

QSAR STUDIES ON THIAZOLES HAVING ANTIPLATELET ACTIVITY

G. A. Kashid^{1*} J. Sarvanan² N. P. Jain¹

¹ Department of Pharmaceutical Chemistry, S N D College of Pharmacy, Babhulgaon, Yeola, Nashik, Maharashtra, India-423401,

² Department of Pharmaceutical Chemistry, P E S College of Pharmacy, Bangalore, Karnataka.

Received 03 April 2015; Revise 12 May 2015; Accepted 15 April 2015

Abstract

Recently thiazole derivatives were identified antiplatelet activity. Using with multiple regression method, partial least square method, principal component regression and kNN-MFA 3D-QSAR models were generated. One of these models was selected on the basis of q^2 and $pred_r^2$ values. The selected model has shown good internal and external prediction for training set of 16 molecules and test set of 4 molecules with validation q^2 and cross validation ($pred_r^2$). QSAR analysis of thiazole derivatives has been applied to develop of relationships between physicochemical properties of chemical substances and their antiplatelet activities, to obtain a reliable statistical model for prediction. The best model shown by 3D QSAR study was obtained from Partial least square regression forward method having r^2 value = 0.9665 and for 2D QSAR study was obtained from Multiple regression forward method r^2 value=0.9151 with good predictive ability. The results obtained from QSAR studies could be used in designing better anticancer agents among the congeners in future.

Keywords: Thiazoles, Antiplatelet Activity, 3D & 2D QSAR.

1 Introduction

In simplest terms Quantitative Structure–Activity Relationship (QSAR) is a method for building computational or mathematical models which attempts to find a statistically significant correlation between structure and function using a chemometric technique.¹ In terms of drug design, structure were refer to the properties or descriptors of the molecules, their substituent's or interaction energy fields function corresponds to an experimental biological/biochemical endpoint like binding affinity, activity, toxicity or rate constants; While chemometric method include MLR, PLS, PCA, PCR, ANN, GA etc. The term 'quantitative structure–property relationship' (QSPR) is used when some

property other than the biological activity is concerned.² Various QSAR approaches have been developed gradually over more than a hundred years of time span and served as a valuable predictive tools, particularly in the design of pharmaceuticals and agrochemicals.³ The methods has evolved from Hansch and Free-Wilson's one or two-dimensional linear free energy relationships via Cramer's three-dimensional QSAR to Hopfinger's fourth, Vedani's fifth and sixth dimensions.⁴ All one, two dimensional and related methods are commonly referred to as 'classical' QSAR methodologies. Every molecule included in the study binds to the same site of the same target receptor. However, the main difference between all these formalisms reside in the manner in which each one of them treats, represents structural properties of the molecules and extracts the quantitative relationships between the properties and activities.⁵ Due to the limited scope and

Corresponding author: Tel:+919822441067,
Mobile:+919405623523, Fax- +912559225042.
Email-girish.kashid@yahoo.com

space for this review, the author will focus only on the 3D-QSAR approaches in drug design.¹ The antiplatelet activity of this dataset is reported as IC₅₀ values.

When there is Vascular Endothelial damage the plates adhere to sub endothelial basement membrane and collagen. On exposure to collagen they release adenosine diphosphate from their granules together with some unstable derivative of arachidonic acid metabolism, i.e. prostaglandins PG, H₂ and Thromboxane A₂. These release substances lead to platelet aggregation, antiplatelet drug also known as platelet aggregation inhibitor.⁶

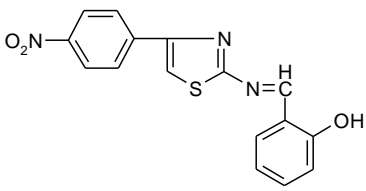
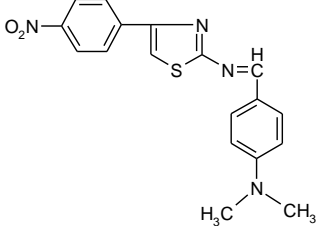
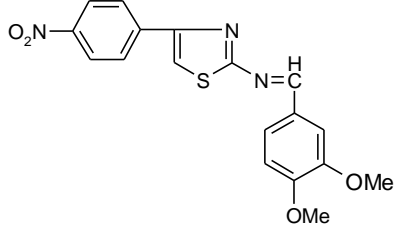
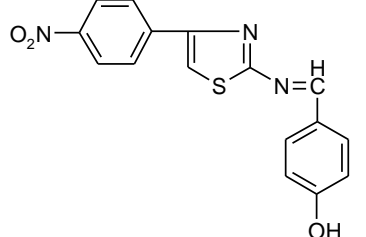
QSAR has shown that for hydrogen bond acceptors aromatic and hydrophobic are the important features for antiplatelet activity. Antiplatelet agents are becoming the area of choice for various researchers. We have taken QSAR studies of antiplatelet drug to determine the activity which in turn depends further more

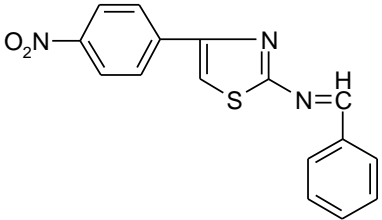
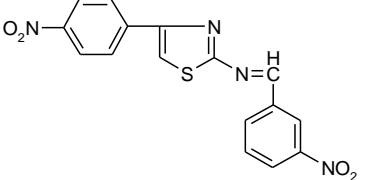
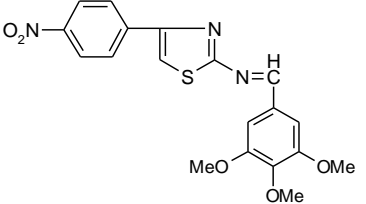
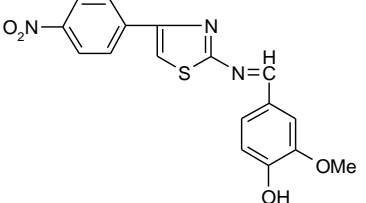
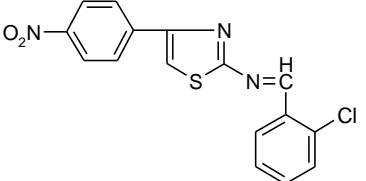
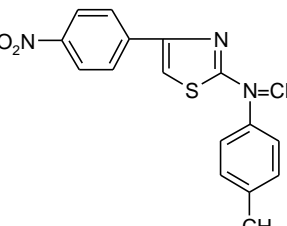
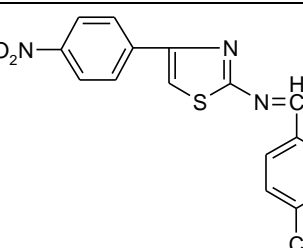
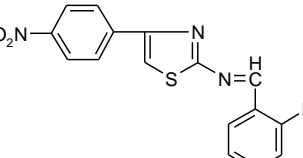
on hydrophobic, steric or electrostatic parameters.⁷

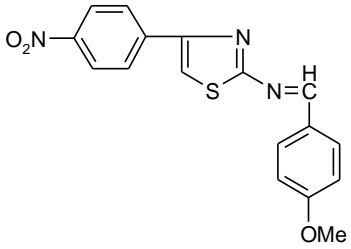
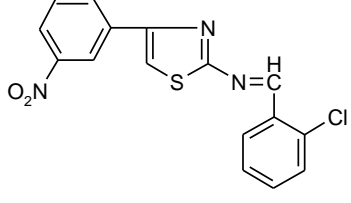
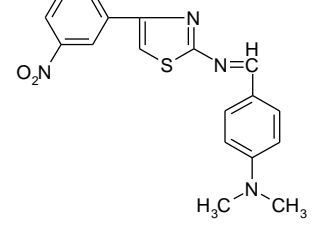
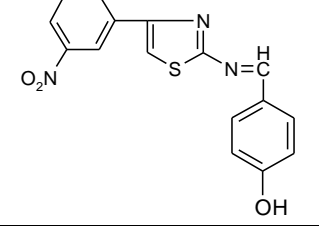
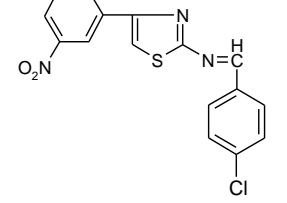
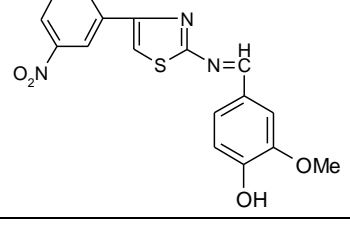
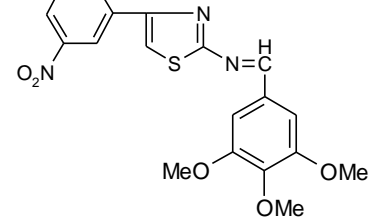
2 Materials and Methods

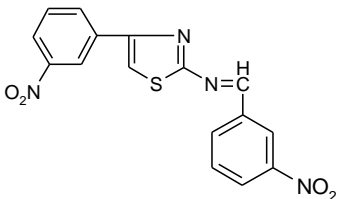
QSAR study involves data set consist of 13 structures of 2-amino-4-(4-nitrophenyl)thiazole derivatives and 7 structure of 2-amino-4-(3-nitrophenyl)thiazole derivatives.⁸ The antiplatelet activity of this dataset is reported as IC₅₀ values. The chemical structures were drawn in the 2D Draw App and converted to 3D, using V Life MDS 3.5 software (V Life sciences Pvt Ltd Pune).⁹ All structures were single point optimized using the MMFF94 force field and Gasteiger-Marsili charges, till gradient of 0.001kcal/A⁰ was reached. The optimized molecule should be aligned by template base alignment facility.¹⁰ The general structures and corresponding substitutions are included in Table 1.

Table 1 Biological Activity and structure of thiazole compounds

Sr. No.	Structure code	Structure	Log IC ₅₀	-Log IC ₅₀
1	GKS1a		1.266156	0.78978
2	GKS1b		2.241272	0.44617
3	GKS1c		1.628685	0.61399
4	GKS1d		1.254872	0.79689

5	GKS1e		2.272377	0.44006
6	GKS1f		1.74138	0.57425
7	GKS1g		1.601868	0.62427
8	GKS1h		1.340313	0.74608
9	GKS1i		1.828332	0.54694
10	GKS1j		2.278257	0.78231
11	GKS1k		1.283276	0.77925
12	GKS 1l		1.261931	0.79243

13	GKS 1m		1.655888	0.6039
14	GKS2a		1.232771	0.81119
15	GKS 2b		2.266764	0.44116
16	GKS2c		2.174814	0.45981
17	GKS2d		1.290875	0.77467
18	GKS2e		2.167744	0.46131
19	GKS2f		2.221686	0.45011

20	GKS2g		1.594701	0.62708
----	-------	---	----------	---------

2.1 Biological Activity Dataset for QSAR Analysis

Antiplatelet activity of these new thiazoles and their structures are mentioned in table no.1 which is essential for 2D & 3D QSAR studies.

2.2 Computational Details

The structures of all 20 compounds were drawn in 2D Draw App (MDS 3.5 2010). The 2D structures were converted to 3D structures by sending them to MDS. Every compound was energy minimized and batch optimized by using Merck Molecular Force Field (MMFF) and charges.¹¹

2.3 Molecular Modeling for 2D QSAR

2.3.1 Descriptor Calculation:

The Physicochemical Descriptor, Alignment Independent can be calculated by using descriptor calculation facility provided in MDS 3.5 Software. Near about several hundred of descriptors are calculated. The column containing zero value reading and invariability are removed by using 'remove invariable column tool'.¹²

2.3.2 Variable Selection:

There are a hundreds of molecular descriptors available for building a QSAR model. Not all of the molecular descriptors are important in determining the biological activity. To find the optimal subset of the descriptors a variable selection method is required, which plays an important role in determining activity.¹³ The variable selection can be done by step wise forward-backward systemic variable selection method. The IC₅₀ value is converted in to log value of IC⁵⁰; which can be used as dependent variable in QSAR analysis. Put all another descriptors as independent variable.¹⁴

2.3.3 Statistical Methods:

A suitable statistical method coupled with a variable selection method allows analyses of this data in order to establish a QSAR model, with the subset of descriptors that are most statistically significant in determining the biological activity.^{15, 16, 17}

2.3.4 Preparation of training set and test set:

The data set can be divided in to two sets i.e. training set and test set. The training set consists of 16 compounds while the test set consists of 4 compounds.^{18,19} Optimized molecules should be aligned by template base alignment facility. The general structures and corresponding substitutions are included in Table 1.

2.4 Molecular Modeling for 3D-QSAR:

Preparation of Training Set and Test Set: for 3D QSAR data set can be divided in to training set and test set.^{20, 21} the training set consists of 16 compounds while the test set consists of 4 compounds. The optimized molecule should be aligned by template base alignment facility.^{22, 23, 24,25} General structures and corresponding substitutions are included in Table 1. Descriptor calculation, variable selection and statistical methods are same as 2D QSAR of same molecule

3 Result and Discussion

For the development of QSAR models of thiazole compounds, a total number of 20 thiazole derivatives have been considered for the QSAR study using software Vlife MDS 3.5. For these 20 molecules, dataset are divided into training and test sets for an effective QSAR modeling. For selection of training and test sets, we were ensured that the molecules have uniform spread (training and test) in terms of both activity and chemical space. Biological activity was selected as dependent variable and remaining all the variables were selected as independent variables. The training and test-set molecules for this group of compounds are selected by the random selection method and the models are validated by both internal and external validation procedures. Some statistically significant QSAR models were chosen for discussion.

The derived models in 2D QSAR from multiple linear regression (MLR) with forward stepwise shows good correlation between

biological activity and parameters Quadrupole2, MomInertiaX, ZcompDipole, QMDipoleY as the coefficient of determination, $r^2 = 0.9660$, $r^2 = 0.758$, capable of explaining 72% of variance in the observed activity values. All the descriptors contributed well for the generation of model. The low standard error of r^2 $se = 0.0324$, r^2 $se = 0.0795$ demonstrates accuracy of the model. The leave-one-out procedure was used for internal validation of the model. The model showed an internal predictive power cross validated r^2 ($q^2 = 0.9262$, $q^2 = 0.6168$) of 65% values reflect good internal predictive power of

the model. In addition, the randomization test shows confidence of 99 % that the generated model is not random and hence it is chosen as the QSAR model. The F-test= 78.1905, 20.3746 shows the overall statistical significance level of 99 % of the model which means the probability of failure of the model is 1 in 10,000. The descriptors show positive correlation among the parameters selected for the derived QSAR model. The positive coefficients suggest that inclusion of such carbon atoms in the molecules lead to increased antiplatelet activity.

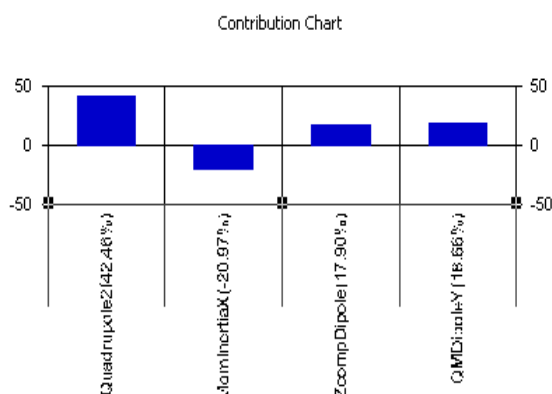
Table No. 2 Evaluation of Molecular Modeling for 2D & 3D-QSAR

QSAR Methods	2D-QSAR (Parameters)				
	Sets	Selected Descriptors	Coefficient	Constant	Statistics
Multiple regressions Forward Method	Training Set Size = 16, Test Set Size = 4	Quadrupole2, MomInertiaX, ZcompDipole, QMDipoleY	0.0082(± 0.0001), -0.0001(± 0.0000), 0.0377(± 0.0086), 0.0952(± 0.0212)	0.8550	n = 16, Degree of freedom = 11, $r^2 = 0.9660$, $q^2 = 0.9262$, F test = 78.1905, r^2 se = 0.0324, q^2 se = 0.0478 pred_ r^2 = -3.5849, pred_ r^2 se = 0.3011.
Principle Component Regression forward Method:	Training Set Size = 16, Test Set Size = 4	Quadrupole2, ZcompDipole	0.0057,-0.0110, 0.0253	1.6419	Optimum Components = 2, n = 16 Degree of freedom = 13, $r^2 = 0.7581$, $q^2 = 0.6168$ F test = 20.3746, r^2 se = 0.0795, q^2 se = 0.1001, pred_ r^2 = -1.1447, pred_ r^2 se = 0.2059 (Fig 3 & 4)
QSAR Methods	3D-QSAR (Parameters)				
	Sets	Selected Descriptors	Coefficient	Constant	Statistics
Multiple regressions Forward Method	Training Set Size = 16, Test Set Size = 4	E_944, S_505, S_911, E_716	0.1838(± 0.0004), 0.1580(± 0.0008), 4.3740(± 0.6071), 0.0357(± 0.0005)	0.5412	n = 16, Degree of freedom = 11, $r^2 = 0.9660$, $q^2 = 0.9262$, F test = 78.1905, r^2 se = 0.0324, q^2 se = 0.0478 pred_ r^2 = -3.5849, pred_ r^2 se = 0.3011.

Model-I	Equation-1: LOGMIC=0.1838(±0.0004)E_944+0.1580(±0.0008)S_505+4.3740(±0.6071)S_911+0.0357(±0.0005) E_716+ 0.541 (Fig-5,6 &7)				
Principle Component Regression forward Method:	Training Set Size = 16, Test Set Size = 4	E_944, S_749, S_1124, E_529	-0.1071, 3.5072, -0.1799, 0.1171	0.7638	Optimum Components = 3, n = 16, Degree of freedom = 12, $r^2 = 0.9526$, $q^2 = 0.9083$, F test = 80.4572, $r^2_{se} = 0.0366$, $q^2_{se} = 0.0509$, $pred_r^2 = -2.7414$, $pred_r^2_{se} = 0.2720$
Model-II	Equation-2: LOGMIC=0.1071E_944+3.5072S_749-0.1799 S_1124+0.1171E_529+0.7638(Eqn2) (fig-8, 9, 10)				
Partial Least Square Regression forward Method	Training Set Size = 16, Test Set Size = 4	E_944, S_749, S_1124, E_539	:-0.0972, 3.6200, -0.2435, 0.2332	0.7655	Optimum Components = 3 n = 16 Degree of freedom = 12, $r^2 = 0.9665$, $q^2 = 0.9278$, F test = 115.4957, $r^2_{se} = 0.0308$, $q^2_{se} = 0.0452$ $pred_r^2 = -3.2191$, $pred_r^2_{se} = 0.2889$
Model-III	Equation-3: LOGMIC=0.0972E_944+3.6200S_749-0.2435S_1124+0.2332E_539+0.7655 (Eq ⁿ 3) (Fig-11, 12 & 13)				

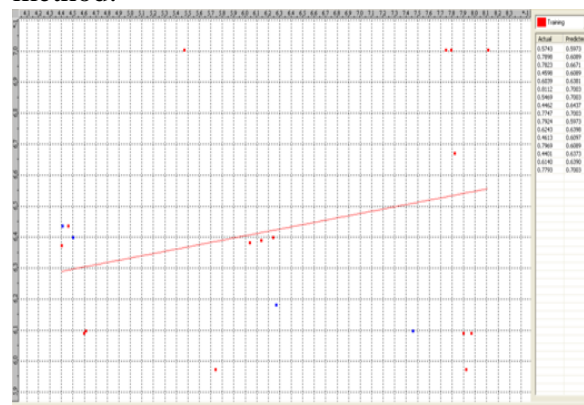
Quadrupole2 descriptor signifies magnitude of first tensor of quadrupole moments. Its positive contribution in the QSAR model implies that will lead to increase potency. Its positive value suggests that increasing the number of such atom that increase the dipole moment will lead

Fig-1 Contribution chart for multiple regression method:



to better antiplatelet potency. The MomInertiaX, ZcompDipole, QMDipoleY descriptor are type of dipole interaction and its contribution for the antiplatelet activities indicate that optimum groups provide good antiplatelet activity.

Fig-2 Fitness plot for multiple regression method:



The derived models in 3D QSAR from multiple linear regression (MLR) with forward stepwise shows good correlation between biological activity and parameters. With coefficient of determination $r^2=0.9660$, $r^2=0.9526$, $r^2=0.9665$, which is capable of explaining variance in the observed activity values. The model selection criterion is the value 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 randomization tests suggest that the proposed QSAR model has a probability of less than 0.01 of being generated by chance. E_944, S_505, S_911, E_716, S_749, S_1124, E_529, E_944, E_539 are steric descriptors and electrostatic descriptors contributing to models. The q^2 value obtained (0.9262, 0.9083, 0.9278, 0.6111) are the indicative power of the models. Values of r^2 , q^2 , F test, r^2 se, q^2 se, $pred_r^2$, $pred_r^2$ se prove that QSAR equation are obtained is statistically significant and shows that the predictive power of the model is 70%

Fig-3 Contribution chart for partial component regression forward:

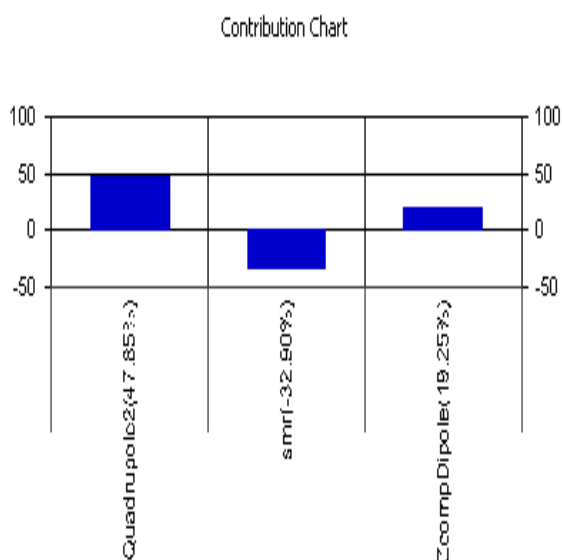


Fig-4 Fitness plot for partial component regression forward:

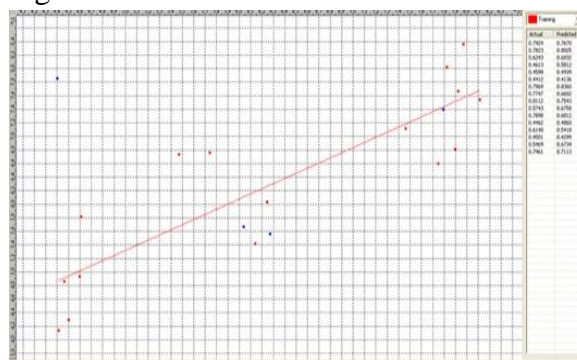


Fig-5 Show molecule w.r.t. model I

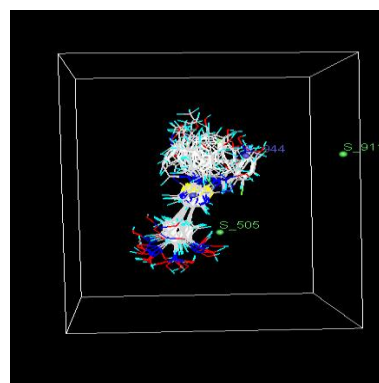


Fig-6 Contribution chart w.r.t.model I

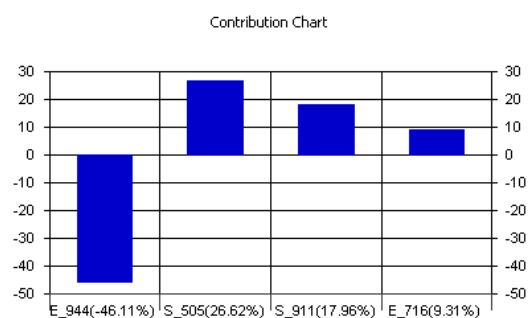


Fig-7 Fitness plot w.r.t. model I

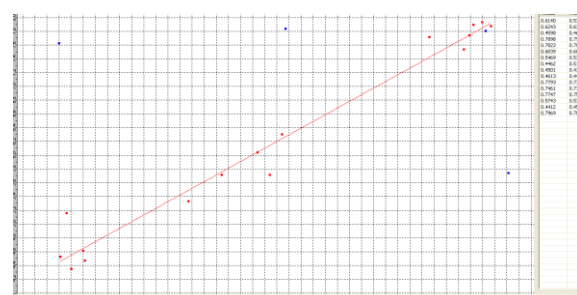
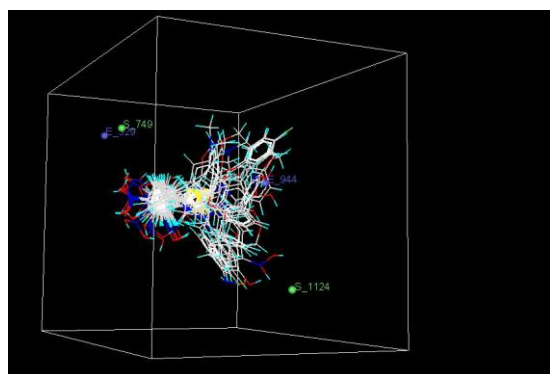


Fig-8 Show molecule w.r.t. model II



(internal validation) and 65 % (external validation). Steric descriptors indicate that steric potential is favorable for activity and less bulky substituent is preferred in that region. Steric and electrostatic field energy of interactions between probe (CH₃) and compounds at their corresponding spatial grid points show in 3D view (fig no.5,8,11 &14). The contributions of steric and electrostatic fields indicate that both fields are more important.

The steric effect, as shown in fig no. 5,8,11 &14 with green color ball around the phenyl ring at ortho or meta position, implies about the preferred substitution (less or more bulky group)

Fig-9 Contribution chart w.r.t. model II

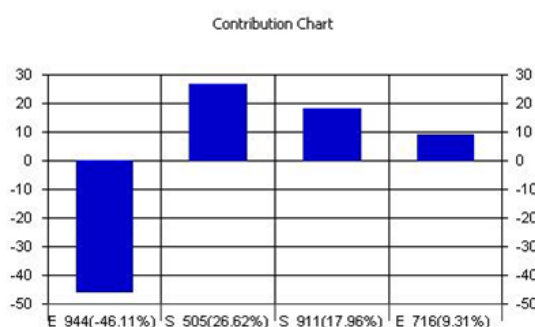


Fig- 10 Fitness plot w.r.t.model II

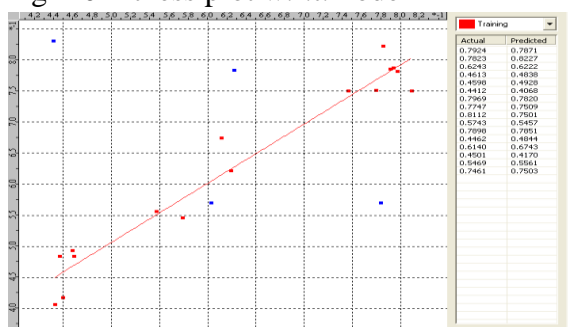
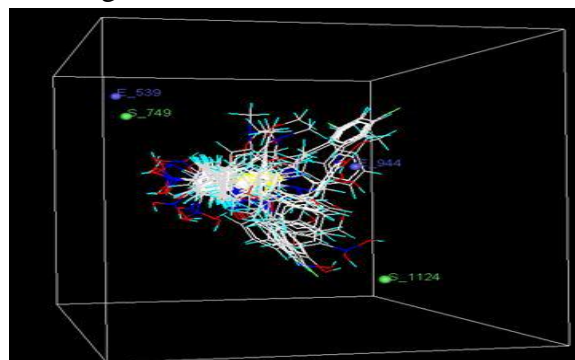


Fig-11 Show molecule w.r.t. model III



to produce higher antiplatelet activity. Electrostatic descriptor with positive coefficient around position of the phenyl ring corroborates that electropositive (electron-withdrawing) group is preferred at 4-position of phenyl ring.

4 Conclusions

Equation 1 explains ~97 % ($r^2 = 0.9665$) of the total variance in the training set as well as it has internal (q^2) and external ($pred_r^2_{se}$) predictive ability of ~92 % and ~28% respectively. The equation 2 explains ~95 % ($r^2 = 0.9083$) of the total variance in the training set as well as it has internal (q^2) and external ($pred_r^2_{se}$)

Fig- 12 contribution chart w.r.t.model III

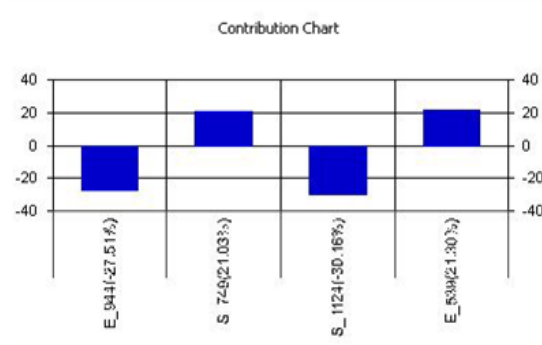


Fig- 13 Fitness plot w.r.t.model III

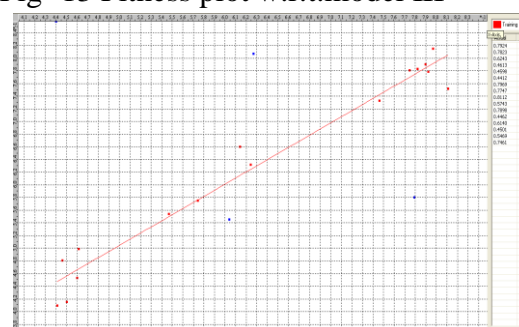


Fig-14 show molecule for kNN Method
Forward:

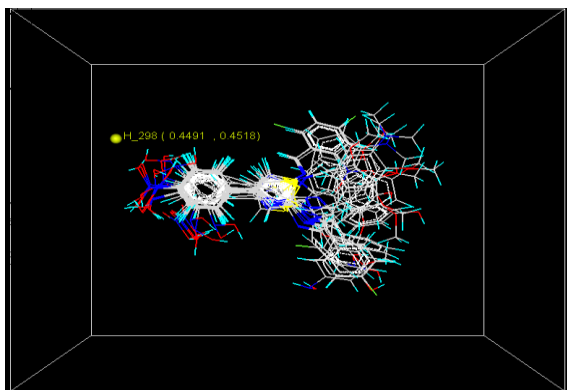
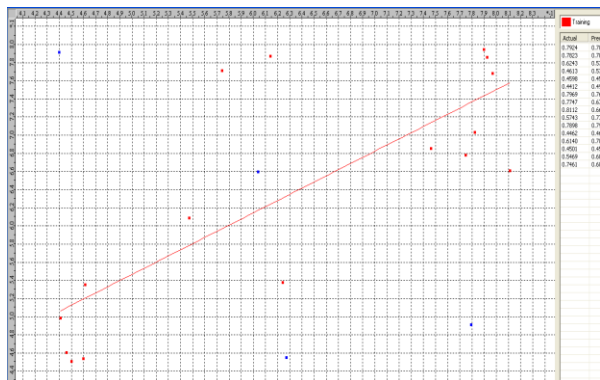


Fig-15 Fitness plot for kNN Method
Forward:



predictive ability of ~90 % and ~27% respectively. Equation 3 explains ~96% ($r^2 = 0.9665$) of the total variance in the training set as well as it has internal (q^2) and external (pred_r2se) predictive ability of ~92 % and ~28% respectively. From kNN Forward method, the descriptor range is, H_{298} 0.4491 to 0.4518, that means Positive range of hydrophobic descriptor indicates that positive hydrophobic potential is favorable for increase in the antiplatelet activity, hence a less bulky substituent group is Preferred in that region. 2D & 3D-QSAR models with moderate to high predictive ability of thiazole derivatives were derived. The role of hydrophobicity as a 3D property was confirmed and also Electrostatic and Steric effects were found to contribute to antiplatelet activity. The obtained models may help design of new active thiazole as antiplatelet activity.

DECLARATION OF CONFLICT INTEREST:

Author declares no conflict of interest.

5 Acknowledgements

The authors gratefully acknowledge the "SND College of Pharmacy" for providing VLife MDS software and more on. Authors like to acknowledge Principal of the institute for providing facilities to carry out the work. The authors also thank the anonymous reviewers whose valuable comments and suggestions greatly helped in improving the manuscript.

REFERENCES

1. Ajmani S, Jadhav K and Kulkarni SA. Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. *Journal of Chemical Information & Modeling* 2006;46:24–31.
2. Bhadoriya KS, Jain SV, Bari SB, Chavhan ML and Vispute KR. 3D-QSAR study of indol-2-yl ethanones derivatives as novel indoleamine 2, 3-dioxygenase (IDO) inhibitors. *European Journal of Chemistry* 2012;9:1753–59.
3. Ivanka M. Tsakovska QSAR and 3D-QSAR of Phenothiazine Type Multidrug Resistance Modulators in P388/ADR Cells. *Bioorganic & Medicinal Chemistry* 2003;11:2889–99.
4. Clark M, Cramer III RD and Van ON, Validation of the general purpose Tripose 5.2 force field. *Journal of Computational Chemistry* 1989;10:982–1012.
5. Cramer III RD, Patterson DE and Bunce JD. Comparative molecular field analysis (CoMFA) 1. Effect of shape on binding of steroids to carrier proteins. *Journal of American Chemical Society* 1988;110:5959–67.
6. Akira M, Masakazu S, Mari A, Makoto T. 2-Acylimino-3H-thiazoline Derivatives: A Novel Template for Platelet GPIIb/IIIa Receptor Antagonists. *Bioorganic & Medicinal Chemistry Letters* 2001;11:1031–35.
7. Doweiko AM. QSAR: dead or alive? *J Comput Aided Mol Des* 2008;22:81–89.
8. Yadav JS, Subba Reddy BV, Gopala Rao Y, Narsaiah AV. First example of the coupling of α -diazoketones with thiourea: a novel route for the synthesis of 2-aminothiazoles. *Tetrahedron Letters* 2008;49:2381–83.

9. Gaurav A, Yadav MR, Giridhar R, Gautam V and Singh R. 3D-QSAR studies of 4-quinolone derivatives as high-affinity ligands at the benzodiazepine site of brain GABAA receptors. *Medicinal Chemistry Res* 2010; Available from: URL: <http://dx.doi.org/10.1007/s00044-010-9306-5>.
10. Ghosh P and Bagchi MC. QSAR modeling for quinoxaline derivatives using genetic algorithm and simulated annealing based feature selection. *Current Medicinal Chemistry* 2009;16:4032–48.
11. Halgren TA. Merck molecular force field: III. Molecular geometries and vibrational frequencies for MMFF94. *Journal of Computational Chemistry* 1996;17:553–86
12. Hansch C. A quantitative approach to biochemical structure activity relationships. *Account of Chemical Research* 1969;2:232–39.
13. Jain SV, Bhadoriya KS and Bari SB. QSAR and flexible docking studies of some aldose reductase inhibitors obtained from natural origin. *Medicinal Chemistry Research* 2012;21:1665–76. Available from: URL: <http://dx.doi.org/10.1007/s00044-011-9681-6>.
14. Jain SV, Ghate M, Bhadoriya KS, Bari SB, Chaudhari A and Borse JS. 2D, 3D-QSAR and docking studies of 1,2,3-thiadiazole thioacetanilides analogues as potent HIV-1 non-nucleoside reverse transcriptase inhibitors. *Organic Medicinal Chemistry Letters* 2012;2:22. Available from: URL: <http://dx.doi.org/10.1186/2191-2858-2-22>.
15. Nakka S and Guruprasad L. The imidazolidone analogs as phospholipase D1 inhibitors: analysis of the three-dimensional quantitative structure–activity relationship. *Medicinal Chemistry Research* 2012;21:2517–25. Available from: URL: <http://dx.doi.org/10.1007/s00044-011-9773-3>.
16. Radhika V, Kanth SS and Vijjulatha M. CoMFA and CoMSIA studies on inhibitors of HIV-1 integrase-bicyclic pyrimidinones. *Journal of Chemistry*. 2010; 7(S1):S75–S84.
17. Sahu NK, Bari SB and Kohli DV. Molecular modeling studies of some substituted chalcone derivatives as cysteine protease inhibitors. *Medicinal Chemistry Research* 2011; Available from: URL: <http://dx.doi.org/10.1007/s00044-011-9900-1>.
18. Scior T, Medina-Franco JLDQ, Martinez-Mayorga K, Yunes Rojas JA, Bernard P. How to recognize and workaround pitfalls in QSAR studies: a critical review. *Current Medicinal Chemistry* 2009;16:4297–13.
19. Sharma MC and Kohli DV. A comprehensive structure–activity analysis 2,3,5-trisubstituted 4,5-dihydro-4-oxo-3H-imidazo [4,5-c] pyridine derivatives as angiotensin II receptor antagonists: using 2D- and 3D-QSAR approach. *Medicinal Chemistry Research* 2012 Available from: URL: <http://dx.doi.org/10.1007/s00044-012-0040-z>.
20. Shen M, Xiao Y, Golbraikh A, Gombar VK and Tropsha A. An in silico screen for human S9 metabolic turnover using KN nearest neighbor QSPR method. *Journal of Medicinal Chemistry* 2003;46:3013–20.
21. Tanneeru K, Reddy BM and Guruprasad L. Three-dimensional quantitative structure–activity relationship (3D-QSAR) analysis and molecular docking of ATP-competitive triazine analogs of human mTOR inhibitors. *Med. Chem. Res.* 2011; Available from: URL: <http://dx.doi.org/10.1007/s00044-011-9629-x>.
22. VLife MDS 3.5, Molecular Design Suite. Vlife Sciences Technologies Pvt. Ltd. Pune. 2010.
23. Zheng W and Tropsha A. Novel variable selection quantitative structure property relationship approach based on the k-nearest neighbor principle. *Journal of Chemistry Information and Computer Science* 2000;40:185–94.

24. Mukesh CS, Kohli DV. Insight into the structural requirement of aryltriazolinone derivatives as angiotensin II AT1 receptor: 2D and 3D-QSAR k-Nearest Neighbor Molecular Field Analysis approach. *Medicinal Chemistry Research* 2012;21(10):2837-53.
25. Verma J, Khedkar VM. 3D-QSAR in drug design--a review. *Current topics in medicinal chemistry* 2009;10(1):95-115.