14. Klebe, G.; Supuran, C. T.; Weber, A.; Bohm, M.; Scozzafava, A.; Sotriffer, C. A. *J. Chem. Inf. Model.* **2006**, *46*, 2737.
15. Avila, C. M.; Romeiro, N. C.; Silva, G. M. S. D.; Anna, S. M. R.; Barreiro, E. J.; Fraga, A. M. C. *Bioorg. Med. Chem.* **2006**, *14*, 6874.
16. Pan, X.; Tan, N.; Zeng, G.; Han, H. J.; Huang, H. *Bioorg. Med. Chem.* **2006**, *14*, 2771.
17. Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 457.