PHARMACIA
Vol. -2 , Issue - 1, Year - June 2013
ISSN 0976-9692

2D AND 3D QSAR STUDIES ON 4-THIAZOLIDINONES CONTAINING INDOLIN-2-ONE MOIETY AS POTENTIAL ANTITUMOR AGENT

Shaikh Anwar[1], Thombare Shraddha[1]

[1]MET's Institute of Pharmacy, Bhujbal Knowledge City, Adgaon, Nashik-422003, India

Abstract

Quantitative structure–activity relationship (QSAR) analysis for recently synthesized 4-thiazolidinones and indolin-2-one hybrid derivatives was studied for their cytotoxic activities. The statistically significant 2D-QSAR models (r^2 = 0.8363; q^2 = 0.5888; F test = 19.4139; r^2 se = 0.2637; q^2se = 0.4180; pred_r^2 = 0.4118; pred_r^2se = 0.5771) were developed using multiple linear regression (MLR) methodology with stepwise (SW) forward-backward variable selection method, when the number of descriptors in equation was set to five, indicating the descriptors of T_O_S_5, T_2_O_6, SdssCE-index, chi6chain and T_N_F_4 mainly influence the cytotoxic activity. The results of the 2D-QSAR model were further compared with 3D-QSAR model generated by kNN-MFA, (k-Nearest Neighbor Molecular Field Analysis), investigating the substitutional requirements for the favorable receptor–drug interaction and providing useful information in the characterization and differentiation of their binding sites. The results derived may be useful in further designing novel antitumor agent prior to synthesis.

Keywords: 4-thiazolidinone, indolin-2-one, cytotoxic activities, QSAR, KNN-MFA.

Introduction

4-Thiazolidinone derivatives are an important group of heterocyclic compounds possessing a variety of biological effects [1], including antitumor [2-4], anti-inflammatory [5], antimicrobial [6], antiviral [7], anticonvulsant [8], antifungal [9], and antibacterial [10] activities and so on. Among them, 5-benzylidene-4-thiazolidinone derivatives have been reported to show marked antitumor activities with different biotargets and mechanism, such as phosphatase of a regenerating liver (PRL-3) [11], Sphingosine Kinase (SK) [12], JNK-stimulating phosphatase-1 (JSP-1) [13] and nonmembrane protein tyrosine phosphatase (SHP-2) [14]. Moreover, 5-benzylidene-4 thiazolidinone derivatives exhibited potent antitumor activities against non-small cell lung cancer cell line H460, paclitaxel-resistant H460taxR, human colon cancer cell line HT-29 and human breast cancer cell line MDA-MB-231 [15]. The indolin-2-one ring system belongs to the privileged structure in modern medicinal chemistry, particularly in discovery of new antitumor and anti-angiogenic agents. Various kinase inhibitors containing indolin-2-one moiety have been intensively studied for the inhibition of VEGFR, c-Kit, FLT3, PDGFR-a/b, and CSF-1-R [16]. Sunitinib, a multitargeted receptor tyrosine kinase inhibitor, interfering with tumor blood vessel formation, is approved by the FDA for the treatment of advanced renal cell carcinoma (RCC) and gastrointestinal stromal tumors (GIST) [17]. BIBF1120, a triple angiokinase inhibitor reported by Boehringer, is currently in phase III clinical trials in non-small cell lung cancer [18]. Indirubin, an active ingredient of a traditional Chinese medicine recipe, has been applied to treat chronic myelocytic leukemia [19]. Based on these facts, substituted indolin-2-one attached to the 5-benzylidene-4-thiazolidinone scafford, and the combination of two privileged structures in one molecule leads to drug like molecules. Indolin-2-one group was introduced at the 2-position of the 4-thiazolidonone ring and a basic side chain was introduced at the 3-position of the 4-thiazolidinone ring, in order to improve the solubility and bioavailability of these compounds. Recently synthesized 4-thiazolidinone and indolin-2-one hybrid derivatives were discovered with the cytotoxic activities [20].

Corresponding author-
Dr. Shaikh Anwar
Bhujbal Knowledge City, MET's Institute of Pharmacy, Adgaon, Nashik-422003, India
E-mail: pharmacy_2003@rediffmailmail.com

Traditional computer-assisted quantitative structure–activity relationship (QSAR) studies pioneered by C.z Hansch et al. [21] have been proved to be one of the useful approaches for accelerating the drug design process [22], which helped to correlate the bioactivity of compounds with structural descriptors [23]. QSAR which has become an accepted tool for establishing quantitative relationship between biological activity and descriptors representing physicochemical properties of the compounds in a series using statistical methods and it helps to predict the biological activities of newly designed analogues contributing the drug discovery processes [24]. Recently synthesized 4-thiazolidinone and indolin-2-one hybrid derivatives were discovered with the cytotoxic activities (Table 1) [20]. To gain further insights into the structure-activity relationships of these derivatives and understand the mechanism of their substitutional specificity, statistical significant 2D and 3D QSAR models were developed between antitumor activities and structural descriptors of compounds using molecular design suite software (VLifeMDS 4.1) [25] that can help the medicinal chemist to design new candidates as potential antitumor agents prior to synthesis.

Materials And Methods

Selection of molecules

Data set of 38 [4-thiazolidinone and indolin-2-one hybrid derivatives] (Table 1) collected from published literature [20] were taken for present study. The affinity data of antitumor activities were converted into the corresponding pKi values to get the linear relationship in equation using the following formula: pKi = -log Ki, where Ki value represents antitumor activity in IC50 (μM). Molecules were rationally divided into the training set and test set using sphere exclusion algorithm, random selection and manual selection method, captured structural features of the training set molecules, were used for activity prediction.

Molecular modeling

All computational experiments were performed using on HCL computer having genuine Intel Pentium Dual Core Processor and Windows XP operating system using the software Molecular Design Suite (vlifeMDS 4.1). Structures were drawn using the 2D draw application and converted to 3D structures and subjected to an energy minimization and geometry optimization using Merck Molecular Force Field (MMFF), force field and charges followed by Austin Model-1 with 10000 as maximum number of cycles, 0.01 as convergence criteria (root mean square gradient) and 1.0 as constant (medium's dielectric constant which is 1 for in vacuo) in dielectric properties. The default values of 20.0 and 10.0 Kcal/mol were used for electrostatic and steric energy cutoff.

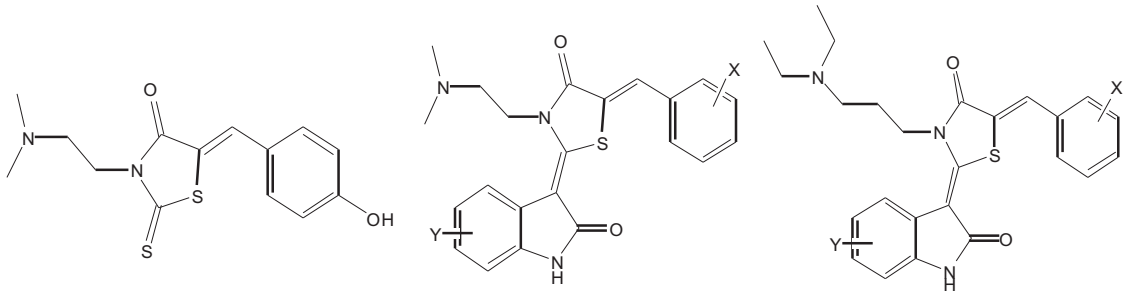

Table 1: Chemical structures and antitumor activities of 4-thiazolidinone and quinolin-2-one hybrid derivatives

Compounds No.	X	Y	Ki	pKi
1(a)*	4- hydroxyl	-	46.13	1.7517
2(b)*	4- hydroxyl	H	0.98	0.1746
3(b)	4- hydroxyl	5-CH_3	1.77	0.4481
4(b)	4- hydroxyl	5-F	1.27	0.6673
5(b)	4- hydroxyl	6-F	1.43	0.3125
6(b)	4- hydroxyl	5-Cl	2.03	0.1811
7(b)	4- hydroxyl	5-Br	1.11	0.1741
8(b)	4-methoxy	6-F	2.20	0.1838
9(b)	3,4-difluoro	5-F	44.7	1.4984
10(b)	3,4-difluoro	6-F	6.4	1.1418
11(b)	3,4-difluoro	5-CH_3	59.67	1.2761
12(b)	2,4-difluoro	6-F	3.20	0.3464
13(b)	2,4-dimethoxy	5-Cl	10.6	1.4408
14(b)	2,4-dimethoxy	6-F	20.53	1.2128
15(b)	2,5-dimethoxy	6-F	7.15	1.2173
16(b)	2,5-dimethoxy	H	63.7	1.2747
17(b)	3,4 dioxymethylene	H	0.82	0.0668
18(b)	3,4-dioxymethylene	5-CH_3	4.03	0.2496
19(b)	3,4-dioxymethylene	6-F	2.81	0.0093
20(b)	3,4,5-trimethoxy	5-F	1.70	0.3447
21(c)*	3,4-dioxymethylene	H	1.02	0.05221
22(c)	3,4-dioxymethylene	5-CH_3	1.86	0.2349
23(c)	3,4-dioxymethylene	5-F	0.075	0.3472
24(c)	3,4-dioxymethylene	6-F	1.91	-0.0052
25(c)	3,4-dioxymethylene	5-Cl	1.00	0.2227
26(c)	3,4-difluoro	5-F	1.30	0.6587
27(c)	3,4-difluoro	6-F	1.12	0.3062
28(c)	2,4-difluoro	H	5.47	0.3892
29(c)	2,4-difluoro	6-F	2.33	0.3318
30(c)	2,4-difluoro	5-Cl	6.10	0.5598
31(c)	3,4,5-trimethoxy	H	1.67	0.0347
32(c)	3,4,5-trimethoxy	5-CH_3	2.67	0.2175
33(c)	3,4,5-trimethoxy	5-F	0.78	0.3297
34(c)	3,4,5-trimethoxy	6-F	1.41	-0.0227
35(c)	3,4,5-trimethoxy	5-Cl	2.99	0.2053
36(c)	4-hydroxy	5-CH_3	0.26	0.1311
37(c)	2-hydroxy	5-CH_3	11.33	1.4103
38(c)	2-hydroxy	5-Br	51.67	1.3962

Ki = antitumor activity of compounds, pKi = all antitumor activities are expressed as –log (Ki), which is pKi. (a)*- compounds containing nucleus 'a', (b)*- compounds containing nucleus 'b', (c)*- compounds containing nucleus 'c'.

2D-QSAR analysis

Calculation of descriptors

Number of descriptors was calculated after optimization or minimization of the energy of the data set molecules. Various types of physicochemical descriptors were calculated: Individual (Molecular weight, H-Acceptor count, H-Donor count, XlogP, slogP, SMR, polarisablity, etc.), retention index (Chi), atomic valence connectivity index (ChiV), Path count, Chi chain, ChiV chain, Chain PathCount, Cluster, Pathcluster, Kappa, Element count (H, N, C, S count etc.), Distance based topological (DistTopo, ConnectivityIndex, WienerIndex, Balaban Index), Estate numbers (SsCH3count, SdCH2count, SssCH2count, StCHcount, etc.), Estate contribution (SsCH3-index., SdCH2-index, SssCH2-index , StCH index), Information theory based (Ipc, Id etc.) and Polar surface area. More than 200 alignment independent descriptors were also calculated using the following attributes. A few examples are T_2_O_7, T_N_N_5, T_2_2_6, T_C_O_1, T_O_Cl_5 etc. The invariable descriptors (the descriptors that are constant for all the molecules) were removed, as they do not contribute to QSAR.

Generation of training and test sets

In order to evaluate the QSAR model, data set was divided into training and test set using sphere exclusion, random selection and manual selection method. Training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the predictive power of the model which is not included in model generation.

Sphere Exclusion method: In this method initially data set were divided into training and test set using sphere exclusion method. In this method dissimilarity value provides an idea to handle training and test set size. It needs to be adjusted by trial and error until a desired division of training and test set is achieved. Increase in dissimilarity value results in increase in number of molecules in the test set.

Random Selection Method: In order to construct and validate the QSAR models, both internally and externally, the data sets were divided into training [90%-60% (90%, 85%, 80%, 75%, 70%, 65% and 60%) of total data set] and test sets [10%-40% (10%, 15%, 20%, 30%, 35% and 40%) of total data set] in a random manner. 10 trials were run in each case.

Manual data selection method: Data set is divided manually into training and test sets on the basis of the result obtained in sphere exclusion method and random selection method. Generation of 2D-QSAR models

MLR was used for model generation. All the calculated descriptors were considered as independent variable and biological activity as dependent variable. Multiple regression is the standard method for multivariate data analysis. It is also called as ordinary least squares regression (OLS). This method of regression estimates the values of the regression coefficients by applying least squares curve fitting method. For getting reliable results, data set having typically 5 times as many data points (molecules) as independent variables (descriptors) is required. The regression equation takes the form

$$Y = b1*x1 + b2*x2 + b3*x3 + c$$

Where, Y is the dependent variable, the 'b's are regression coefficients for corresponding 'x's (independent variable), 'c' is a regression constant or intercept. Multiple linear regression is among the most widely used mapping methods in QSAR in last decades.

3D-QSAR analysis

kNN-MFA

A novel three- dimensional QSAR approach has been developed based on the principles of the k-nearest neighbor method. The method employs different variable selection procedures with either simulated annealing, stepwise forward or genetic algorithm. Molecular Design Suite (VLifeMDS 4.1) allows user to choose probe, grid size, and grid interval for the generation of descriptors. The variable selection methods along with the

corresponding parameters are allowed to be chosen, and optimum models are generated by maximizing q^2. k-Nearest Neighbor Molecular Field Analysis (kNN-MFA) requires suitable alignment of given set of molecules. This is followed by generation of a common rectangular grid around the molecules. The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values are considered for relationship generation and utilized as descriptors to decide nearness between molecules.

Nearest Neighbor (kNN) Method

The k-NN methodology relies on a simple distance learning approach whereby an unknown member is classified according to the majority of its k-nearest neighbors in the training set. The nearness is measured by an appropriate distance metric (e.g., a molecular similarity measure calculated using field interactions of molecular structures). The standard k-NN method is implemented simply as follows: Calculate distances between an unknown object (u) and all the objects in the training set; select k objects from the training set most similar to object u, according to the calculated distances; and classify object u with the group to which the majority of the k objects belongs. An optimal k value is selected by optimization through the classification of a test set of samples or by leave one out cross-validation [26]. The variables and optimal k values are chosen using different variable selection methods as described below.

kNN-MFA with Simulated Annealing

Simulated Annealing (SA) [27] is another stochastic method for function optimization employed in QSAR. Simulated annealing (SA) is the simulation of a physical process, 'annealing', which involves heating the system to a high temperature and then gradually cooling it down to a preset temperature (e.g., room temperature). During this process, the system samples possible configurations distributed according to the Boltzmann distribution so that at equilibrium, low energy states are the most populated.

kNN-MFA with Stepwise (SW) Variable Selection

This method employs a stepwise variable selection procedure combined with kNN to optimize the number of nearest neighbors (k) and the selection of variables from the original pool as described in simulated annealing.

kNN-MFA with Genetic Algorithm

Genetic algorithms (GA) [28] first described by Holland mimic natural evolution by modeling a dynamic population of solutions. The members of the population, referred to as chromosomes, encode the selected features. The encoding usually takes form of bit strings with bits corresponding to selected features set and others cleared. Each chromosome leads to a model built using the encoded features. By using the training data, the error of the model is quantified and serves as a fitness function. During the course of evolution, the chromosomes are subjected to crossover and mutation. By allowing survival and reproduction of the fittest chromosomes, the algorithm effectively minimizes the rror function in subsequent generations.

Alignment rules

Molecular alignment was used to visualize the structural diversity in the given set of molecules. This was followed by generation of common rectangular grid around the molecules. The template structure, i.e. unsubstituted 4-thiazolidinone was used for alignment by considering the common elements of the series as shown in Figure 1. The reference molecule 18 is chosen high protective effect which made it a valid lead molecule and therefore was chosen as a reference molecule. After optimizing, the template structure and the reference molecule were used to superimpose all molecules from the series using the template alignment method. kNN-MFA method requires suitable alignment of given set of molecules after optimization; alignment was carried out by template based alignment method Stereoview of aligned molecules in training set and test set is shown in Figure 2.

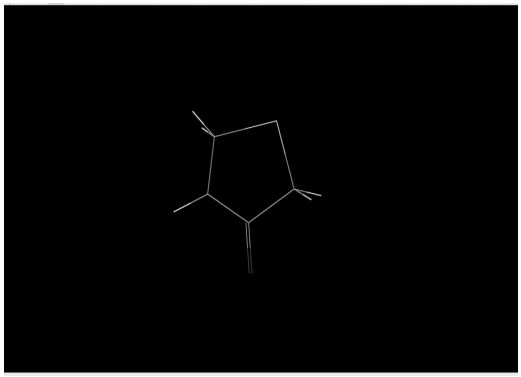

Fig. 1: Template molecule: Unsubstituted 4-thiazolidinone

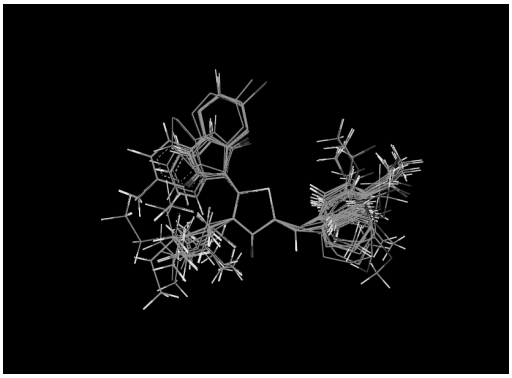

Fig. 2: Template based alignment of 4-thiazolidinones and indolin 2-one hybrid derivatives

Creation of interaction energies

Methyl probe with charge 1 and energy cut-off for electrostatic 10 Kcal/mol and for steric 30 Kcal/mol, dielectric constant 1 and charge type Gasteiger-marsili were used to calculate steric and electrostatic fields. The fields were computed at each lattice intersection of a regularly spaced grid of 2.0A° within defined three-dimensional region. The total grid points generated were 4675, steric and electrostatic fields generated were scaled by the standard method in the software.

Generation of training and test sets

In order to evaluate the QSAR model, data set was divided into training and test set using sphere exclusion, random selection and Manual selection method. Training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the predictive power of the model which is not included in model generation.

Results And Discussion

2D-QSAR models

Different sets of 2D-QSAR models were generated using the multiple linear regression analysis in conjunction with stepwise forward-backward variable selection method. Different training and test set were constructed using sphere exclusion, random and manual selection method. Training and test set were selected if they follow the unicolumn statistics, i.e., maximum of the test is less than maximum of training set and minimum of the test set is greater than of training set, which is prerequisite for further QSAR analysis. This result shows that the test is interpolative i.e., derived from the min-max range of training set. The mean and standard deviation of the training and test set provides insight to the relative difference of mean and point density distribution of the two sets.

The models for antitumor activity are shown in Table 2.

Model	r^2	q^2	pred_r^2	r^2se	q^2se	pred_r^2se	F test
1	0.8363	0.5888	0.4118	0.2637	0.4180	0.5771	19.4139
2	0.8319	0.5906	0.4060	0.2699	0.4212	0.5739	18.8121
3	0.6264	0.6227	0.3846	0.2764	0.4076	0.5578	18.0912

Table 2: Statistical results of 2D-QSAR models for antitumor activity

The selection of the best model is based on the values of r^2 (squared correlation coefficient), q^2 (cross-validated correlation coefficient), pred_r^2 (predicted correlation coefficient for the external test set), F (Fisher ratio) reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F–test indicate that the model is statistically significant. r^2 se, q^2se and pred_r^2 se are the standard errors terms for r^2, q^2 and pred_r^2 respectively. The statistically significant 2D-QSAR model is shown as follows.

Model-1 (Test set size: 4,9,12,15,16,18,19,23,26,32,33,35,37)

pKi = + 1.3163(± 0.0405) T_O_S_5 - 0.1402(± 0.0008) T_2_O_6 - 0.1505(± 0.0021) SdssCE-index - 11.8788(± 3.5811) chi6chain - 0.3422(± 0.0477) T_N_F_4 + 3.4167

Statistics:

n = 25; Degree of freedom = 19; r^2 = 0.8363; q^2 = 0.5888; F test = 19.4139; r^2 se = 0.2637; q^2 se = 0.4180; pred_r^2 = 0.4118; pred_r^2se = 0.5771

In this equation n is the number of molecules (Training set) used to derive the QSAR model. The QSAR model was obtained by using manual method (biological activity sorted in ascending manner) of training and test set data selection, where 25 of the total molecules were selected for training set while remaining 13 were selected as test set molecules.

The Mode-1 explains 83.63 % (r^2= 0.8363) of the total variance in the training set as well as it has internal (q^2) and external (pred_r^2) predictive ability of 58.88% and 41.88 % respectively. The F test = 19.4139 shows the statistical significance of 99.98% of the model. In addition randomization test shows confidence of 99.9% that the generated model is not random and hence it can be selected as the QSAR model.

To determine the reliability and significance of this generated model, the leave-one-out test and randomization tests were employed. From the cross- validation test, q^2 (0.5888) indicated that the result obtained for above best 2D-QSAR model was not by chance correlation.

Model	k-NN	q^2	pred_r^2	q^2se	DOF
1	2	0.7773	0.3005	0.1478	18
2	2	0.6766	0.3813	0.2309	18
3	2	0.5377	0.4428	0.3121	18

Table 3: Statistical results of 3D-QSAR models for antitumor activity

Compounds No.	Actual Activity	Predicted Activity Model-1 (2D-QSAR)	Predicted Activity Model-2 (3D-QSAR)
1(a)	1.7517	1.7517	ND
2(b)	0.1746	-0.009	0.0807
3(b)	0.4481	0.248	0.7961
4(b)	0.6673	0.104	0.1757
5(b)	0.3125	0.155	0.2883
6(b)	0.1811	0.307	0.1537
7(b)	0.1741	0.045	0.2089
8(b)	0.1838	0.342	0.4769
9(b)	1.4984	1.65	ND
10(b)	1.1418	0.806	0.5272
11(b)	1.2761	1.776	ND
12(b)	0.3464	0.505	0.3225
13(b)	1.4408	1.025	0.7955
14(b)	1.2128	1.312	ND
15(b)	1.2173	0.854	0.8744
16(b)	1.2747	1.804	ND
17(b)	0.0668	-0.086	0.1090
18(b)	0.2496	0.605	0.4623
19(b)	0.0093	0.449	0.5409
20(b)	0.3447	0.23	0.2918
21(c)	0.05221	0.009	0.2882
22(c)	0.2349	0.27	0.1764
23(c)	0.3472	-1.125	0.0144
24(c)	-0.0052	0.281	0.2882
25(c)	0.2227	0	0.0124
26(c)	0.6587	0.114	0.0206
27(c)	0.3062	0.049	0.0524
28(c)	0.3892	0.738	0.9359
29(c)	0.3318	0.367	0.2895
30(c)	0.5598	0.785	0.6734
31(c)	0.0347	0.223	0.4409
32(c)	0.2175	0.427	0.4361
33(c)	0.3297	-0.108	0.4624
34(c)	-0.0227	0.149	0.2372
35(c)	0.2053	0.476	0.5265
36(c)	0.1311	-0.059	0.2403
37(c)	1.4103	1.054	ND
38(c)	1.3962	1.713	ND

ND = Not determine

Table 4 represents the predicted biological activity by the Model-1 for training and test set respectively.

Table 4: Actual and predicted activities of Model-1 (2D-QSAR) and Model-2 (3D-QSAR) for antitumor activity

In the QSAR equation, the positive coefficient value of T_O_S_5 [count of number of oxygen atoms (single, double or triple bonded) separated from sulphur atom by 5 bond distance in molecule] on the biological activity indicated that higher T_O_S_5 value leads to better antitumor activity.

The negative coefficient value of T_2_O_6 [count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from oxygen atom by 6 bonds], SdssCE-index

[electrotopological state indices for number of carbon atom connected with double and two single bonds], chi6chain [ ], T_N_F_4 [count of number of nitrogen atoms (single, double or triple bounded)separated from any fluorine atom(single or double bonded) by 4 bond distance in molecule] descriptors indicated that decrease value lead to better antitumor activity. Figure 3 represents the contribution chart showing contribution of the various descriptors playing important role in determining the antitumor activity. It reveals that the descriptors T_O_S_5, contributes 37% and T_2_O_6, SdssCE-index, chi6chain, T_N_F_4 inversely contribute 16%, 12%, 17%, 21% respectively to biological activity. Figure 3 represent the data fitness plot of actual vs predicted activity for Model-1, provides an idea about how well the Model-1 was trained and how well it predicts the activity of the external test set. From Figure 4 it can be seen that the model is able to predict the activity of the training set quiet well as well as external test set, providing confidence of the Model-1.

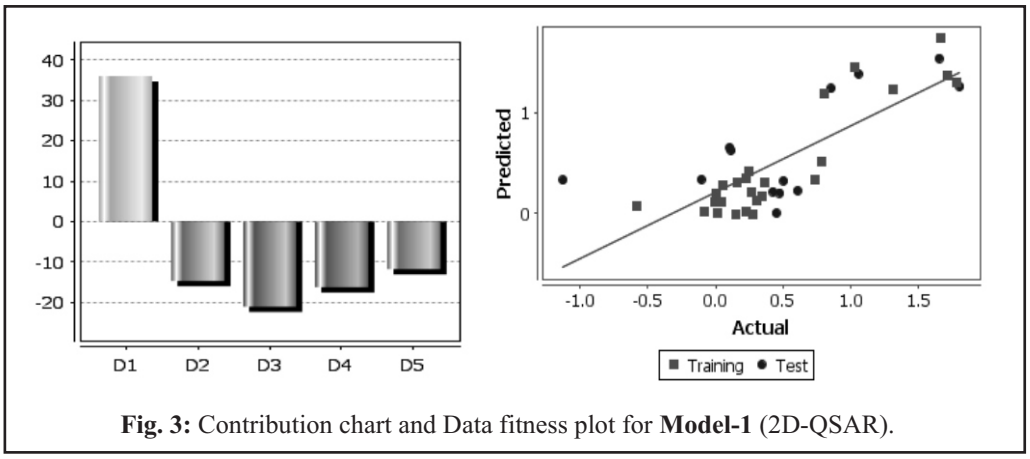

Fig. 3: Contribution chart and Data fitness plot for **Model-1** (2D-QSAR).

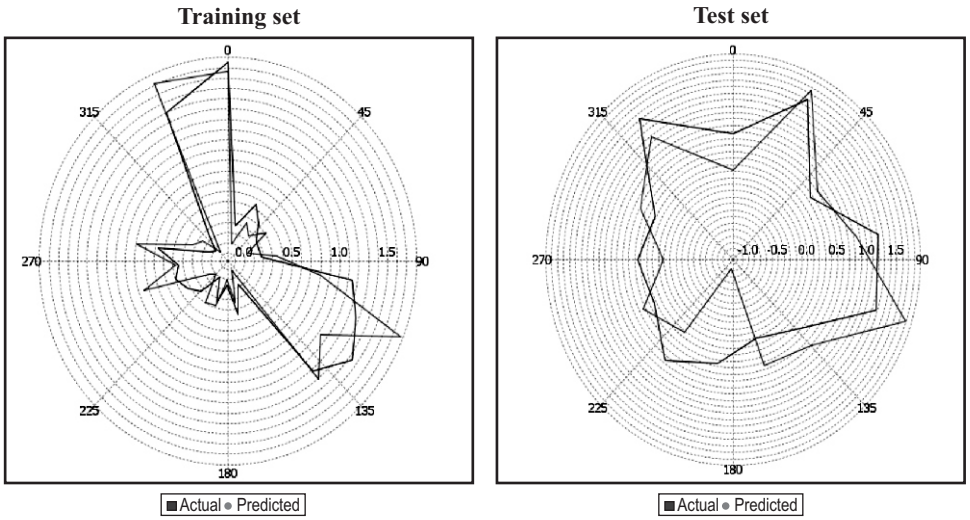

Fig. 4 : Graph between actual and predicted biological activity for training and test sets for Model-1(2D-QSAR).

3D-QSAR model

3D-QSAR was studied on 31 molecules. Because when study was carried out on 38 molecules satisfactory results were not obtained. So those molecules which were potent enough were used for 3D-QSAR study and remaining compounds (1, 9, 11, 14, 16, 37, and 38) were excluded.

Different sets of 3D-QSAR model were generated by kNN-MFA include both the electrostatic, steric descriptors along with their range to indicate their importance for interaction in molecular field. Different training and test set were constructed using sphere exclusion, random and manual selection method. Training and test set were selected if they follow the unicolumn statistics, i.e., maximum of the test is less than maximum of training set and minimum of the test set is greater than of training set, which is prerequisite for further QSAR analysis. This result shows that the test is interpolative i.e., derived from the min-max range of training set. The mean and standard deviation of the training and test set provides insight to the relative difference of mean and point density distribution of the two sets.

The statistical significant 3D-QSAR models for antitumor activity are given in Table 3. The selection of the best model is based on the values of q^2 (the internal predictive ability of the model) and that of pred_r^2 (the ability of the model to predict the activity of external test set). As the cross-validated correlation coefficient (q^2) is used as a measure of reliability of prediction, the correlation coefficient suggests that our model is reliable and accurate. The statistical significant 3D-QSAR model is given below.

Model-2 (Test set size 2,4,7,15,16,18,19,20,25,28,31)

pKi = S_1369 (-0.115, -0.100) E_1021 (-1.437, -1.334)

Statistics:

n= 17, Degree of freedom = 18, q^2 = 0.7773, q^2se = 0. 1478, pred_r^2 =0.3813, pred_r^2 se =0.5231

The Model-2 explains values of k (2), q^2 (0.7773), pred_r^2 (0.3813), q^2_se (0.1478), and pred_r^2 se (0.5231) prove that QSAR equation so obtained is statistically significant and shows the predictive power of the Model-2 is 77.73% (internal validation). Table 4 represents the predicted biological activity by the Model-2 for training and test set respectively. Figure 5 represents the data fitness plot of actual vs predicted activity for Model-2, provides an idea about how well the model was trained and how well it predicts the activity of the external test set.

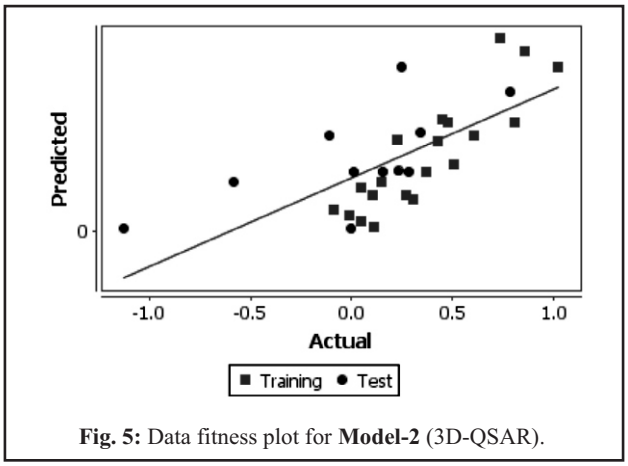

Fig. 5: Data fitness plot for **Model-2** (3D-QSAR).

From **Figure 6** it can be seen that the model is able to predict the activity of the training set quiet well as well as external test set, providing confidence of the model.

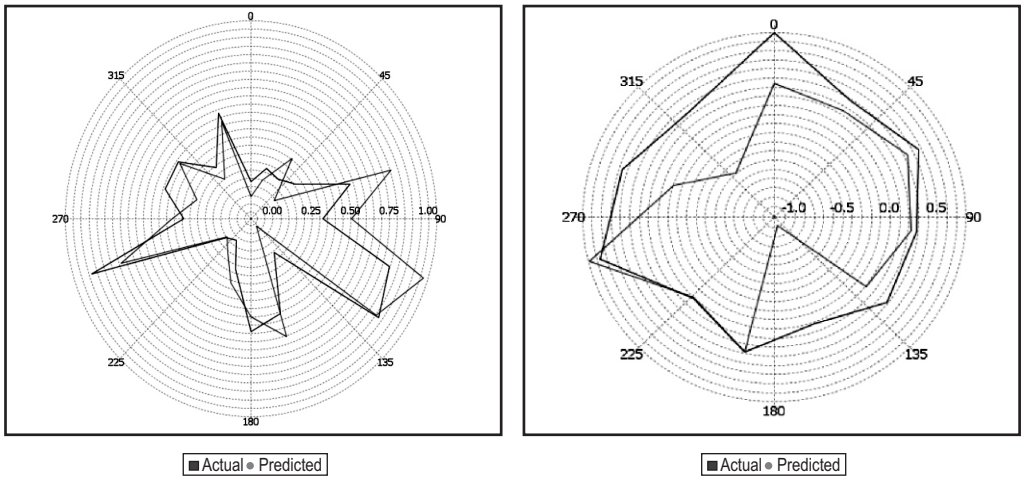

Fig. 6: Graph between actual and predicted biological activity for training and test sets for Model-2 (3D-QSAR).

Result plot in which 3D-alignment of molecules with the important steric and hydrophobic points contributing in the model with ranges of values shown in the parenthesis represented in Figure 7. It shows the relative position and ranges of the corresponding important steric and electrostatic fields in the model provides guideline for new molecule design as follows-

(a) Steric filed, S_1369 (-0.115, -0.100) has negative range indicates that negative steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.

(b) Electrostatic field, E_1021 (-1.437, -1.334) has negative range indicates that negative electrostatic potential is favorable for increase in the activity and hence more electronegative substituent group is preferred in that region.

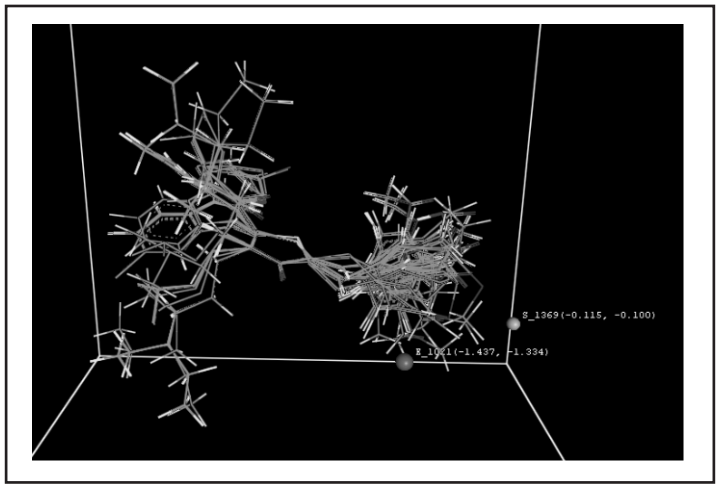

Fig. 7: 3D-alignment of molecules with the important steric and electrostatic points contributing Model-2 (3D-QSAR) with ranges of values shown in parenthesis.

Taking clues from the above mentioned guidelines and looking at the developed model field plot and corresponding important steric and electrostatic field range which shows the ranges are towards negative side (S_1369) meaning less bulky substituent group is preferred and negative side for (E_1021) meaning more electronegative substituent group is preferred at the respective sites.

In this presented work, statistically significant 2D/3D-QSAR models were generated with the purpose of deriving structural requirements for the antitumor activities of some 4-thiazolidinone and indolin-2-one hybrid derivatives. The validation of models was done by cross-validation test, randomization test and external test set prediction. The best 2D-QSAR models indicated that the descriptors T_O_S_5, T_2_O_6, SdssCE-index contributing 37%, 17%, 21% respectively and chi6chain, T_N_F_4 contributing 16%, 12% respectively to biological activity. The best 3D-QSAR model constructed by kNN-MFA method, providing useful information on substitutional requirements for the favorable antitumor activity and characterization and differentiation of their binding sites. In conclusion, the information provided by the robust 2D/3D-QSAR models use for the design of new molecules and hence, this method is expected to provide a good alternative for the drug design.

Acknowledgment

The author wishes to express gratitude to V-life Science Technologies Pvt. Ltd. for providing the software for the study. Also the authors are thankful to the principal Dr. P. S. Gide and the trustee Bhujbal Knowledge City for providing the necessary facilities to carry out the research work.

References

1. Tomasic T, Masic LP, Rhodanine as a Privileged Scaffold in Drug Discovery, Curr. Med. Chem. 2009; 16:1596-1629.
2. Lesyk R, Zimenkovsky B, Atamanyuk D, Jensen F, Kiec-Kononowicz K, Gzella A, Anticancer thiopyrano [2,3-d] [1,3] thiazol-2-ones with norbornane moiety. Synthesis, cytotoxicity, physicochemical properties, and computational studies, Bioorg. Med. Chem. 2006; 14:5230-5240.
3. Havrylyuk D, Zimenkovsky B, Vasylenko O, Zaprutko L, Gzella A, Lesyk R, Synthesis of novel thiazolone-based compounds containing pyrazoline moiety and evaluation of their anticancer activity, Eur. J. Med. Chem. 2009;44:1396-1404.
4. Kaminskyy D, Zimenkovsky B, Lesyk R, Synthesis and in vitro anticancer activity of 2,4-azolidinedione-acetic acids derivatives, Eur. J. Med. Chem. 2009; 44:3627-3636
5. Ottana R, Maccari R, Brreca ML, Bruno G, Rotondo A, Rossi A, Chiricosta G, Di Paola R, Sautebin L, Cuzzocrea S, Vigorita MG, 5-Arylidene-2-imino-4-thiazolidinones : Design and synthesis of novel anti-inflammatory agent, Bioorg. Med. Chem. 2005; 13:4243-4252.
6. Vicini P, Geronikaki A, Anastasia K, Incerti M, Zani F, Synthesis and antimicrobial activity of novel 2- thiazolylimino-5-arylidene-4-thiazolidenones, Bioorg. Med. Chem. 2006; 14:3859-3864.
7. Elbarbary AA, Khodair AI, Pedersen EB, Nielsen C, Synthesis and evaluation of antiviral activity of 2'-deoxyuridines with 5-methylene-2-thiohydantoin substituents in 5-position, Monatsh. Chem. 1994; 125:593-598.
8. Rydzik E, Szadowska A, Kaminska A, Synthesis of benzylidene derivatives of 3-o,3-m and 3-p-chlorophenyl hydantoin and the study of their anticonvulsant action, Acta Pol. Pharm. 1984; 41:459-464.
9. Liu HL, Li ZC, Anthonsen T, Synthesis and Fungicidal Activity of 2-limino-3-(4-arylthiazol-2-yl)-thiazolidin-4-ones and Their 5-Arylidene Derivatives, Molecules 2000; 5:1055-1061.
10. Samir B, Wesam K, Ahmed AF, Synthesis and antimicrobial evaluation of some new thiazole, thiazolidinone and thiazoline derivatives starting from 1-chloro-3,4-dihydronaphthalene-2-carboxaldehyde, Eur. J. Med. Chem. 2007; 42:948-954.
11. Ahn JH, Kim SJ, Park WS, Cho SY, Ha JD, Kim SS, Kang SK, Jeong DG, Jung S-K, Lee S-H, Kim HM, Park SK, Lee KH, Lee CW, Ryu SE, Choi J-K, Synthesis and biological evaluation of rhodanine derivatives as PRL-3 inhibitors, Bioorg. Med. Chem. Lett. 2006; 16:2996-2999.
12. French KJ, Schrecengost RS, Lee BD, Zhuang Y, Smith SN, Eberly JL, Yun JK, Smith CD, Discovery and Evaluation of Inhibitors of Human Sphingosine Kinase, Cancer Res. 2003; 63:5962-5969.
13. Cutshall NS, O'Day C, Prezhdo M, Rhodanine derivatives as inhibitors of JSP-1, Bioorg. Med. Chem. Lett. 2005; 15:3374-3379
14. Geronikaki A, Eleftheriou P, Vicini P, Alam I, Dixit A, Saxena AK, 2-Thiazolylimino/Heteroarylimino-5-arylidenethiazolidinones as New Agents with SHP-2 Inhibitory Action, J. Med. Chem. 2008; 51:5221-5228.

15. Ottana R, Carotti S, Maccari R, Landini I, Chiricosta G, Caciagli B, Vigorita MG, Mini E, Invitro antiproliferative activity against human colon cancer cell lines of representive 4-thiazolidinones. Part-1, Bioorg. Med. Chem. Lett. 2005; 15:3930-3933.

16. Krug M, Hilgeroth A, Recent advances in the development of multi-kinase inhibitors Mini-Rev. Med. Chem. 2008; 8:1312-1327.

17. Sun L, Liang C, Shirazian S, Zhou Y, Miller T, Cui J, Fukuda JY, Chu J-Y, Nematalla A, Wang XY, Chen H, Sistla A, Luu TC, Tang F, Wei J, Tang C, Discovery of 5-[5-fluro-2-oxo-1,2-dihydroindol-(3Z)- ylienemethyl]-2,4-dimethyl-1. H-pyrrole-3-carboxylic acid (2-diethylaminoethyl) amide, a Novel Tyrosine Kinase Inhibitor Targeting Vascular Endothelial and Platelet-Derived Growth Factor Receptor Tyrosine Kinase, J. Med. Chem. 2003; 46:1116-1119.

18. Roth GJ, Heckel A, Colbatzky F, Handschuh S, Kley J, Lehmann-Lintz T, Lotz R, Tontsch-Grunt U, Walter R, Hilberg F, Design Synthesis and Evaluation of Indolinones as Triple Angiokinase Inhibitors and the Discovery of a Highly Specific 6-Methoxycarbonyl-Substituted Indolinone (BIBF 1120), J. Med. Chem. 2009; 52:4466-4480.

19. Xiao Z, Hao Y, Liu B, Indirubin and Meisoindigo in the Treatment of Chronic Myelogenous Leukemia in China, Leuk. Lymphoma. 2002; 43:1763-1768.

20. Wang S, Zhao Y, Zhang G, Lv Y, Zhang N, Goag P, Design, synthesis and biological evaluation of novel 4-thiazolidinones containing indolin-2-one moiety as potential antitumor agent, Eur. J. Med. Chem. 2011; 46:3509-3518.

21. Hansch C, Kurup A, Garg R, Gao H, Chem-Bioinformatics and QSAR: A Review of QSAR Lacking Positive Hydrophobic Terms, Chem. Rev. 2001; 101:619–672.

22. Lill MA, Multi-dimensional QSAR in drug discovery, Drug Discovery Today. 2007; 12:1013–1017.

23. Yang GF, Huang XQ, Development of Quantitative Structure-Activity Relationships and Its Application in Rational Drug Design, Curr. Pharm. Des. 2006; 12:4601–4611.

24. MMC. Ferreira, Multivariate QSAR, J. Braz Chem Soc. 2002; 13:742.

25. VLifeMDS, Molecular Design Suite, Vlife Sciences Technologies Pvt. Ltd., Pune, India. www.vlifesciences.com

26. Sharaf MA, Illman DL, Kowalski BR, Chemometrics, Welly, New York, 1986.

27. Kirkpatrick S, C.D. Gelatt J, Vecchi MP, In Readings in computer vision: issues, problems, principles, and paradigms, Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1987.

28. Michalewicz Z, Genetic algorithms + data structures = evolution programs (3rd ed.), Springer-Verlag: London, UK, 1996.