Available online at www.pelagiaresearchlibrary.com

Pelagia Research Library

Pelagia Research Library

Der Pharmacia Sinica, 2013, 4(3):121-129

Der Pharmacia Sinica
ISSN: 0976-8688
CODEN (USA): PSHIBD

Structural indents for 5-hydroxyaurone derivatives as potent anticancer agents against HUVEC cancer cell lines-kNN MFA approach

Anupama Parate*, Rajesh Sharma and Isha Maheshwari

School of Pharmacy, Devi Ahilya University, Takshshila Campus, Khandwa Road, Indore, (M.P), India

---

ABSTRACT

Human umbilical vein endothelial cells are involved with endothelial cell function in human beings. Aurones exhibit good anticancer activities against human umbilical vein endothelial cell lines and hence provide for a demanding and potential area of development of aurone derivatives as anticancer agents. In the present research an attempt has been made to correlate the anticancer activity of the aurones to the physicochemical descriptors through QSAR studies. kNN- MFA and other advanced methods of analysis were opted for a set of seventeen aurone derivatives to figure out the correlationship between steric and electrostatic descriptors. In 2D-QSAR, amongst the several models obtained, statistically significant model 1 was selected with r^2=0.9095, q^2=0.4682 and pred_r^2=0.6132 with four descriptors Ssss-N count, SdSc-count, T_F_2_1 and StSc-count, which showed significant contributions in the model. The model was found to be significant within the corresponding limits of internal and external validation parameters. In 3D-QSAR, statistically significant model 2 was obtained using simulated annealing algorithm, with the corresponding q^2 and pred_r^2 values 0.7101 and 0.6304 respectively, showing internal and external predictivities 71% and 63% respectively. The four descriptors which contributed significantly to the model were E_299, S_851, S_564 and E_760. The models were found to show good statistical and predictive significance which can further be used for guiding ligand modification for the development of potential anticancer agents.

Keywords: Anticancer, QSAR, Aurones, HUVEC, Knn-MFA

INTRODUCTION

Cancer is a disease which is caused either directly or indirectly by physical, environmental, metabolic, chemical and genetic factors [1].Nearly 90% of all cancers are suggested to be caused by potentially controllable factors [2]. Cancer can be prevented, delayed, or reversed by diet control or use of pharmaceutical agents capable of controlling it[3, 4]. One such cancer is caused due to the uncontrolled growth of the human umbilical vein cells.

Human Umbilical Vein Endothelial Cells (HUVEC) are the cells obtained from normal human umbilical vein. These cell systems are helpful for investigations such as macromolecule transport [5], blood coagulation [6], and fibrinolysis [7]. In order to find an effective cure, many conventional anticancer agents have been used and they have been demonstrated in recent years to be beneficial. Antitumor activities can also be achieved by targeting activated endothelial cells [8]. Anticancer potential of a class of naturally occurring flavonoids, aurones, has been demonstrated here. Aurones, (Z)-2-benzylidenebenzofuran-3(2H)-ones, are the naturally occurring yellow pigments of the plants that are structurally isomeric of flavones [9], mostly impart color to flowers and fruits in

plants. Aurones possess a wide range of biological activities like anticancer [10], antihyperlipidaemic [11], antimalarial [12] etc. Among hydroxylated derivatives of aurones, 4-hydroxyaurones and 6-hydroxyaurones are more commonly studied but reports for 5-hydroxyaurones and 7-hydroxyaurones are currently limited. An extensive QSAR study has thus been carried out on the given set of 5-hydroxyaurone derivatives in an attempt to interpret the role of various substituent on the nucleus, so that they might prove helpful in the design of novel molecules as anticancer agents.

MATERIALS AND METHODS

All computations and molecular modeling studies were carried out on Windows 7 workstation using V-Life molecular Design Suite (V-Life MDS) version 3.5 [13]. The IC_{50} values of seventeen molecules were taken from the reported data and converted to pIC_{50} for QSAR analyses. QSAR models were built using V-Life MDS 3.5 (V-life Sciences Technologies Pvt. Ltd. Pune, India). All the structures were sketched using the 2D draw application of Chemoffice and converted to 3D structures. Three-dimensional structures were drawn for each molecule and the molecular geometries optimized using Monte Carlo conformational search [14], energy minimization and geometry optimization were conducted using the MM2 method with the root mean square (RMS) gradient set to 0.001 kcal/mol A. The following set of molecules (Table 1) were further used to perform QSAR. Energy minimized and geometry optimized structures [15] of the molecules are motioned in the 3-dimensional space so as to obtain appropriate conformations of the molecules [16] which were further aligned by template alignment method. The template structure, i.e., benzene ring was used for alignment by considering the common elements of the series as shown in Figure 1. In this alignment method, a template structure is defined and it is used as a basis for alignment of a set of molecules, and a reference molecule is chosen on which the other molecules of the dataset get aligned considering the chosen template. The descriptors were calculated after alignment which included the various structural, topological, electro topological and physicochemical descriptors. The descriptors which showed no significant variations were removed by removing the invariable columns. In 3D-QSAR, a total of 2048 three dimensional descriptors were calculated. These included various steric and electrostatic descriptors. Field descriptors were calculated, using Tripos force field [17] steric and electrostatic field types with cut-offs of 10.0 and 30.0 kcal/mol, respectively. Charge type selected was Gasteiger and Marsili [18]. Steric and electrostatic fields are computed at each grid point considering MMFF charges [19]. A value of methyl probe of charge +1 was assigned and the cutoff 10.0 kcal/ mol electrostatic and 30.0 kcal/mol steric were used. The distance-dependent dielectric constant was assigned a value of 1. The steric and electrostatic field interaction energies calculated are used to predict a relationship between steric and electrostatic properties to the biological activities of the compounds. The dataset was selected and the test and training set were chosen on the basis of manual data selection method keeping in mind the distribution of training: test as 70:30 i.e. 12 training: 5 test. In order to check the distribution pattern of the molecules, statistical parameters (with respect to the biological activity) i.e. mean, maximum, minimum and standard deviation are calculated for the training and test sets. As can be seen from Table 2 and Table 3 (for 2D and 3D respectively), the minimum of test set is greater than the minimum activity of training set and the maximum activity of the test set is less than the maximum activity of the training set, this indicates that the test set lies within the range of activity of the training set. Also Figure 2 shows the distribution of actual and predicted activities of test and training compounds (3D-QSAR). A k-nearest neighbor (k-NN) classification model was developed. An optimal k value is selected by optimization through the classification of a test set of samples or by leave-one-out cross-validation. Among several search algorithms, stepwise (SW) forward variable selection method[20], genetic algorithms [21] and simulated annealing [22] the variables and optimal k values were selected All the calculated descriptors were subjected to analyses by using Stepwise (SW), Genetic Algorithm (GA) and Simulated Annealing (SA) algorithms.

Stepwise Forward Variable Selection Method: In stepwise (SW) forward variable selection algorithm; independent variables are added one at a time, and the model is examined at each step using partial least square regression. Here, F is kept equal to 4 for inclusion and F=3.99 for exclusion for the forward–backward selection method. The cross-correlation limit was set at 0.5, the number of variables at 5, and the term selection criteria at q^2. The variance cutoff was set at 0.0, and auto scaling was set for scaling option.

Simulated annealing: During this process of simulated annealing (SA), the system adopts all the low energy states which are most populated [23]. Here, all the calculated descriptors remaining after removal of invariable columns were subjected to SA algorithm coupled with k-nearest neighbour (k-NN) methods for building a QSAR model based on the training set.

Model evaluation is done next to validate the model both internally and externally so as to check the effectiveness and the predictive ability of the model. Internal validation is carried out using leave-one-out (q^2, LOO) method. In this method, biological activity of each molecule is predicted once, eliminating it out of the system [24]. For external validation, activity of each molecule in the test set was predicted using the model generated from the training set. Pred_r^2 determines the ability of the model to effectively predict the biological activity values for a external test set. Developed quantitative models were evaluated using following statistical measures: n, number of observations (molecules); k, number of variables (descriptors); Number of components; Number of nearest neighbors, number of k-nearest neighbor in the model; r^2, coefficient of determination; q^2, cross-validated r^2 (by leave one out); pred_r^2, r^2 for external test set; F-test, pred_r^2_se, standard error of external test set prediction. The r^2 and q^2 values are used as determining factors in checking the effectiveness of the model [25].

RESULTS

A. 2D-QSAR- The following model was developed for 2D-QSAR by SA.
The statistical result of 2D-QSAR model along with the contribution of the descriptors is tabulated in Table 4. A brief idea of the requirement of different physicochemical parameters and their contributions (positive or negative influence on biological activity), required for potential anti-HUVEC anticancer activity, was obtained from the 2D QSAR analysis.

pIC_{50} = 0.8829 SsssN-Count - 0.6607 SdssC-count - 0.7194 T_F_2_1- 0.4824 Sts-Count + 6.4439

B. 3D QSAR
The models were obtained using all the three methods, SW, SA and GA. Model obtained using Stepwise Linear Regression showed good q^2 and pred_r^2 values but, since only a single descriptor contributed to the model, the model was unable to describe the exact structural requirements of the molecule.

Next, model was developed using SA with kNN which led to statistically significant results. Table 5 reports the statistically significant parameters of both the models. 3D data points were generated around the molecule (Figure 3) which helps to define the requirements which would promote the anticancer activity. Four parameters both electrostatic and steric, E_299, S_851, S_564 and E_760, were found to show significant contributions to the model. This helps to generate the steric and electrostatic fields which correlate these parameters to the activity profile of the pharmacophore.

DISCUSSION

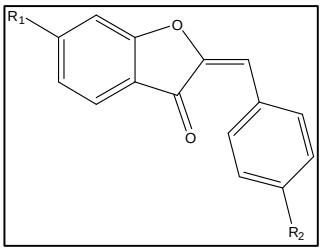
Four 2D descriptors **SsssN Count** (which signifies the number of N-atoms connected with three single bonds), **SdssC- Count** (which defines the total number of carbon atoms connected with one double bond and two single bonds present in the molecule), **T_F_2_1** (which defines the total number of fluorine atoms separated from any doubly bonded atom separated by a distance of one bond) and **StSC-Count** (which defines the total number of carbon atoms with a triple bond and a single bond present in the molecule).

SsssN-Count showing a positive contribution can be correlated to the fact that the presence of dimethylamino and diethylamino group in the molecule enhances the activity, greater the length of the side chain attached to the nitrogen atom, greater is the activity. The second parameter SdssC count contributed negatively which explains why amide group containing compound showed poor activity profile. On the other hand, the remaining two descriptors, T_F_2_1 and StSC count also contributes negatively which proves why flouro and cyano containing molecules show poor activities.

In 3D-QSAR the steric (S) and electrostatic (E) descriptors identify the regions which are associated with altering the biological activity. S_564, E_299, S_851 and E_760 are the steric and electrostatic field energy of interactions between probe (CH_3) and compounds at their corresponding spatial grid points of 299, 564, 851 and 760. Since all the descriptors lie in the area of the substituted benzyl ring, it can be concluded that derivatives of these benzyl ring with appropriate steric and electrostatic substituent can prove effective. The specific descriptor values obtained show the negative and the positive contribution of these descriptors to the biological activity. Thus, more steric substituent at S_564 and S_851, while in case of electrostatic descriptors, more electrostatic substituent are preferred at E_299 and less electrostatic substituent at E_760 is preferred. The predictive ability

of this SA kNN MFA model was evaluated by predicting the biological activities of the test set molecules. The plots of predicted versus activity have been reported. The internal and external predictability of the above 3D-QSAR model using the test set was determined by q^2 and pred_r^2, which is 0.7101 and 0.6304 respectively. So the above results indicate that 3D-QSAR model generates 71.01% and 63.04% internal and external model prediction, respectively. The plot of contributions of steric and electrostatic field interactions indicates relative regions of the local fields (steric and electrostatic) around the aligned molecules. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table VI. Also Figure 4 and Figure 5 show the plots of predicted versus actual activities of the given compounds for 3D and 2D-QSAR respectively.

Table 1. HUVEC anticancer activities of substituted aurones



Comp.	R_1	R_2	IC_{50} (µm)	pIC_{50}
1	H	OH	9.35	5.017
2	H	OMe	9.86	5.006
3	H	F	39.45	4.403
4	H	Cl	7.93	5.1
5	H	Br	8.97	5.04
6	H	CF_3	5.35	5.27
7	H	CN	22.74	4.64
8	H	$CONH_2$	35.58	4.448
9	H	NH_2	7.33	5.134
10	H	NMe_2	1.35	5.86
11	H	NEt_2	0.25	6.602
12	H	Me	9.35	5.03
13	H	i-Pr	6.38	5.19
14	H	t-Bu	3.49	5.457
15	Ac	NMe_2	3.18	5.49
16	Ac	NEt_2	0.23	6.638
17	Ac	t-Bu	46.4	4.33

Table 2. Unicolumn Statistics

Parameters	Test set	Training set
Minimum	4.480	4.3300
Maximum	5.900	6.6380
Mean	5.3405	5.1465
Standard Dev.	0.7247	0.6526

Table 3. Unicolumn Statistics

Parameters	Test set	Training set
Minimum	4.4030	6.6380
Maximum	5.900	4.3000
Mean	5.2786	5.2227
Standard Dev.	0.6236	0.7292

Table 4. Statistical results of 2D QSAR equation generated by MLR method

S. No.	Statistical Parameters		2D-results
1	r^2		0.9095
2	r^2_se		0.1802
3	q^2		0.4682
4	q^2_se		0.4368
5	pred_r^2		0.6132
6	pred_r^2se		0.7142
7	F-test		17.5969
8	n		12
9	Degree of Freedom		7
10	Contributing descriptors	1. 2. 3. 4.	Ssss-N count(0.8829) SdssC count(-0.6607) T_F_2_1(-0.7194) StSC count(-0.4824)

Table 5. Statistical parameters for 3D QSAR

Parameters	QSAR(kNN_SA)	QSAR (kNN_SW)
kNN	3	2
Degree of freedom	7	10
No of training molecules	12	12
q^2	0.7101	0.6031
q^2_se	0.3926	0.4594
pred_r^2	0.6304	0.5951
pred_r^2se	0.3810	0.3462
Descriptors	E_299(0.1839, 0.3314) S_851(0.0064, 3.3695) S_564(0.6078, 30.0000) E_760(-2.3273, 0.1019)	S_571(0.1514-.5471)

Table 6. Table of Actual and Predicted activities as per 2D and 3D models

Comp.	Actual values[a]	Predicted Values[b]	Predicted Values[c]
1	5.017	5.0863	5.12244
2	5.006	4.8673	5.12244
3	4.403	4.69875	4.403002
4	5.1	4.86359	5.12244
5	5.04	4.84624	5.12244
6	5.27	5.45699	5.12244
7	4.64	4.81858	5.12244
8	4.448	4.92252	4.64002
9	5.134	5.04338	4.46169
10	5.86	5.25555	5.12244
11	6.602	6.638	6.00537
12	5.03	4.93851	6.00537
13	5.19	5.2682	5.12244
14	5.457	5.26764	5.12244
15	5.49	6.1971	5.34462
16	6.638	6.602	5.34462
17	4.33	5.45619	4.46169

a. Reported pIC_{50} values of the given compounds
b. pIC_{50} values predicted by 3D-QSAR model statistically validated by Simulated annealing method.
c. pIC_{50} values predicted by 2D-QSAR model statistically validated by Multiple Regression method.

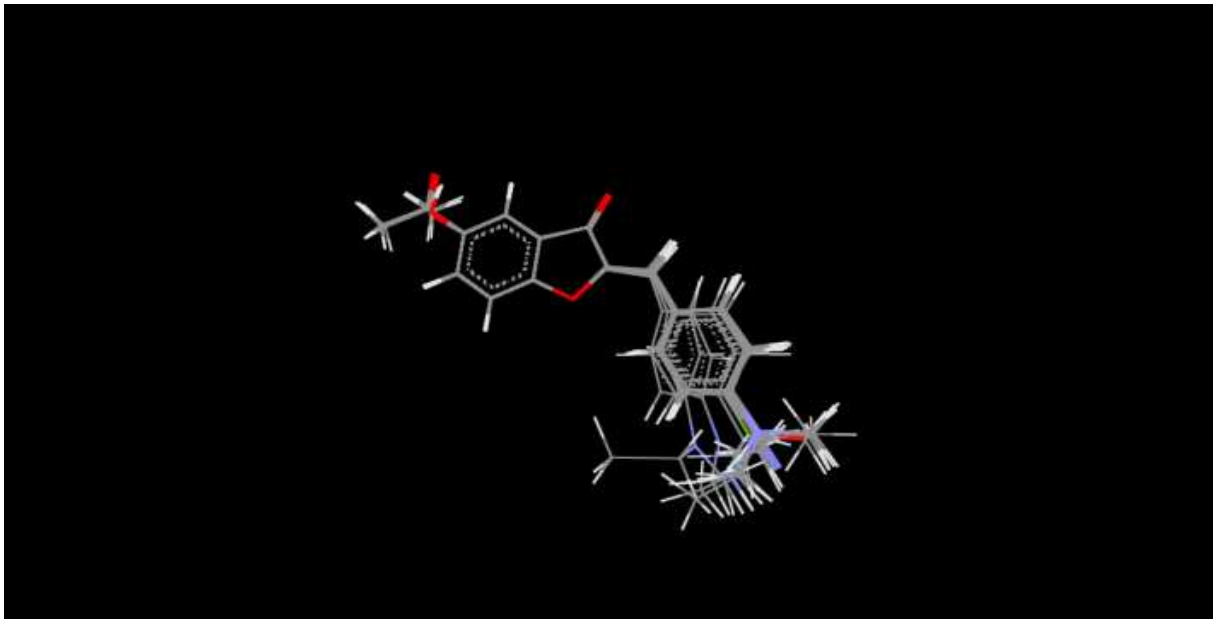
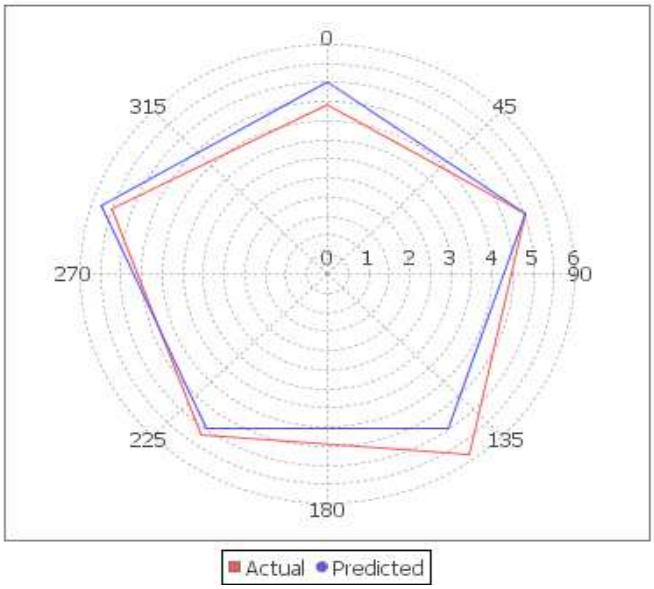


Figure 1: Template based Alignment



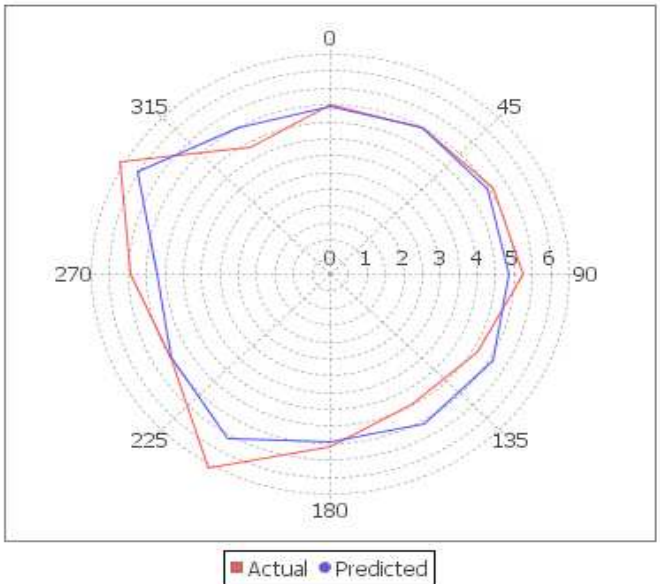


Figure 2: Actual and Predicted biological activity ranges for test and training respectively

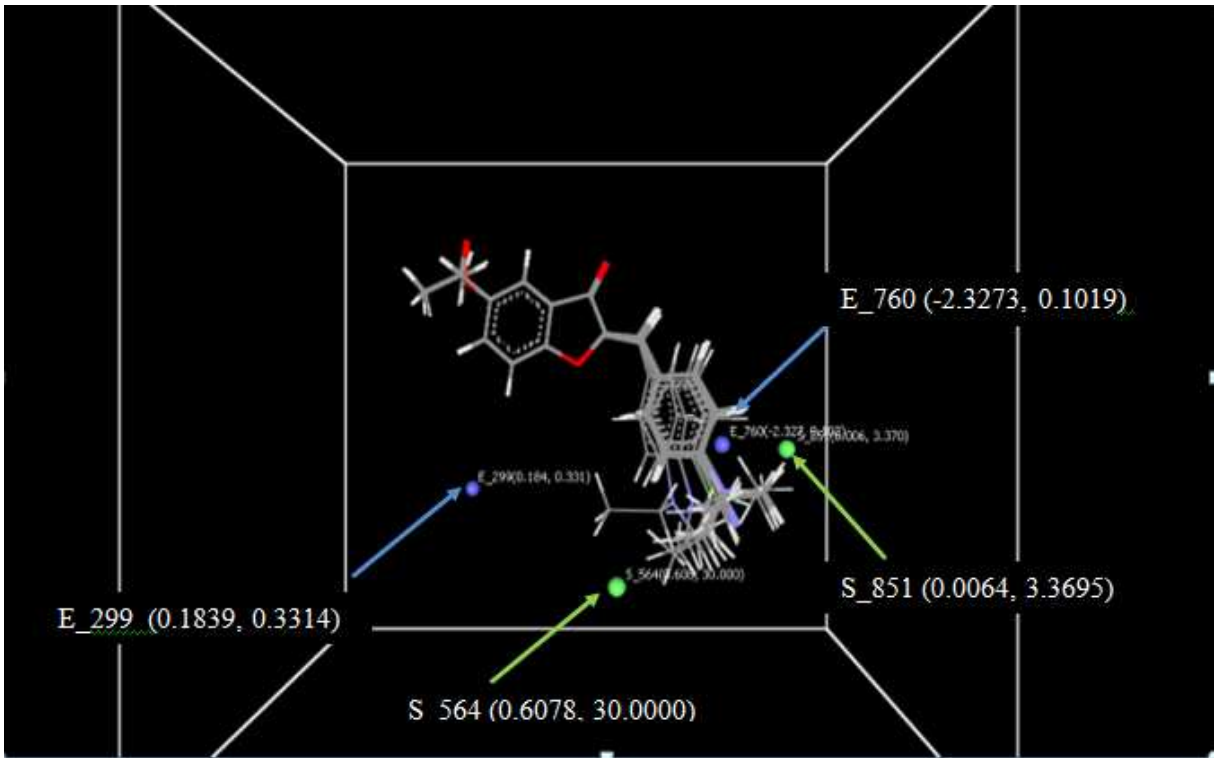


Figure 3: Contribution plot for steric and electrostatic field interactions in 3D-QSAR

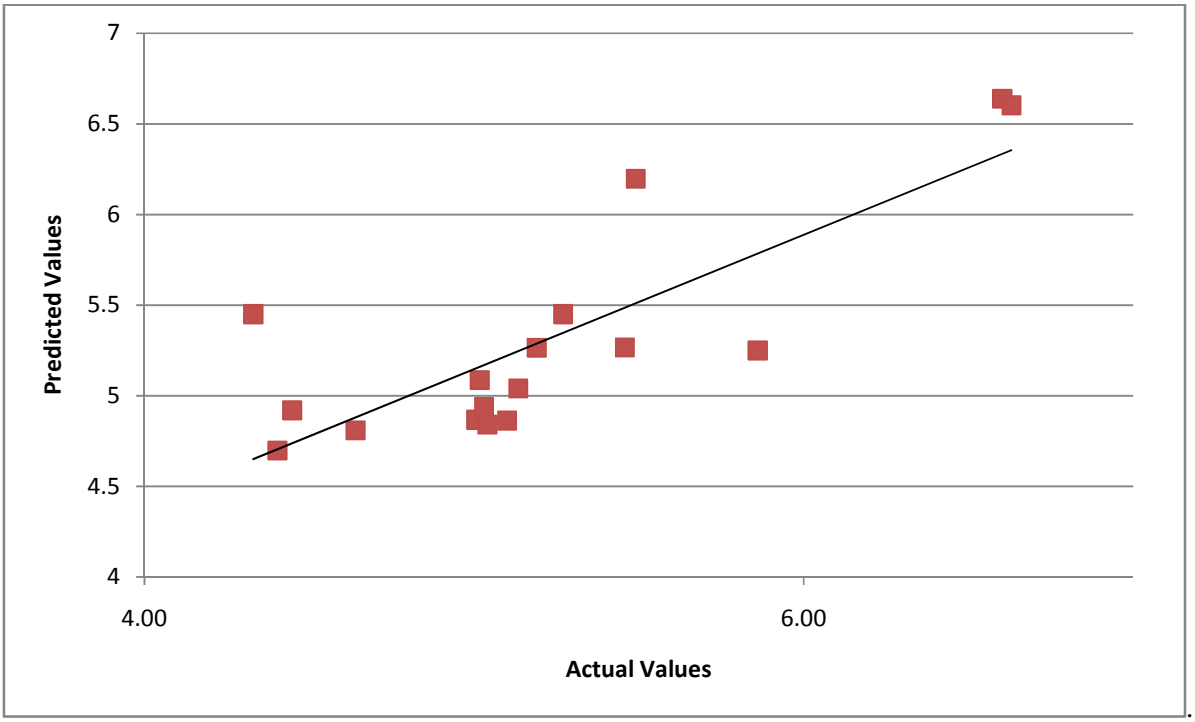


Figure 4: Predicted versus Actual activity plot for 3D-QSAR

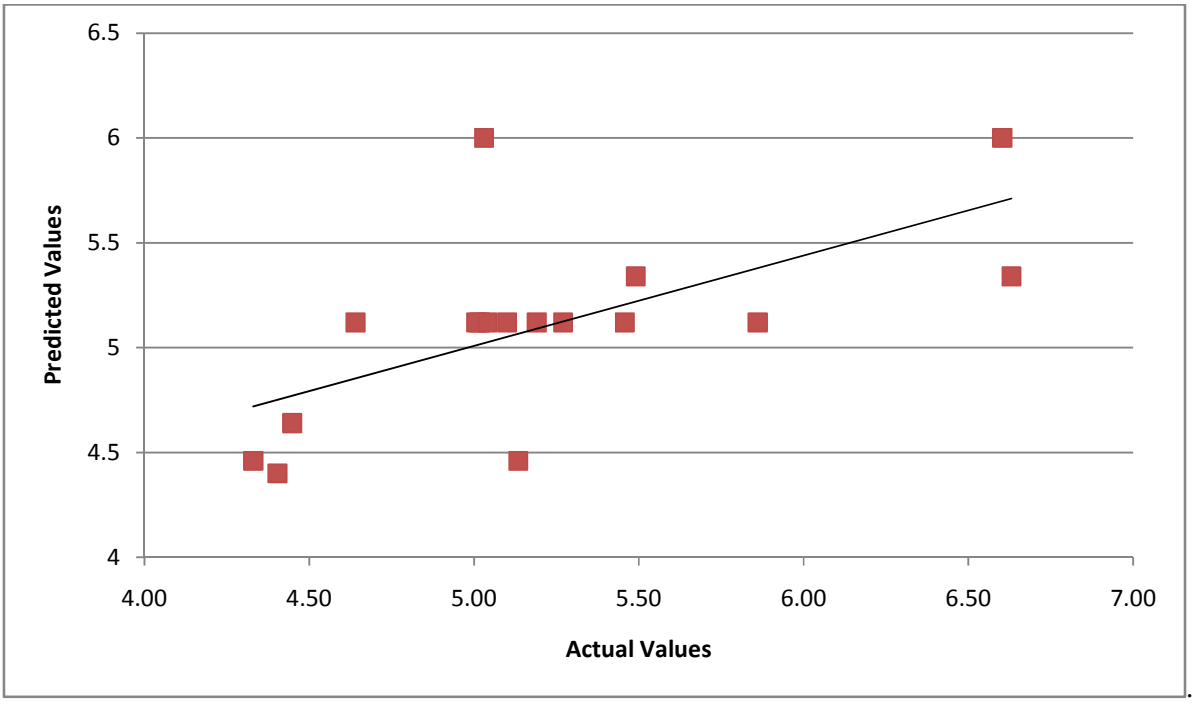


Figure 5: Predicted versus Actual activity plot for 2D-QSAR

CONCLUSION

In 2D-QSAR, study unveils key structural requirements for anticancer activity. The analysis signifies the importance of various physicochemical and topological descriptors and their related contribution towards activity. This generated model can prove effective in accounting for the changes to be accommodated so as to obtain good biological activity results. On the other hand, the importance and utility of the new 3D-QSAR method discussed here has been established using the V-Life software. 3D-QSAR model which was generated by kNN-MFA using the simulated annealing (SA) method has been reported. This model has shown good internal as well as external predictivity values. This study shows how chemical features for a set of compounds along with their activities ranging over several orders of magnitudes can be used to generate QSAR equation that can successfully predict the activity.

Acknowledgements
The authors would like to thank V-Life, Pune for providing the software and Head of Department, School Of Pharmacy, DAVV, Indore for providing with all the necessary facilities to work. Also, would like to thank AICTE for providing with the fellowship which helped us to carry out our work efficiently.

REFERENCES

[1] S. Rahman, F. Salehin, A. Iqbal, *BMC Compl. and Alt. Med*. 11, **2011**, 76-82.
[2] R. Doll, R. Peto, *J. Natl. Cancer Inst*. 66, **1981**, 1191-1308.
[3] L. W. Wattenberg, *Cancer Res*. 45, **1985**, 1-8.
[4] M. B. Sporn, N. M. Dunlop, D. L. Newton, J. M. Smith, *Fed. Proc*. 35 , **1976**, 1332-38.
[5] L. W. Tania, E. S. Sarah, M. Reusch, B. Wintroub, *In Vitro Cell. Biol*. 26, **1990**, 759-768.
[6] B. Furie, B. C. Furie, *Cell* 53 , **1988**, 505-518.
[7] J. Wojta, M. Gallicchio, H. Zoellner, E. L. Filonzi, J. A. Hamilton, K. Mcgrath, Blood 12 , **1993**, 3285-3292.
[8] L. Cheng, Y. Zhang, S. Liu, H. Chen, X. Cheng, Z. Lu, G. C. Zheng, *Eur. Jour. of Med. Chem.* 45 , **2010**, 5950-5957.
[9] B. P. Bandgar , S. A. Patil, B. L. Korbad, S. C. Biradar, S. J. N. Nile, C. N. Khobragade, *Eur. Jour. of Med.* Chem. 45 , **2010**, 3223-3227.
[10] R. Valclavikova, E. Kondrova, M. Ehrlichova, A. Boumendjel, J. Kovar, P. Stopka, P. Soucek, I. *Bioorg. Med. Chem.* 16 , **2008**, 2034-2042.
[11] M. A. F. Jahromi, A. B. Ray, *Jour. of Nat. Prod*. 56 , **1993**, 989-994.
[12] F. Souard, S. Okombi, C. Beney, S. Chevalley, A. Valentin, A. Boumendjel, *Bioorg. & Med. Chem.* 18 , **2010**, 5724–5731.
[13] V-Life MDS, Version 3.5, **2008**, V-Life Sciences Technologies Pvt. Ltd., Pune, India.
[14] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* 21 , **1953**, 1087–1092.
[15] S. Ajmani, K. Jhadav, S. A. Kulkarni, *Chem. Inform. Mod*. 46 , **2006**, 24–31.
[16] K. Baumann, *J. Chem. Inf. Comput. Sci*. 42 , **2002**, 26–35.
[17] M. Clark, R. D. Cramer, O. N. Van, *J. Comput. Chem.10* , **1989**, 982–1012.
[18] J. Gasteiger, M. Marsili, *Tetra*. 36 , **1980**, 3219–3228.
[19] T.A. Halgren, *J. Comput. Chem.* 17 , **1996**, 553–586.
[20] R.B. Darlington, Regression and linear models McGraw-Hill, New York, **1990**.
[21] K. Hasegawa, T. Kimura, K. Funatsu, *Quant. Struct. Act. Rel*. 18 , **1999**, 262–272.
[22] W. Zheng, A. Tropsha, *J. Chem. Inform. Comput. Sci*. 40 , **2000**, 185–194.
[23] S. Kirkpatrick, C.E. Gelatt Jr., M.P. Vecchi, Optimization by simulated annealing, Sci. 220 , **1983**, 671-680.
[24] R. D. Cramer, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc*. 110 , **1988**, 5959–5967.
[25] A. Parate, S. C. Chaturvedi, *Med. Chem. Res.* 19 , **2010**, 375–391.