ScienSage

Journal of Advanced Scientific Research

Available online through http://www.sciensage.info

ISSN
0976-9595
Research Article

QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP STUDIES OF β-CARBOLINE 3-(SUBSTITUTED-CARBOHYDRAZIDE) DERIVATIVES FOR ANTITUMOR ACTIVITY

Avneesh Awasthy, Om Prakash Agrawal*

Patel College of Pharmacy, Madhyanchal Professional University, Ratibad, Bhopal, Madhya Pradesh, India

**Corresponding author: om11agra85@gmail.com*

ABSTRACT

β-Carboline moieties are important structural subunits which occur as components of many biologically interesting molecules for antitumor activity. Quantitative structure-activity relationship (QSAR) studies have been performed on β-carboline derivatives to explore the structural necessities for antitumor activity. 2D QSAR studies were done using VALSTAT drug designing module to explain the structural requirements for the anti-tumor activity. The 2D-QSAR was performed using the Step Wise K Nearest Neighbour Molecular Field Analysis [(SW) kNNMFA] technique with the partial least-square (PLS) method on a database. Obtained best 2DQSAR model having high predictive ability with q2 =0.743, r2= 0.721, pred_r2 =0.708 and standard error =0.346, explaining the majority of the variance in the data with partial least square (PLS) components. The results of the present study may be useful on the designing of more potent compounds as antitumor drugs.

Keywords: QSAR, β-Carboline, 2D-QSAR, VALSTAT, PLS.

1. INTRODUCTION

Synthetic and naturally occurring compounds containing the β-carboline nucleus possess a large spectrum of important pharmacological properties, including potent antitumor activity [1-5]. The potential of β-carboline compounds as anticancer agents have stimulated studies into their synthesis and structure-activity-relationship (SAR) with an aim to the improvement of their antitumor potential [6-8]. SAR studies on a variety of synthetic β-carboline derivatives have demonstrated that the introduction of appropriated substituents into position-1,-2,-3 and -9 of the β-carboline skeleton resulted in more potent antitumor derivatives, with reduced toxicity. The anticancer mode of action of these alkaloids has been also widely investigated [9-11]. Multiple mechanisms, such as DNA intercalation and inhibition of Topoisomerases I and II, IkB kinase (IKK), cyclin-dependent kinases (CDKs), mitogen activated protein kinase-activated protein kinase 2 (MK-2), polo-like kinase (PLK1) and kinesin-like protein Eg5 were pointed out from these investigations [12-14].

2. MATERIAL AND METHODS

A series of β-carboline derivatives was selected from a reported article which presented the synthesis of novel derivatives of this compound and tested their anti-cancer potential against various cancer cell lines [15, 16]. Structure build-up, physico-chemical property determination, and sequential multiple regression analysis was performed on the reported series [17].

2.1. Biological Activity Calculation

The observed potency (IC_{50} values) against renal cancer cell line (786-O) for all 26 compounds were altered from micromolar concentration to molar concentration and subsequently these values for renal cell lines from the reported series [N-(substituted-benzylidene) β-carboline-3-carbohydrazide derivatives [8] were used to derive the biological activity values in the form of (Log $1/IC_{50}$). Although, the series presented a total of 51 compounds, but about twenty-five compounds which were shown having the IC_{50} values greater than 100 micro-molar concentration (>100), were eliminated because their IC_{50} values were not exactly defined. These structures along with their activity (Log $1/IC_{50}$) values are mentioned in the table 1.

2.2. Structure Building & Energy Minimization

The structures of the remaining twenty-six compounds were fabricated by means of Chemdraw Ultra 7.0.1 of

Chem-office Ultra 7.0.1 suite software, which is a product of Cambridge soft corporation, U.S.A. These structures were then saved in MDL (.mol) format which is followed by energy minimization using Chem3D ultra 7.0.1 by the means of MM_2 (Molecular Mechanics) force fields and followed by MOPAC-Closed shell (AM-1) pro force fields using 0.100 as root mean square gradient.

2.3. Physico-Chemical Property Calculation

The properties of all these compounds were simultaneously computed using Chem3D ultra. Subsequently, all these calculated properties were arranged in Microsoft Excel 2007 sheet and subjected to the statistical software VALSTAT. The different properties of the molecules computed were log P, connolly accessible area, connolly molecular area, connolly solvent accessible volume, molecular weight, ovality, principle moment of inertia X, Y, Z, molecular refractivity, partition coefficient, bending energy, charge-dipole energy, dipole-dipole energy, molecular topological index, shape attribute, shape coefficient, stretch energy, stretch-bend energy, bending energy, torsion energy, van der waal forces, sum of valence degrees.

2.4. QSAR Model Development

Dataset of compounds was separated into training and test set which was randomly carried out by VALSTAT software. The compounds which were selected by the software for training set were 3, 8, 11, 12, 13, 15, 16, 21, 22, 23, 26, 27, 31, 34, 38, 42, 44, 47, 48, 53, 54 and for test set were 4, 6, 20, 33, 46. The training set of compounds was used for development of suitable models whereas the test set of compounds was used for cross validation of the various models developed through training set.

The QSAR model was fabricated using Sequential Linear Multiple Regression method. An Inter-Correlation matrix between all parameters was developed and it is mentioned in the table 2. The observed, calculated, predicted and residual activity values for training set of compounds are mentioned in the table3. Fig. 1 shows a graph between experimental and calculated values of training set of compounds. Fig. 2 shows graph between predicted and experimental values of training set compounds. The predicted, experimental and predicted residual activity for test set of compounds is given in table 4. Fig. 3 shows graph between predicted and experimental values of test set compounds.

Model Validation: The developed models were validated using following methods-

- External Validation
- Internal Validation (Leave-one-out method)

The Cross-validated regression coefficient value was calculated by the following formula.

$$Q^2 = 1 - \frac{PRESS}{\sum_{i=1}^{N} (Z_i - Z_m)^2}$$

Where PRESS = predicted residual sum of squares,

Z_i = activity for training set,

Z_m = mean observed value, corresponding to the mean of the values for each cross-validation group.

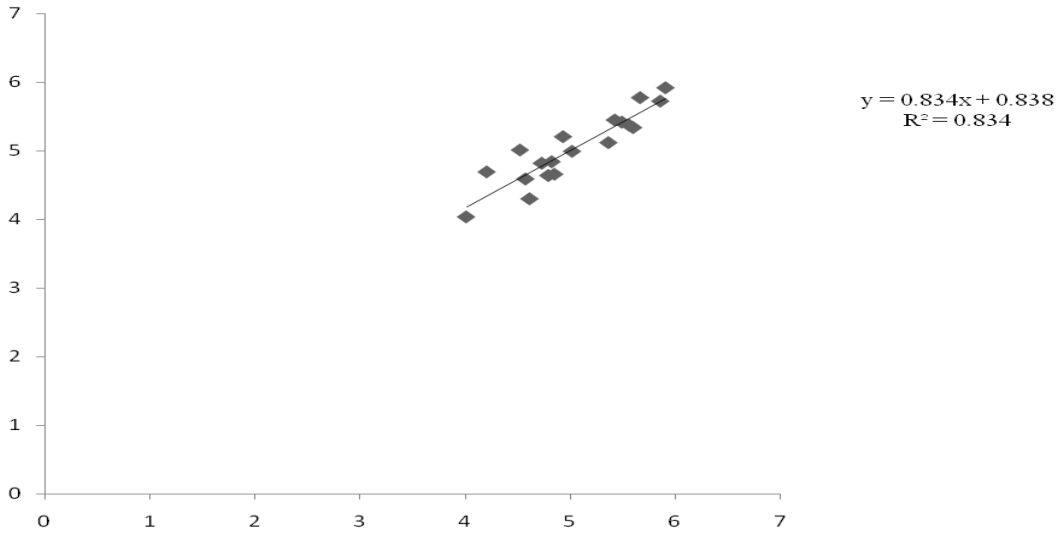


Fig. 1: Graph between experimental and calculated activity values for training set of compounds

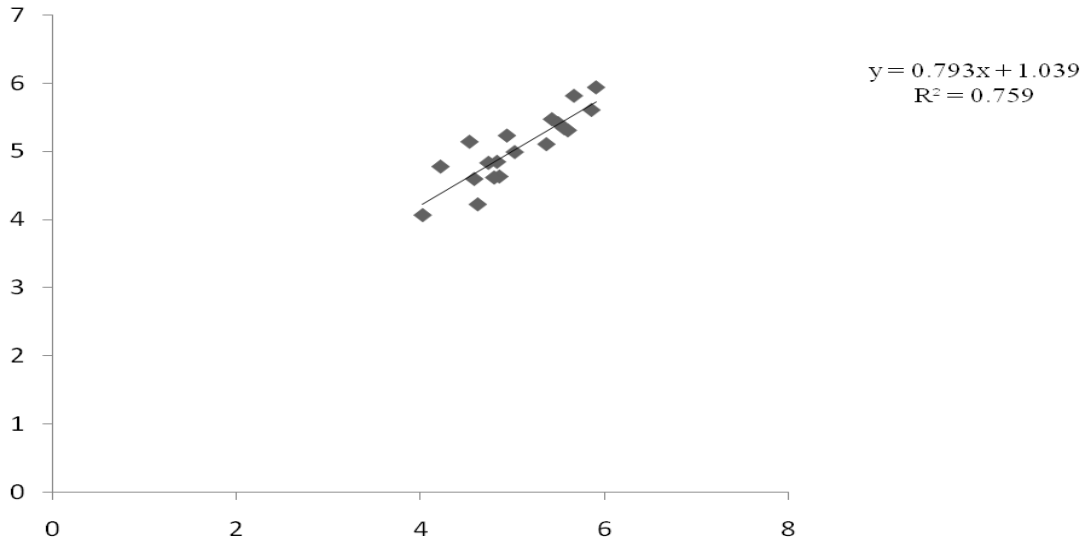


Fig. 2: Graph between experimental and predicted activity values for training set of compounds

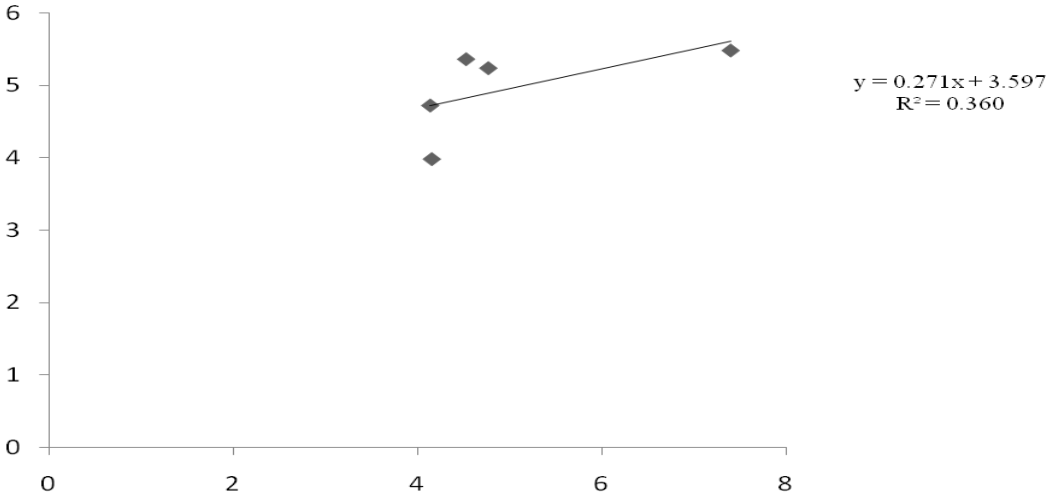
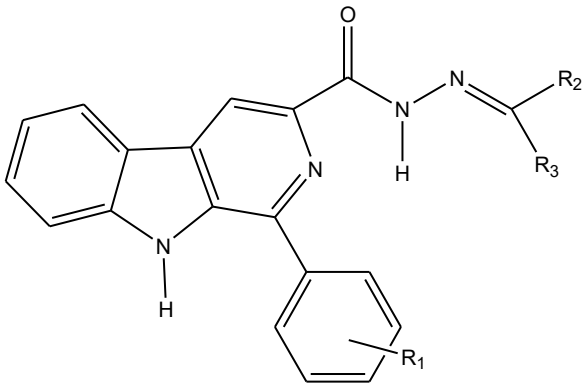
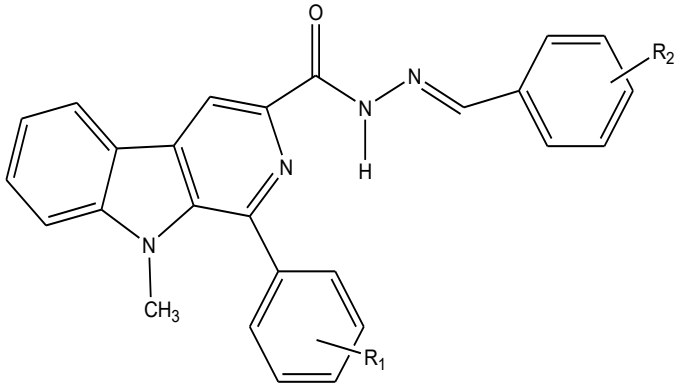


Fig. 3: Graph between experimental and predicted activity values for test set of compounds



Basic structure for compounds 3-48



Basic structure for compounds 53, 54

Table 1: 3-(Carbohydrazidesubstituted) - β-carboline derivatives with their experimental activities

C_{pd}	R_1	R_2	R_3	BA
3	3-NO_2	C_6H_5	H	5.425
4	3-NO_2	4-$N(CH_3)C_6H_5$	H	7.398
6	3-NO_2	2-ClC_6H_5	H	4.769
8	4-OCH_3	C_6H_5	H	4.027
11	4-OCH_3	2-ClC_6H_5	H	5.560
12	4-OH	4-$OCH_3C_6H_5$	H	4.208
13	4-OH	C_6H_5	H	5.365
15	4-OH	4-$NO_2C_6H_5$	H	5.019
16	4-OH	2-ClC_6H_5	H	5.492
20	H	4-$NO_2C_6H_5$	H	4.527
21	H	2-ClC_6H_5	H	5.665
22	4-NO_2	4-$OCH_3C_6H_5$	H	4.851
23	4-NO_2	C_6H_5	H	4.524
26	4-NO_2	2-ClC_6H_5	-	5.560
27	4-OCH_3	Cyclohexyl	-	4.731
31	4-OH	Cyclohexyl	-	4.825
33	4-NO_2	Cyclohexyl	-	4.138
34	4-OCH_3	Cyclohexyl	-	4.610
38	2-Cl	Cyclopentyl	-	4.013
42	2-Cl	CH_3	-	4.933
44	H	CH_3	CH_3	4.794
46	3-OH, 4-OCH_3	CH_3	CH_3	4.155
47	4-OH	CH_3	CH_3	4.615
48	4-$N(CH_3)_2$	CH_3	CH_3	4.576
53	3-NO_2	4-$N(CH_3)_2C_6H_5$	-	5.906
54	4-OCH_3	C_6H_5	-	5.857

Table 2: Inter-Correlation matrix amongst descriptors

	LogP	CAA	CMA	CSEV	MWt	OV	PMOIX	PMOIY	PMOIZ	MOREF	PC	BE	ChDi	DiDi	NVDW	SE	SBE	TE	VDW	MTI	SA	SC	SVD
LogP	1																						
CAA	0.605	1																					
CMA	0.65	0.9909	1																				
CSEV	0.723	0.9654	0.9871	1																			
MWt	0.683	0.9529	0.95587	0.9375	1																		
OV	0.375	0.9297	0.90373	0.8236	0.8815	1																	
PMOIX	0.484	0.9346	0.94384	0.9114	0.9098	0.904	1																
PMOIY	0.54	0.9703	0.94418	0.9005	0.9434	0.936	0.88	1															
PMOIZ	0.479	0.9668	0.95062	0.9025	0.9383	0.955	0.921	0.978	1														
MOREF	0.709	0.9678	0.97572	0.9603	0.9826	0.89	0.905	0.952	0.937	1													
PC	0.727	0.1294	0.18756	0.3031	0.118	0.143	0.01	0.032	0.043	0.1974	1												
BE	0.177	0.3607	0.3566	0.3651	0.2129	0.29	0.298	0.288	0.317	0.2429	0.4232	1											
ChDi	0.059	0.3814	0.42334	0.421	0.4291	0.369	0.411	0.401	0.44	0.446	0.2009	0.0867	1										
DiDi	0.39	0.1476	0.1453	0.141	0.2798	0.134	0.02	0.283	0.179	0.3049	0.1865	0.3656	0.0655	1									
NVDW	0.284	0.0332	0.00094	0.0154	0.2257	0.053	0.03	0.116	0.008	0.1431	0.0538	0.423	0.176	0.3743	1								
SE	0.068	0.4855	0.45601	0.3571	0.5275	0.658	0.577	0.546	0.608	0.4775	0.6947	0.3651	0.3726	0.0965	0.2034	1							
SBE	0.105	0.3945	0.3631	0.3545	0.3429	0.34	0.332	0.38	0.322	0.3371	0.0533	0.3012	0.3201	0.0892	0.0757	0.1178	1						
TE	0.233	0.1206	0.12176	0.0923	0.1251	0.182	0.023	0.237	0.146	0.232	0.231	0.1227	0.0051	0.7806	0.0867	0.0085	0.1012	1					
VDW	0.757	0.9094	0.94153	0.9627	0.8629	0.763	0.828	0.838	0.84	0.9111	0.4046	0.3811	0.3501	0.1969	0.0805	0.2768	0.2185	0.2059	1				
MTI	0.609	0.9721	0.97401	0.9507	0.9544	0.906	0.925	0.964	0.948	0.9824	0.133	0.275	0.5007	0.2548	0.0539	0.4958	0.3762	0.2238	0.89106	1			
SA	0.627	0.9735	0.97121	0.9399	0.984	0.929	0.92	0.971	0.963	0.9901	0.0966	0.2683	0.46	0.2462	0.1586	0.5377	0.363	0.1873	0.87431	0.9824	1		
SC	0.19	0.191	0.13986	0.1077	0.1788	0.208	0.149	0.173	0.161	0.136	0.041	0.1976	0.3605	0.0447	0.2564	0.1069	0.0321	0.1899	0.07444	0.1028	0.15	1	
SVD	0.484	0.9137	0.89433	0.8317	0.9421	0.946	0.87	0.952	0.945	0.9346	0.0823	0.2001	0.4629	0.265	0.2496	0.6415	0.3457	0.2182	0.75083	0.9337	0.97	0.161	1

Table 3: Experimental, Calculated, Predicted activity values for training set compounds

Compound	Observed	Calculated	Residual	Predicted	Pred_Residual
3	5.42481	5.45187	-0.02706	5.46998	-0.045168
11	5.5986	5.34112	0.257479	5.30671	0.291889
12	4.20775	4.69608	-0.48833	4.77606	-0.568312
13	5.36452	5.12146	0.243056	5.10425	0.260266
15	5.01863	4.99608	0.022554	4.98876	0.029874
16	5.49214	5.4208	0.071344	5.41219	0.079954
21	5.66555	5.77724	-0.11169	5.81323	-0.147684
22	4.85109	4.66209	0.188999	4.63038	0.220709
23	4.52462	5.01469	-0.49007	5.13963	-0.615011
26	5.56067	5.37762	0.183047	5.33636	0.224307
27	4.73166	4.82166	-0.09	4.83041	-0.098754
31	4.82507	4.84394	-0.01887	4.8452	-0.020132
38	4.01305	4.04024	-0.02719	4.06502	-0.051971
42	4.93293	5.20844	-0.27551	5.22993	-0.297001
44	4.79425	4.64326	0.150994	4.61525	0.179004
47	4.61475	4.30385	0.310901	4.22267	0.392081
48	4.57659	4.59239	-0.0158	4.59421	-0.01762
53	5.90658	5.92131	-0.01473	5.93595	-0.029372
54	5.85699	5.72611	0.130875	5.6048	0.252185

Table 4: Experimental, Predicted activity values of test set of compounds

C_{pd}	Observed	Predicted	Pred residual
20	4.52739	5.36056	-0.83317
6	4.768785	5.23678	-0.467995
4	7.39794	5.48049	1.91745
33	4.137869	4.71912	-0.581251
46	4.154902	3.97681	0.178092

3. RESULTS AND DISCUSSION

The development of QSAR model was carried out by performing sequential multiple linear regression analysis on the selected series (Barbosa et al., 2011) to find out better β-carboline derivative which results in the following models.

BA = [3.80861(±0.57417)] + LogP [0.431076 (± 0.101695)] +BE [-0.0376416(± 0.0165881)] +ChDi [0.455172(± 0.225939)] ------- (1)

BA = [4.17537(± 0.605439)] + LogP [0.305977 (± 0.1277)] +BE [-0.0477708(± 0.0171435)] +MTI [2.15155e-005 (± 1.15673e-005)] -------- (2)

BA = [2.29594(± 0.830433)] +MOREF [0.355035 (± 0.0783638)] +BE [-0.0683345 (± 0.0206988)] +SE [-0.00578236 (± 0.00299006)] ------ (3)

BA = [1.97921(± 0.894114)] +MWt [0.0113072 (± 0.0025127)] +BE [-0.0685135(± 0.0208124)] +SE [-0.00666666(± 0.00313194)] ------- (4)

BA = [3.7067(± 0.598276)] +LogP [0.355948 (± 0.117565)] +PMOIX [0.00022756 (± 0.000135149)] +BE [-0.0482381(± 0.0175505)] ---- (5)

Among the above given models model number 1 showed two compounds (compounds no. 8, compound no. 34) as outliers. Therefore these compounds were selectively taken out of the test set and the new optimized model (model no.6) having 19 training set compounds (3, 11, 12, 13, 15, 16, 21, 22, 23, 26, 27, 31, 38, 42, 44, 47, 48, 53, 54) was generated.

This optimized model no.6 was considered as the best model on the basis of significant statistical data obtained which has high cross validated correlation coefficient (Q^2) value and less standard error for prediction and it is given as follows:

BA = [3.72949(± 0.364258)] +LogP [0.497787(± 0.0659259)] +BE [-0.0431053(± 0.0108995)] +ChDi [0.369448(± 0.143659)] ---------- (6)

Fraction contribution of Log P = 0.537934

Fraction contribution of Bending energy = -0.28221
Fraction contribution of Charge-Dipole energy = 0.179856

The values for different statistical parameters obtained for the various models developed are given in the table 5. An inter-correlation chart between the parameters used in this model is mentioned in table 6.

Thus, it is understood that the best model (model 6) indicates that the biological activity is positively correlated with LogP and Charge-Dipole energy and negatively correlated with bending energy. Hence in order to increase the biological activity, the properties like LogP, and Charge-dipole energy should be increased, whereas bending energy which is showing a negative value in the equation should be decreased. Thus it is concluded that the biological activity will be increased if substituents that bring about changes in the molecule as stated above are affixed to it.

Table 5: Statistical data for developed models

Models	N	R	R^2	Q^2	Pred_R^2	SDEP	S_{PRESS}	F	Std. Error
1	21	0.780	0.609	0.461	0.329	0.412	0.459	8.816	0.391
2	21	0.773	0.597	0.440	0.259	0.420	0.467	8.404	0.396
3	21	0.770	0.592	0.419	0.357	0.428	0.476	8.249	0.399
4	21	0.768	0.590	0.397	0.3108	0.437	0.485	8.146	0.400
5	21	0.764	0.585	0.407	0.406	0.433	0.481	7.975	0.403
6*	19	0.913	0.834	0.757	0.337	0.265	0.298	25.115	0.247

Model no 6 is the best model*

Where N = No. of compounds in the training set; R = correlation coefficient, Std. error = Standard Error for Regression; Q^2 = cross validated R^2; Pred_ R^2 = Predicted R^2; S_{PRESS} - Standard error for prediction; SDEP = Standard deviation of prediction; F value = F-ratio between mean square regression and mean residualsquare.

Table 6: Inter-Correlation between parameters used in the best MODEL

Variables	Log P	BE	ChDi	VIF*
Log P	1.000000			1.05795
BE	0.212562	1.000000		1.01951
ChDi	0.077143	0.095813	1.000000	1.0614

VIF = Variance Inflation Factor (VIF)*

4. CONCLUSION

The best model designed exposed that the Log P values and Charge-Dipole energy of the molecules are positively associated to the biological activity whereas the bending energy values showed negative relationship. The best model developed also shows a greater control of Log P values and Charge-Dipole energy on biological activity than bending energy. Therefore, one should keep in mind that only those groups which impart the above mentioned changes must be attached to the molecules for escalating the biological activity. This study may prove to be helpful in further studies related with the synthesis of newer potent derivatives of β-carboline.

5. REFERENCES

1. Ma C, Cao R, Shi B, et al. *J. Med. Chem.*, 2010; **45**:5513-5519.
2. Mwense M, Wang XZ, Buontempo FV, et al. *SAR QSAR Environ. Res.*, 2006; **17**:53-73.
3. Cao R, Peng W, Chen H, et al. *J. Med. Chem.*, 2005; **40**:249-257.
4. Carroll AR, Davis RA, Quinn RJ, Eudistomin V. *J. Nat. Prod.*, 1998; **61**:959-960.
5. Castro AC, Dang LC, Soucy F, et al. *Bioorg. Med. Chem. Lett.*, 2003; **13**:2419-2422.
6. Baiget J, Lang S, MacKay SP, et al. *Beilstein J. Org. Chem.*, 2011; **7**:1407-1411.
7. Banks MN, Houston JG. Burger's Medicinal Chemistry, *Drug Discovery and Development*. 6th Ed. USA: Wiley Publications, 2010; 38-67.
8. Barbosa VA, Savariz FC, Formagio ASN, et al. *Bioorg. Med. Chem.*, 2011; **19**:6400-6408.
9. Barsanti PA, Wang W, Ni Z, et al. *Bioorg. Med. Chem. Lett.*, 2010; **20**:157-160.
10. Badre A, Boulanger A, Abou-Mansour E, et al. *J. Nat. Prod.*, 1994; **57**:528-533.
11. Benigni R, Giuliani A. *Bioinformatics* 2003; **19**:1194-1200.
12. Borman S. *Chem Eng News* 1990; **68**:20-23.

13. Cao R, Chen H, Peng W, et al. *J. Med. Chem.*, 2005; **40**:991-1001.
14. Herraiz T, González D, Ancín-Azpilicueta C, et al. *Food Chem. Toxicol.*, 2010; **48**:839-845.
15. Cao R, Chen Q, Hou X, et al. *Bioorg. Med. Chem.*, 2004; **12**:4613-4623.
16. Cao R, Guan X, Shi B, et al. *J. Med. Chem.*, 2010; **45**:2503-2515.
17. Kohli DV, Sharma MC, Sahu NK, et al. *Dig. J. Nano. Biostru.*, 2009; **4**:339-347.