Available online at www.derpharmachemica.com

Scholars Research Library

Der Pharma Chemica, 2010, 2(6): 125-133
(http://derpharmachemica.com/archive.html)

ISSN 0975-413X

2D and 3D-QSAR Analysis of some 1, 3-disubstituted Urea Derivatives for Antiproliferative Activity

Sanmati K. Jain[2], Ravindra D. Dubey[1]*, S. Mallick[2], S. Nag[2], A. Yadav[2], A. Mishra[3] and Naveen K. Mahobia[4]

[1]*Institute of Pharmacy, RITEE, Mandir Hasaud, Chhatauna, Raipur, Chhattisgarh, India*
[2]*SLT Institute of Pharmaceutical Sciences, Guru Ghasidas Vishwavidyalaya, Bilaspur, Chhattisgarh, India*
[3]*Shri Shankracharya Institute of Pharmaceutical Sciences, Bhilai Chhattisgarh, India*
[4]*Shri Leuva Patel Trust Pharmacy Mahila College, Amreli, Gujarat, India*

ABSTRACT

The 2D and 3D quantitative structure activity relationship study was performed on a series of some 1,3-disubstituted urea derivatives possessing antiproliferative activity for establishing quantitative relationship between biological activity and their physicochemical properties. Several statistical regression expressions were obtained with 2D-QSAR study using stepwise multiple linear regression (MLR), partial least squares regression (PLSR), and partial component regression (PCR) analysis and two statistical significant models were generated (r^2 = 0.8195, 0.8283 for model 1 and 2 respectively) indicating that biological activity is influenced by the descriptors T_N_O_3 and SsOHE-index. With 3D- QSAR anlysis one statistical significant models were generated (r^2 = 0.8668 and $predr^2$ = 0.8307) by using kNN Random data selection method.

Key words: QSAR, antiproliferative activity, statistical regression.

INTRODUCTION

1,3-disubstituted Urea derivatives are wide range of biological activities, such as antiproliferative [1] anticonvulsant activity. [2] colchicine-blinding antagonist [3] and CXCR3 antagonist. [4] Some urea derivatives are antitumerogenic and exhibit cell growth inhibition (GI50) on numerous cancer cell lines and have developed various chemoresistance mechanisms, and to be potent inhibitors of the PTK activity of a number of trans-membrane growth factor receptors and cellular oncogene products, particularly epidermal growth factor receptor (EGFR). [5] Many trans-membrane growth factor receptors possess intracellular PTK activity, with initiation of this

activity following external binding of a growth factor being the first step in the cellular signal transduction pathway which controls mitogenesis and cell proliferation. [6][7] The over-expression or inappropriate expression of normal or mutant PTK activity in these receptors can thus result in loss of growth control and the unregulated cell proliferation associated with malignancy. [8][9] Therefore, selective interruption of signal transduction in tumor cells by specific inhibitors of PTK activity has recently emerged as a major new approach for the design of tumorspecific drugs. [10][11] Some of the kinase inhibitors used in treating cancer are inhibitors of tyrosine kinase. [12] They can interfere with the repair of DNA double-strand breaks. [13] Several researches reported that urea-like compounds on the human-leukemia K562 cell line found that incorporating a heterocycle, especially a morpholine ring can highly improve their activity. No molecular modeling QSAR studies have been done till date on this class of compounds, so we attempt to subject this new class of compounds for an effective QSAR analysis.

MATERIALS AND METHODS

In the present study a dataset of 24 molecules of 1,3-disubstituted urea derivatives were used which were reported to have antiproliferative activity. Antiproliferative activity data listed in Table No. 01 (n=24), have been used in this study. The activity data IC50 (μM), determined with human tumor cell line (K562). All compounds have been converted to the logarithmic scale [pIC50 (moles)], and then used for subsequent QSAR analysis as the response variables. pIC50 data saved as .txt.file.

All work done in Structure of the molecule drawn in the 2D Draw application in Tool menu of QSARPlus of Molecular Design Suit [MDS] software. [14] Then 2D structures where exported to QSAR Plus window (2D structure converted to 3D structure). After the conversion, structures are saved as .mol2 file in QSARPlus 3D window. Energy minimization is done by using the force field or minimizing energy with the help of MMFF force field which result in the optimization of the geometry of the molecule. After drawing all the molecules they were again optimized by using batch minimize option using MMFF force field and optimized molecules were used to calculate the physicochemical and alignment descriptors.

For model development in 2D-QSAR analysis three methods Random selection method, Manuel data selection method, Sphere Exclusion method were used for creation of training and test set and 10 trials were run in each case. After the creation of training and test set, minimum and maximum value of the test and training set is checked, using the QSAR tool, then the different statistical methods like Multiple liner regression (MLR), Partial least squares regression (PLSR), Partial component regression (PCR) were used for model building.

Table 01: General structure of the molecules of 1,3-disubstituted urea derivatives and their biological activities and data set of 24 molecules.

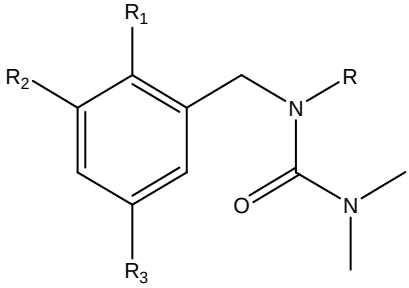
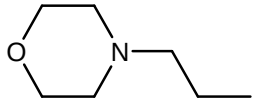
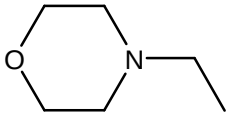
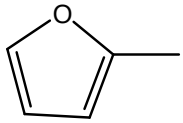
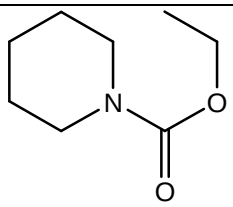
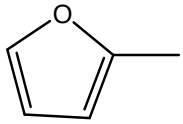
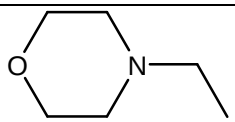
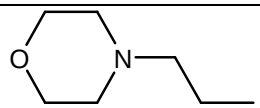
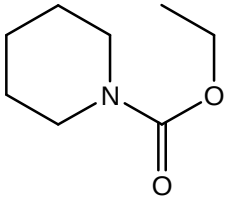
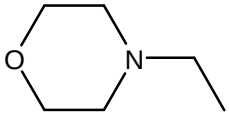
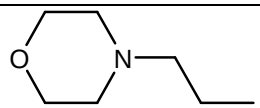
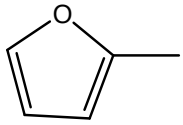
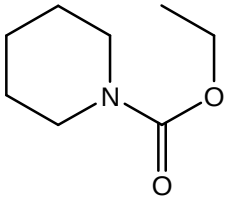
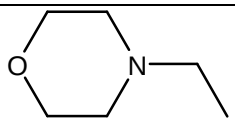
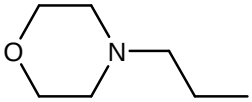
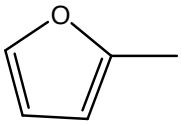
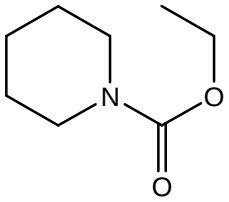
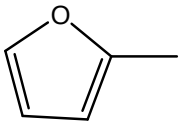
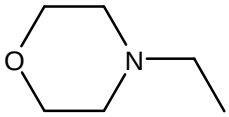
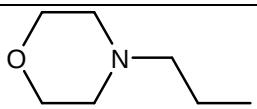
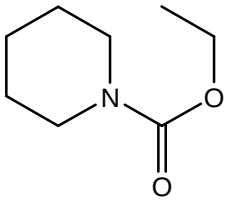
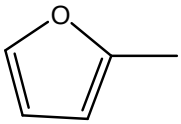
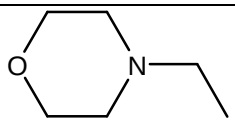
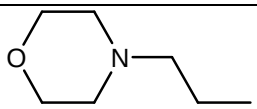
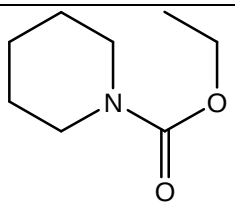


Figure-1: General structure of 1,3-disubstituted urea derivatives

Molecule no.	R	R_1	R_2	R_3	IC50 (μM) K562	log (1/IC50) K562
1		NO_2	H	H	0.42	6.376
2		NO_2	H	H	0.27	6.568
3		NO_2	H	H	40.7	4.390
4		NO_2	H	H	24.9	4.603
5		Cl	H	OH	71.3	4.146
6		Cl	H	OH	3.88	5.411
7		Cl	H	OH	2.92	5.534
8		Cl	H	OH	33.5	4.474
9		H	OH	H	11.8	4.928
10		H	OH	H	7.5	5.124
11		H	OH	H	75.1	4.124
12		H	OH	H	67.2	4.172

13		H	NO_2	H	2.24	5.649
14		H	NO_2	H	1.2	5.920
15		H	NO_2	H	35.4	4.450
16		H	NO_2	H	26	4.585
17		NO_2	H	OH	28.6	4.543
18		NO_2	H	OH	10.8	4.966
19		NO_2	H	OH	8.5	5.070
20		NO_2	H	OH	38.4	4.415
21		Br	H	OH	40.5	4.392
22		Br	H	OH	13.5	4.869
23		Br	H	OH	18.6	4.730
24		Br	H	OH	41.8	4.378

For 3D-QSAR analysis Models were generated by using advanced statistical method kNN. In kNN method Stepwise variable selection (forward-backward), Simulated annealing, and Genetic algorithm method is used. For creation of training and test set same method were applied as 3D-QSAR analysis.

RESULTS AND DISCUSSION

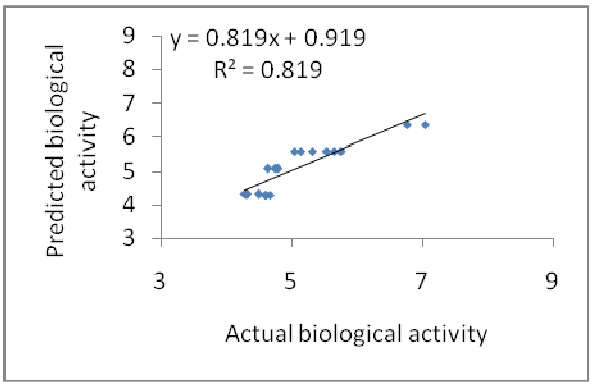
All the 24 molecules of the selected series were subjected to various regression analysis, the following significant 2D-QSAR models with equations were obtained for antiproliferative activity. (Table-02)

Table-02: List of predictive QSAR models with equation generated from various regression methods

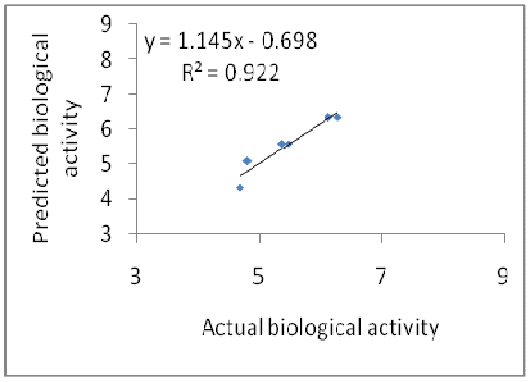
Model no.	Method	Equation
01	Random 75% / Trial-2 / MLR	Log1/IC50 = T_N_O_3. 1.28199 + SsOHE-index. -0.0803342 + 5.08293 n = 18 Degree of freedom = 15 F test = 34.0557 r2 = 0.8195 q2 = 0.7099 pred_r2 = 0.8858 r2 se = 0.3633 q2 se = 0.4606 pred_r2se = 0.2599
02	Random 90% / Trial-10 / MLR	Log1/IC50 = T_N_O_3. 1.26516 + SsOHE-index. -0.0723667 + 5.03928 n = 21 Degree of freedom = 18 F test = 43.4202 r2 = 0.8283 q2 = 0.7533 pred_r2 = 0.9120 r2 se = 0.3470 q2 se = 0.4160 pred_r2se = 0.2310

The QSAR models presented in table no. 02, in which n is the number of data points used in the model. Pred_ r^2 is the predicted r^2 for external test set and se is the standard error of estimate (smaller is better). From the table no. 02, the equation of Model-01 explains 81% (r2=0.81) of the total variance in the training set as well as it has internal (q2) and external (pred_r2) predicative ability of 70% and 88% respectively. Model-02 explains 82% (r2= 0.82) of the total variance in the training set as well as it has internal (q2) and external (pred_r2) predicative ability of 75% and 91% respectively.

Graphs were plotted between the actual and the predicted biological activities for model 1 and 2 shown in Figure-2 and 3 respectively.

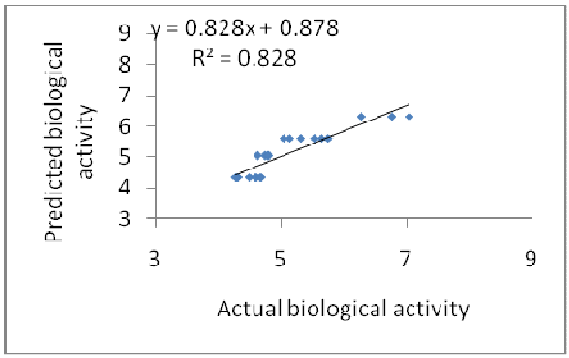


(Training set)

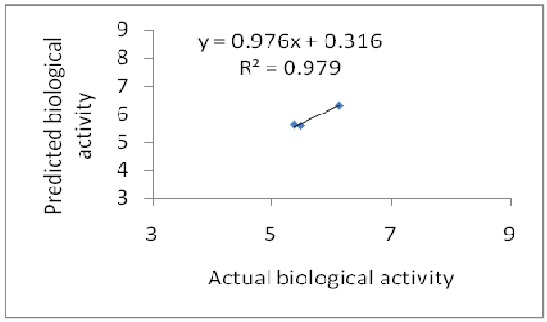


(Test set)

Figure-2: Graph between actual and predicted biological activity for training and test set of Model-01



(Training set)



(Test set)

Figure-3: Graph between actual and predicted biological activity for training and test set of Model-02

The two models with their r^2 values produced from actual vs. predicted activity graph is shown in Figure 2 and 3. The Figure represents the models which are having $r^2 > 0.81$, indicating their predictive ability in predicting the activity of the test set molecules.

Table-03: Correlation matrix for descriptors used in model-01

	T_N_O_3	*SsOHE-index*
T_N_O_3	1	
SsOHE-index	0.060541	1

Table-04: Correlation matrix for descriptors used in model-02

	T_N_O_3	*SsOHE-index*
T_N_O_3	1	
SsOHE-index	0.004967	1

The correlation matrix is used to see the mutual correlation among the parameters used in the model. The correlation matrix shows that descriptors have low inter-correlation value. The contribution charts for model 01 and 02 are shown in Figure-04 and 05 respectively.

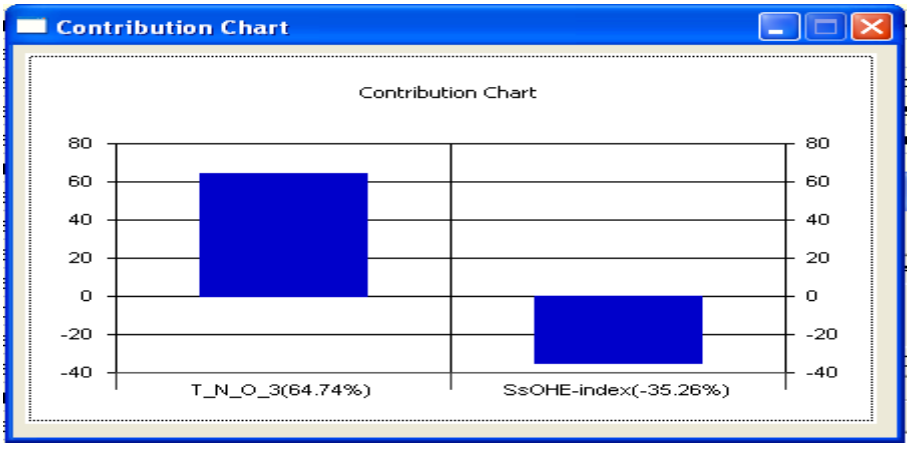


Figure-4: Contribution chart of various descriptors in biological activity for Model-01

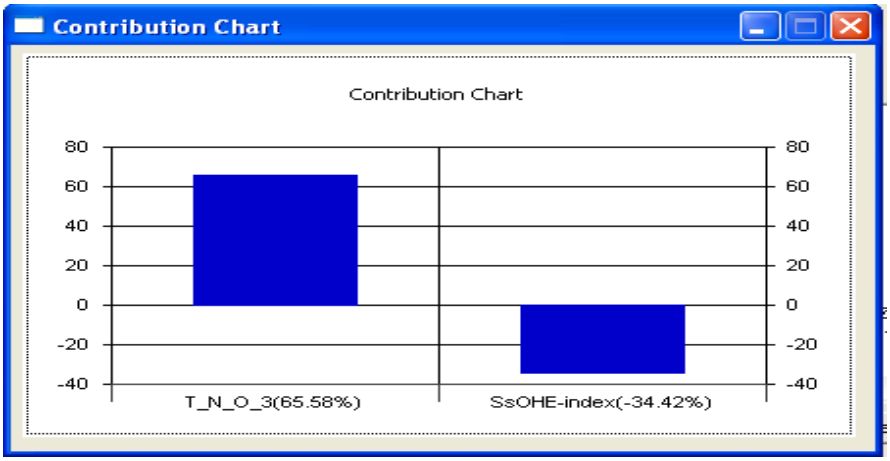


Figure-5: Contribution chart of various descriptors in biological activity for Model-02

In the present study MLR (coupled with stepwise forward- backward variable selection) lead to a statistical significant model. The developed model no. 01 reveals that the descriptors T_N_O_3 play an important role (≈64.74%) in determining antiproliferative activity. The other descriptor i.e. SsOHE- index is inversely proportional to activity (-35.26%). In Model no. 02 the descriptor T_N_O_3 play an important role (≈65.58%) in determining antiproliferative activity. The other descriptor i.e. SsOHE- index is inversely proportional to activity (-34.42%).

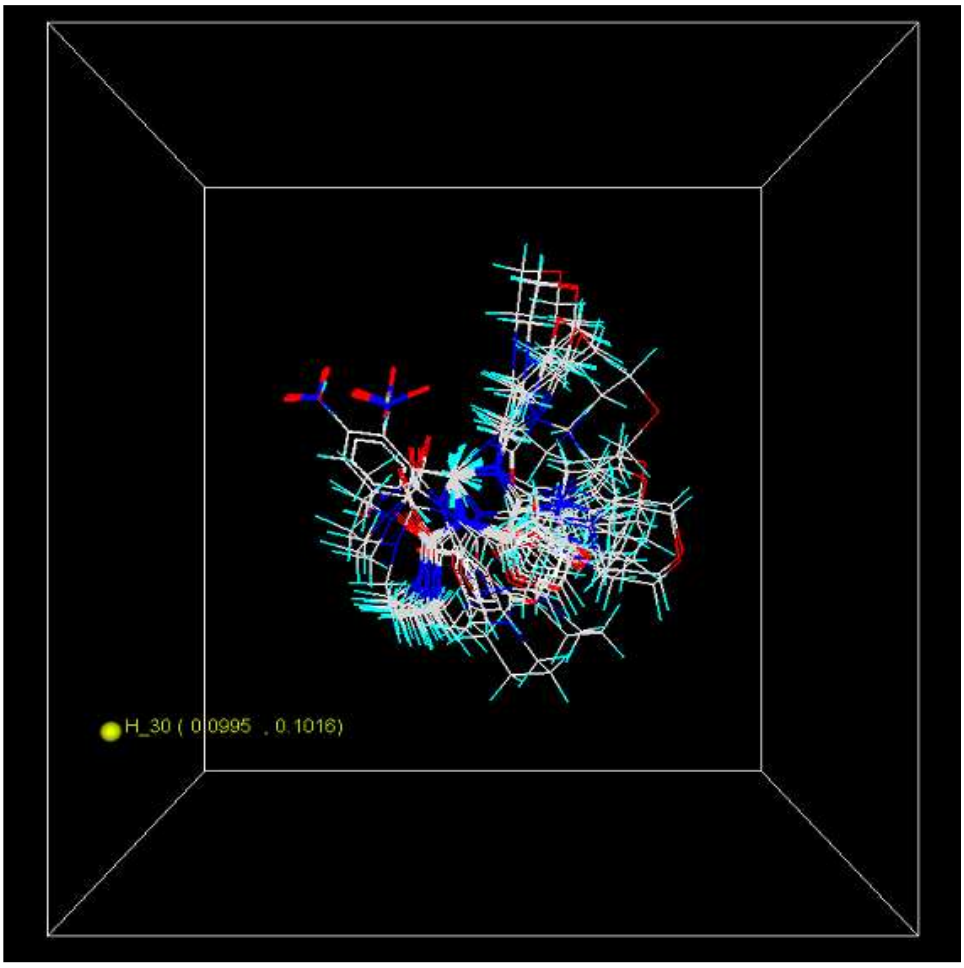


Figure-6: kNN-MFA result plot- 3D-alignment of molecules with the hydrophobic point in parenthesis.

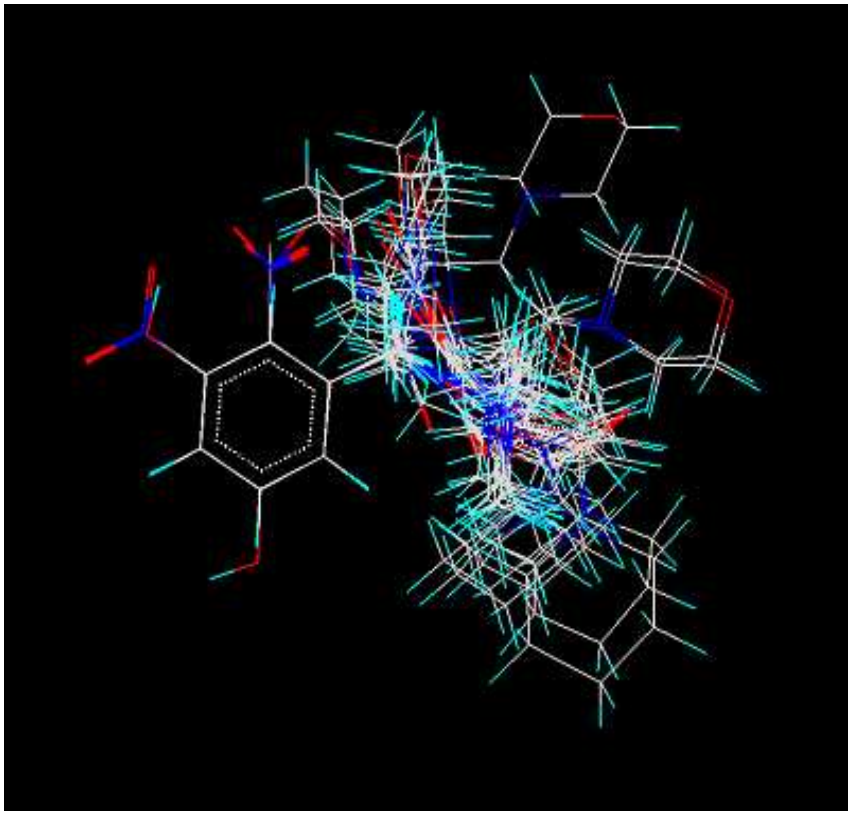


Figure-7: kNN-3D-alignment of molecules.

All the 24 molecules of the selected series of table-01 were subjected to various regression method for 3D-QSAR analysis. For result of the most predictive model generated by kNN stepwise forward – backward method using Random data is as follows selection method.

kNN Method (Trial 4, 85% data set): Training Set Size = 20, Test Set Size = 4.

Selected Descriptors: H_30.
Statistics: k Nearest Neighbour= 2, n = 20, Degree of freedom = 18, q2 = 0.8668, q2_se = 0.3042.
Predr2 = 0.8307, pred_r2se = 0.1741.

Descriptor Range: H_30 0.0995 0.1016.
This model explains very good internal (q2 = 0.8668) as well as external (Predr2 = 0.8307) predictive power of the model. The hydrophobic descriptor at the grid points H_30 play important role in imparting activity. This model indicates that one hydrophobic descriptor is involved. kNN-MFA result plot in which 3D-alignment of molecules with the important hydrophobic point contributing with model with ranges of values shown in parenthesis represented in Figure-6.

CONCLUSION

In the present study an attempt has been made to identify the necessary structural requirements molecule for activity. From the present QSAR analysis, three best models were generated two Model from 2D-QSAR study and Model from 3D-QSAR among which can be used for predicting the activity of the newly designed compounds in finding some more potent molecules.

Finally, it is concluded that the work presented here will play an important role in understanding the relationship of physiochemical parameters with structure and biological activity. By studying the QSAR model one can select the suitable substituent for active compounds with maximum potency.

REFERENCES

[1] H. Q. Li, T. T. Zhu, T. Yan, Y. Luo, H. L. Zhu, *Eur. J. Med. Chem.*, **2009**, 44, 453- 459.
[2] J. A. Shimshoni, M. Bialer, B. Wlodarczyk, R. H. Finnell, B. Yagen, *J Med. Chem.*, **2007**, 50, 6419-6427.
[3] J. S. Fortin, J. M. Lacroix, M. P. Desjardins, P. E. Alexandre, C. G. Rene, *Bioorg. Med. Chem.*, 15, **2007**, 4460-4469.
[4] R. J. Watson, D. R. Allen, H. L. Birch, G. A. Chapman, D. A. Owen, E. J. Thomas, N. Tremayne, S. C. Williams, *Bioorg. Med.Chem. Lett.*, **2008**, 18, 147-151.
[5] P. Bechard, J. Lacroix, P. Poyet, *Eur. J. Med. Chem.*, **1994**, 29, 963-966.
[6] A. Ullrich, J. Schlessinger, *Cell*, **1990**, *61*, 203-212.
[7] S. R. Hubbard, J. H. Till, *Annu. Rev. Biochem.*, **2000**, 69, 373-398.
[8] M. J. Hayman, P. J. Enrietto, *Cancer Cells*, **1991**, 3, 302-307.
[9] I. Dikic, S. Giordano, *Curr. Opin. Cell Biol*., **2003**, 15, 128-135.
[10] F. M. Uckun C. Mao, *Curr. Pharm. Des*., **2004**, 10, 1083-1091.
[11] P. Traxler, J. Green, H. Mett, U. Sequin, P Furet, *J. Med. Chem.*, **1999**, 42, 1018-1026.
[12] "Definition of tyrosine kinase inhibitor - NCI Dictionary of Cancer Terms". http://www.cancer.gov/templates/db_alpha.aspx?CdrID=44833.
[13] Y. Zhao, H. D.Thomas, M. A. Batey, *Cancer Res.*, **2006**, 66, 5354–5362.
[14] QSARPlus of Molecular Design Suit [MDS] VLife Sciences Technology Pvt. Ltd. Pune-411045, Web: www.vlifesciences.com