Asian Journal of Chemistry Vol. 22, No. 10 (2010), 8231-8245

Quantitative Structure-Activity Relationship Study on Pyrrolotriazine Derivatives as Met Kinase Inhibitors

B.K. SHARMA*, YASHWANT† and B. SRIVASTAVA†
Department of Chemistry, S.K. Government College, Sikar-332 001, India
Fax: (91)(1572)270039; Tel: (91)(1572)251039
E-mail: bksharma_sikar@rediffmail.com; yash99yash@gmail.com

Met kinase inhibitory activity of pyrrolotriazine derivatives has been quantitatively analyzed in terms of Dragon descriptors using combinatorial protocol in multiple linear regression (CP-MLR). The study has provided a rational approach for the development of new pyrrolotriazine derivatives as Met kinase inhibitors. The descriptors identified in CP-MLR analysis have highlighted the role of atomic Sanderson's electronegativity in modified Burden eigen value (BELe4) and in respective lag of 2D-autocorrelation (GATS7e), 9th order self-returning walk count (SRW09), structural information content of 4th order neighbourhood symmetry (SIC4) and aromatic ratio (ARR) to explain the Met kinase inhibitory activity of pyrrolotriazine derivatives. Certain structural fragments (C-005 and N-069) and number of five membered rings (nR05) in molecular structures have also shown prevalence to optimize the Met kinase inhibitory activity of pyrrolotriazine derivatives. These guidelines may be used to develop new Met kinase inhibitors based on pyrrolotriazine scaffold.

Key Words: QSAR, Pyrrolotriazine derivatives, Met kinase inhibitory activity, Combinatorial protocol in multiple linear regression.

INTRODUCTION

The Met receptor (a glycosylated, heterodimeric receptor tyrosine kinase) is expressed predominately in the epithelium and endothelium[1]. It is also known as hepatocyte growth factor receptor (HGF). The Met kinase was first identified and derived its name as an activated oncogene in N-methyl-N'-nitro-N-nitrosoguanidine-treated human osteosarcoma cells[2]. A disulfide-linked hetero dimer which can activate Met *via* autocrine or paracrine mechanics is the ligand for the Met receptor. It is observed that HGF-Met signaling triggers a variety of cellular responses such as cell growth, migration and invasion tumor metastasis, angiogenesis, wound healing and tissue regeneration. The deregulated Met kinase signaling has shown implications in a variety of cancers which include prostate[3] (ligand-dependent activation), gastric[4] (gene amplification/overexpression) and hereditary papillary renal cell carcinoma[5] (activating mutation). The pyrrolotriazine derivatives have been reported as novel

†School of Pharmaceutical Sciences, Jaipur National University, Jagatpura, Jaipur-302 025, India.

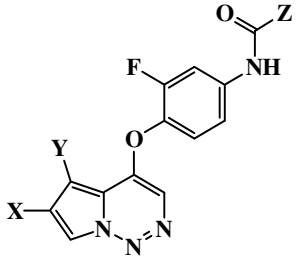
Met kinase inhibitors, in a recent study[6]. This scaffold was identified in VEGFR-2[7] and EGFR/pan-Her[8] programs. In view of the importance of Met kinase inhibitors in the variety of cancers, a quantitative structure-activity relationship is attempted on the Met kinase inhibitory activity of these pyrrolotriazine derivatives. The present study is aimed at rationalizing the substituent variations of these analogues to provide insight for the future endeavours.

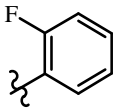
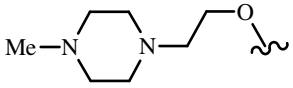
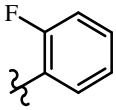
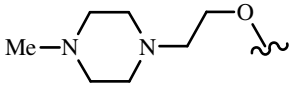
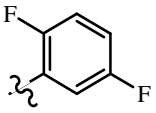
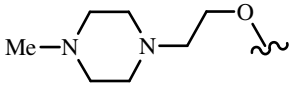
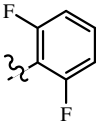
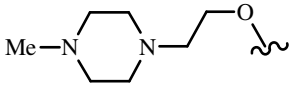
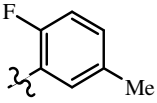
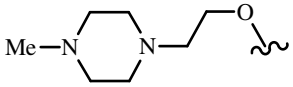
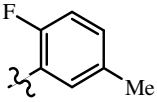
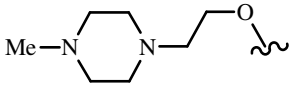
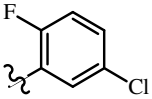
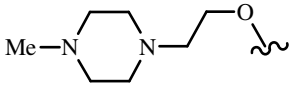
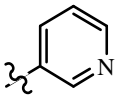
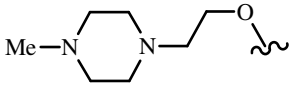
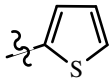
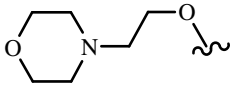
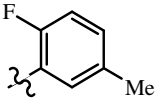
EXPERIMENTAL

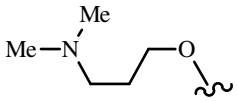
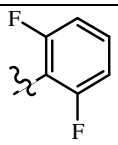
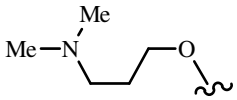
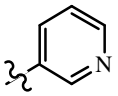
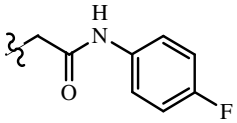
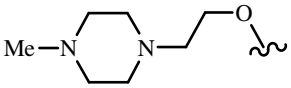
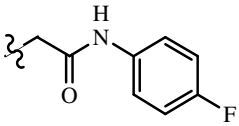
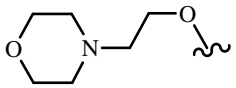
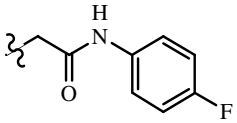
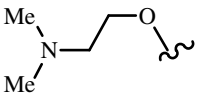
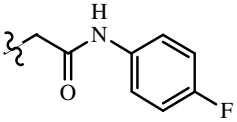
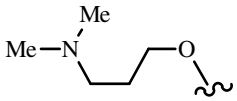
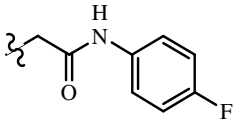
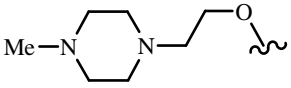
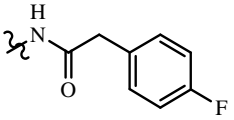
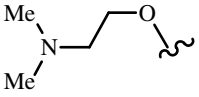
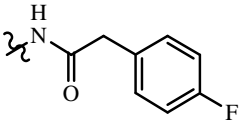
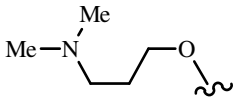
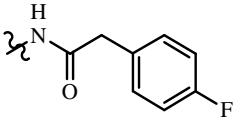
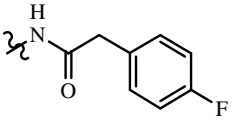
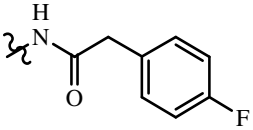
Data set: Reported 36 pyrrolotriazine derivatives are considered as data set for the present study[6]. The structural variations of these analogues are given in Table-1. The biological actions of these compounds were reported in terms of inhibitory activity, IC_{50}, the concentration required to inhibit 50 % of the kinase activity. These inhibitory concentrations were converted to molar scale (pIC_{50}) and are used in QSAR analysis. For the purpose of modeling study all 36 analogues have been divided into training and test sets. Out of the 36 analogues, one fourth compounds (**9**) have been placed in the test set for the validation of derived models. The training and test set compounds are also listed in Table-1.

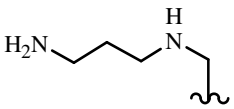
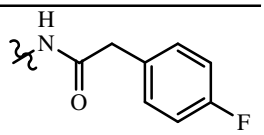
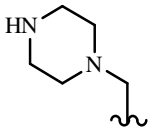
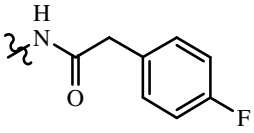
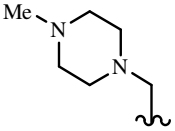
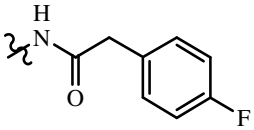
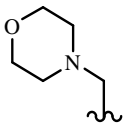
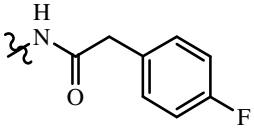
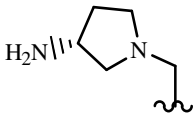
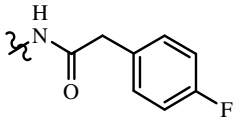
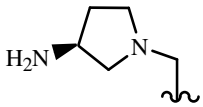
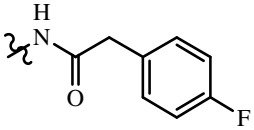
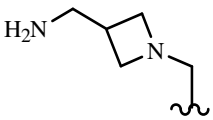
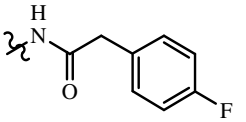
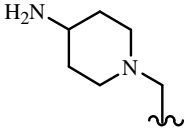
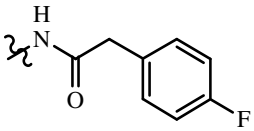
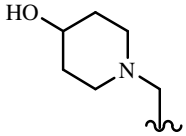
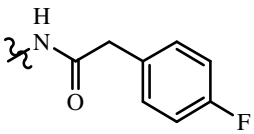
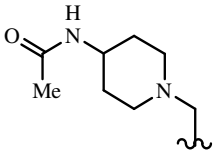
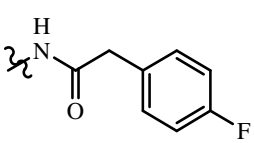
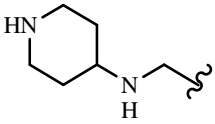
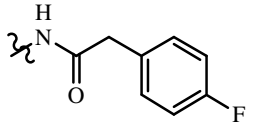
Theoretical molecular descriptors: The structures of the compounds under study have been drawn in 2D ChemDraw[9]. The drawn structures were then converted into 3D modules using the default conversion procedure implemented in the CS Chem3D Ultra. The energy of these 3D-structures was minimized in the MOPAC module using the AM1 procedure for closed shell systems. This will ensure a well defined conformer relationship among the compounds of the study. All these energy minimized structures of respective compounds have been ported to DRAGON software[10] for the computation of descriptors for the pyrrolotriazine derivatives (Table-1). This software offers several hundreds of descriptors from different perspectives corresponding to 0D-, 1D- and 2D-descriptor modules. The outlined modules comprised of 10 different classes, namely, the constitutional (CONST), the topological (TOPO), the molecular walk counts (MWC), the BCUT descriptors (BCUT), the Galvez topological charge indices (GALVEZ), the 2D autocorrelations (2D-AUTO), the functional groups (FUNC), the atom-centered fragments (ACF), the empirical descriptors (EMP) and the properties describing descriptors (PROP). For each of these classes the DRAGON software computes a large number of descriptors which are characteristic to the molecules under multi-descriptor environment. The definition and scope of these descriptor's classes is given in Table-2. The combinatorial protocol in multiple linear regression (CP-MLR)[11] procedure has been used in the present work for developing QSAR models. Before the application of CP-MLR procedure, all those descriptors which are intercorrelated beyond 0.90 and showing a correlation of less than 0.1 with the biological end points (descriptor *vs.* activity, $r < 0.1$) were excluded. This has reduced the total dataset of the compounds from 477 to 106 descriptors as relevant ones for the Met kinase inhibitory activity. A brief description of the computational procedure is given below.

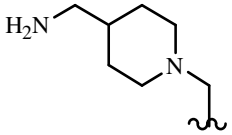
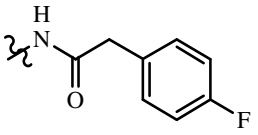
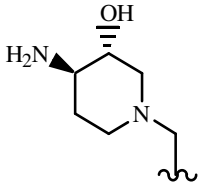
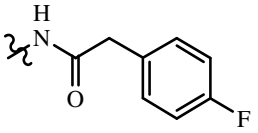
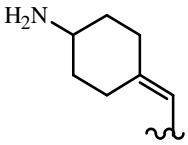
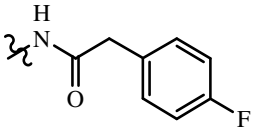
TABLE-1
STRUCTURES OF THE PYRROLOTRIAZINE DERIVATIVES[a]



Compd.	X	Y	Z
1	H	Me	
2		Me	
3		Me	
4[b]		Me	
5[b]		Me	
6		Me	
7		Me	
8		Me	
9		Me	
10		Me	

11		Me	
12[b]		Me	
13	H	CH_3	
14[b]		CH_3	
15		CH_3	
16		CH_3	
17		CH_3	
18[b]		CH_3	
19		CH_3	
20		CH_3	
21	H	Cl	
22	H	Ph	

23[b]	H		
24	H		
25[b]	H		
26	H		
27	H		
28	H		
29[b]	H		
30	H		
31	H		
32	H		
33[b]	H		

34	H		
35	H		
36	H		

[a]As in reference [6]; [b]Compounds included in test set.

TABLE-2
DESCRIPTOR CLASSES [Ref. 10] USED, ALONG WITH THEIR DEFINITION AND SCOPE FOR MODELING THE MET KINASE INHIBITORY ACTIVITY OF PYRROLOTRIAZINE DERIVATIVES

Descriptor class (acronyms)	Definition and scope
Constitutional (CONST)	Dimensionless or 0D descriptors; independent from molecular connectivity and conformations
Topological (TOPO)	2D-descriptor from molecular graphs and independent conformations
Molecular walk counts (MWC)	2D-descriptors representing self-returning walks counts of different lengths
Modified Burden eigen values (BCUT)	2D-descriptors representing positive and negative eigen values of the adjacency matrix, weights the diagonal elements and atoms
Galvez topological charge indices (GALVEZ)	2D-descriptors representing the first 10 eigen values of corrected adjacency matrix
2D-autocorrelations (2D-AUTO)	Molecular descriptors calculated from the molecular graphs by summing the products of atom weights of the terminal atoms of all the paths of the considered path length (the lag)
Functional groups (FUNC)	Molecular descriptors based on the counting of the chemical functional groups
Atom centered fragments (ACF)	Molecular descriptors based on the counting of 120 atom centered fragments, as defined by Ghose-Crippen
Empirical (EMP)	1D-descriptors represent the counts of non-single bonds, hydrophilic groups and ratio of the number of aromatic bonds and total bonds in an H-depleted molecule
Properties (PROP)	1D-descriptors representing molecular properties of a molecule

Model development: The CP-MLR is a 'filter' based variable selection procedure for model development in QSAR studies[11]. Its procedural aspects and implementation are discussed in some of our recent publications[12-16]. It involves selected subset regressions. In this procedure a combinatorial strategy with appropriately placed 'filters' has been interfaced with MLR to result in the extraction of diverse structure-activity models, each having unique combination of descriptors from the dataset under study. In this, the contents and number of variables to be evaluated are mixed according to the predefined confines. Here the 'filters' are significant evaluators of the variables in regression at different stages of model development. Of these, filter-1 is set in terms of inter-parameter correlation cutoff criteria for variables to stay as a subset (filter-1, default value 0.3 and upper limit ≤ 0.79). In this, if two variables are correlated higher than a predefined cutoff value the respective variable combination is forbidden and will be rejected. The second filter is in terms of t-values of regression coefficients of variables associated with a subset (filter-2, default value 2.0). Here, if the ratio of regression coefficient and associated standard error of any variable is less than a predefined cutoff value then the variable combination will be rejected. Since successive additions of variables to multiple regression equation will increase successive multiple correlation coefficient (r) values, square-root of adjusted multiple correlation coefficient of regression equation, r-bar, has been used to compare the internal explanatory power of models with different number of variables. Accordingly, a filter has been set in terms of predefined threshold level of r-bar (filter-3, default value 0.71) to decide the variables' 'merit' in the model formation. Finally, to exclude false or artificial correlations, the external consistency of the variables of the model have been addressed in terms of cross-validated R^2 or Q^2 criteria from the leave-one-out (LOO) cross-validation procedure as default option (filter-4, default threshold value $0.3 \leq Q2 \leq 1.0$). All these filters make the variable selection process efficient and lead to unique solution. In order to collect the descriptors with higher information content and explanatory power, the threshold of filter-3 was successively incremented with increasing number of descriptors (per equation) by considering the r-bar value of the preceding optimum model as the new threshold for next generation.

Model validation: In this study, the data set is divided into training set for model development and test set for external prediction. Goodness of fit of the models was assessed by examining the multiple correlation coefficient (r), the standard deviation (s), the F-ratio between the variances of calculated and observed activities (F). A number of additional statistical parameters such as the Akaike's information criterion, AIC[17,18], the Kubinyi function, FIT[19,20] and the Friedman's lack of fit, LOF[21], (eqs. 1-3) have also been derived to evaluate the best model.

$$AIC = \frac{RSS \times (n + p')}{(n - p')} \quad (1)$$

$$FIT = \frac{r^2 \times (n - k - 1)}{(n + k^2) \times (1 - r^2)} \quad (2)$$

$$LOF = \frac{RSS / n}{\left(1 - \frac{k(d+1)}{n}\right)^2} \quad (3)$$

where, RSS is the sum of the squared differences between the observed and the estimated activity values, k is the number of variables in the model, p' is the number of adjustable parameters in the model and d is the smoothing parameter. The AIC takes into account the statistical goodness of fit and the number of parameters that have to be estimated to achieve that degree of fit. The FIT, closely related to the F-value (Fisher ratio), was proved to be a useful parameter for assessing the quality of the models. The main disadvantage of the F-value is its sensitivity to changes in k (the number of variables in the equation, which describe the model), if k is small and its lower sensitivity if k is large. The FIT criterion has a low sensitivity toward changes in k-values, as long as they are small numbers and a substantially increasing sensitivity for large k-values. The model that produces the minimum value of AIC and the highest value of FIT is considered potentially the most useful and the best. The LOF takes into account the number of terms used in the equation and is not biased, as are other indicators, toward large numbers of parameters. A minimum LOF value infers that the derived model is statistically sound.

The internal validation of derived model was ascertained through the cross-validated index, Q^2, from leave-one-out and leave-five-out procedures. The LOO method creates a number of modified data sets by taking away one compound from the parent data set in such a way that each observation has been removed once only. Then one model is developed for each reduced data set and the response values of the deleted observations are predicted from these models. The squared differences between predicted and actual values are added to give the predictive residual sum of squares, PRESS. In this way, PRESS will contain one contribution from each observation. The cross-validated Q^2_{LOO} value may further be calculated as

$$Q^2_{LOO} = 1 - \frac{PRESS}{SSY} \quad (4)$$

where, SSY represents the variance of the observed activities of molecules around the mean value. In leave-five-out procedure, a group of 5 compounds is randomly kept outside the analysis each time in such a way that all the compounds, for once, become the part of the predictive groups. A value greater than 0.5 of Q^2-index hints toward a reasonable robust model.

The external validation or predictive power of derived model is based on test set compounds. The squared correlation coefficient between the observed and predicted values of compounds from test set, r^2_{Test}, has been calculated as:

$$r^2_{Test} = 1 - \frac{\sum (Y_{Pred(Test)} - Y_{(Test)})^2}{\sum (Y_{(Test)} - \bar{Y}_{(Training)})^2} \quad (5)$$

where, $Y_{Pred(Test)}$ and $Y_{(Test)}$ indicate predicted and observed activity values, respectively of the test-set compounds and $\overline{Y}_{(Training)}$ indicate mean activity value of the training set. r^2_{Test} is the squared correlation coefficient between the observed and predicted data of the test-set. A value greater than 0.5 of r^2_{Test} suggests that the model obtained from training set has a reliable predictive power.

Y-randomization: Chance correlations, if any, associated with the CP-MLR models were recognized in randomization test[22,23] by repeated scrambling of the biological response. The data sets with scrambled response vector have been reassessed by multiple regression analysis (MRA). The resulting regression equations, if any, with correlation coefficients better than or equal to the one corresponding to the unscrambled response data were counted. Every model has been subjected to 100 such simulation runs. This has been used as a measure to express the per cent chance correlation of the model under scrutiny.

RESULTS AND DISCUSSION

In multi-descriptor class environment, exploring for best model equation(s) along the descriptor class provides an opportunity to unravel the phenomenon under investigation. In other words, the concepts embedded in the descriptor classes relate the biological actions revealed by the compounds. For the purpose of modeling study, 9 compounds have been included in the test set for the validation of the models derived from 27 training set compounds. A total number of 106 significant descriptors from 0D-, 1D- and 2D-classes have been subjected to CP-MLR analysis with default 'filters' set in it. Statistical models in three and four descriptor(s) have been derived successively to achieve the best relationship correlating Met kinase inhibitory activity. These models (with 106 descriptors) were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this the optimum r-bar value of the preceding level model has been used as the new threshold of filter-3 for the next generation. The highest significant models in three and four descriptors are given below:

$$pIC_{50} = 3.752(1.085)BELe4 + 0.485(0.108)C\text{-}005 + 1.025(0.216)N\text{-}069 + 0.422$$
$$n = 27,\ r = 0.820,\ s = 0.407,\ F = 15.768,\ AIC = 0.223,\ FIT = 1.314,$$
$$LOF = 0.234,\ Q^2_{LOO} = 0.496,\ Q_{2L5O} = 0.376,\ r^2_{randY}(sd) = 0.312(0.127),$$
$$r^2_{Test} = 0.551 \quad (6)$$

$$pIC_{50} = -24.974(4.475)SIC4 + 3.291(1.155)BELe4 +$$
$$1.281(0.205)N\text{-}069 + 6.501(1.685)ARR + 21.686$$
$$n = 27,\ r = 0.872,\ s = 0.356,\ F = 17.483,\ AIC = 0.184,\ FIT = 1.626,$$
$$LOF = 0.209,\ Q^2_{LOO} = 0.561,\ Q^2_{L5O} = 0.598,\ r^2_{randY}(sd) = 0.369(0.128),$$
$$r^2_{Test} = 0.576 \quad (7)$$

In above and all follow up regression equations, the values given in the parentheses are the standard errors of the regression coefficients. The $r^2_{randY}(sd)$ is the mean random squared multiple correlation coefficient of the regressions in the activity

(Y) randomization study with its standard deviation from 100 simulations. In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation. The signs of the regression coefficients suggest the direction of influence of explanatory variables in the models.

The descriptor BELe4 belongs to BCUT class of Dragon descriptors. The BCUT descriptors are the first 8 highest and the lowest absolute eigen values, BEHwk and BELwk, respectively, for the modified Burden adjacency matrix. Here w refers to the atomic property and k to the eigen value rank. The ordered sequence of the highest and the lowest eigen values reflect upon the relevant aspects of molecular structure, useful for similarity searching. The positive contribution of descriptor BELe4 to the activity advocates that a higher value of this descriptor is beneficiary to the activity. The descriptors C-005 and N-069 belong to ACF class of Dragon descriptors. Atom centered fragments (ACF descriptors) are simple molecular descriptors defined as the number of specific atom types in a molecule and their calculation is based on the knowledge of the molecular composition and atom connectivities. Both the ACF class descriptor C-001 (representing a CH3X fragment in a molecular structure) and N-069 (corresponding to Ar-NH2 or X-NH2) have shown positive influence on the activity suggesting presence of such types of fragments for improved activity. The descriptor SIC4 (structural information content of 4th order neighborhood symmetry) is representative of TOPO class. The TOPO class descriptors are based on a graph representation of the molecule and are numerical quantifiers of molecular topology obtained by the application of algebraic operators to matrices representing molecular graphs and whose values are independent of vertex numbering or labeling. They can be sensitive to one or more structural features of the molecule such as size, shape, symmetry, branching and cyclicity and can also encode chemical information concerning atom type and bond multiplicity. The negative contribution of descriptor SIC4 recommended a lower value of structural information content of 4th order neighbourhood symmetry for better activity. The remaining descriptor ARR (from EMP class of descriptors which are 1D-descriptors represent the counts of non-single bonds, hydrophilic groups and ratio of the number of aromatic bonds and total bonds in an H-depleted molecule) is aromatic ratio. It has shown positive contribution to the activity suggesting a higher value of aromatic ratio for the elevated Met kinase inhibitory activity of pyrrolotriazine derivatives.

The highest significant model in four descriptors could estimate nearly 76 % variance in observed activity of the compounds. Considering the number of observation in the dataset, models with up to five descriptors were explored. Following are some five-descriptor models for the activity.

$$pIC_{50} = -30.427(3.711)SIC4 - 0.001(0.000)SRW09 - 4.142(1.017)GATS7e + 1.841(0.216)N\text{-}069 + 4.958(1.407)ARR + 37.718$$

$$n = 27,\ r = 0.910,\ s = 0.309,\ F = 20.226,\ AIC = 0.150,\ FIT = 1.945,$$
$$LOF = 0.187,\ Q^2_{LOO} = 0.686,\ Q^2_{L5O} = 0.642,\ r^2_{randY}(sd) = 0.417(0.137),$$
$$r^2_{Test} = 0.528 \quad (8)$$

pIC_{50} = -0.573(0.220)nR05 -27.462(3.790)SIC4 - 3.450(0.983)GATS7e + 1.801(0.216)N-069 + 4.681(1.450)ARR + 34.050

n = 27, r = 0.907 s = 0.314, F = 19.420, AIC = 0.155, FIT = 1.867, LOF = 0.194, Q^2_{LOO} = 0.663, Q^2_{L5O} = 0.738, r^2_{randY}(sd) = 0.409(0.130), r^2_{Test} = 0.537 (9)

pIC_{50} = -25.871(4.050)SIC4 + 2.564(1.082)BELe4 - 2.564(1.040)GATS7e + 1.540(0.212)N-069 + 6.550(1.519)ARR + 26.245

n = 27, r = 0.902 s = 0.321, F = 18.431, AIC = 0.162, FIT = 1.772, LOF = 0.202, Q^2_{LOO} = 0.650, Q^2_{L5O} = 0.613, r^2_{randY}(sd) = 0.410(0.121), r^2_{Test} = 0.605 (10)

pIC_{50} = -14.867(6.370)SIC4 + 4.049(1.134)BELe4 + 0.287(0.137)C-005 + 1.294(0.191)N-069 + 5.767(1.607)ARR + 11.298

n = 27, r = 0.896 s = 0.331, F = 17.024, AIC = 0.173, FIT = 1.637, LOF = 0.216, Q^2_{LOO} = 0.575, Q^2_{L5O} = 0.619, r^2_{randY}(sd) = 0.431(0.119), r^2_{Test} = 0.525 (11)

These models have accounted for up to nearly 83 % variance in the observed activities. In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation. The values greater than 0.5 of Q^2-index is in accordance to a reasonable robust QSAR model. The pIC_{50} values of training and test set compounds calculated using equations (8) to (11) have been included in Table-3. The models (8) to (11) are validated with an external test set of 9 compounds listed in Table-1. The predictions of the test set compounds based on external validation are found to be satisfactory as reflected in the test set r^2 (r^2_{Test} > 0.5) values. The plot showing goodness of fit between observed and calculated activities for the training and test set compounds is given in Fig. 1.

The newly participated descriptors in models given above are SRW09 (from MWC class), GATS7e (from 2D-AUTO class) and nR05 (from CONST class). Molecular walk counts (MWC class descriptors) are 2D-descriptors representing molecular walks counts and self-returning walks counts of different lengths. The descriptor SRW09 (9th order self-returning walk count) suggests that a smaller length of 9th order self returning walk would be beneficiary to the activity.

The 2D autocorrelations are molecular descriptors which describe how a considered property is distributed along a topological molecular structure. The 2D-AUTO descriptors have their origin in autocorrelation of topological structure of Broto-Moreau (ATS), of Moran (MATS) and of Geary (GATS). The computation of these descriptors involve the summations of different autocorrelation functions corresponding to the different fragment lengths and lead to different autocorrelation vectors corresponding to the lengths of the structural fragments. Also a weighting component in terms of a physicochemical property has been embedded in this descriptor. As a result these descriptors address the topology of the structure or parts thereof in association with a selected physicochemical property. In these descriptors'

TABLE-3
OBSERVED AND MODELED MET KINASE INHIBITORY ACTIVITY OF PYRROLOTRIAZINE DERIVATIVES

Compd.[a]	pIC_{50} (M)[b]				
	Observed[c]	Calculated			
		Eqn. 8	Eqn. 9	Eqn. 10	Eqn. 11
1	6.12	5.86	5.87	5.79	5.81
2	7.05	6.95	6.98	7.09	7.11
3	7.40	7.19	7.17	7.21	7.05
4[d]	7.52	7.36	7.36	7.45	7.35
5[d]	7.05	7.14	7.15	7.23	7.16
6	7.05	6.88	6.93	7.09	7.19
7	7.15	6.98	7.00	7.08	7.05
8	6.80	7.15	7.16	7.27	7.20
9	6.72	6.52	6.42	6.70	6.45
10	6.29	6.72	6.76	6.81	6.54
11	7.52	7.42	7.43	7.43	7.54
12[d]	7.15	7.17	7.19	7.21	7.37
13	6.11	6.53	6.48	6.35	6.23
14[d]	7.40	7.10	7.13	7.17	7.17
15	6.89	6.73	6.79	6.80	6.59
16	6.85	7.32	7.33	7.15	7.22
17	7.52	7.15	7.19	7.12	7.32
18[d]	7.30	7.11	7.11	7.08	6.96
19	6.82	7.27	7.26	7.00	6.98
20	7.40	7.15	7.15	7.01	7.10
21	5.70	5.87	5.86	5.48	5.38
22	6.59	6.52	6.52	6.82	6.89
23[d]	5.85	6.04	6.01	5.88	5.85
24	6.39	6.39	6.36	6.37	6.39
25[d]	6.07	6.69	6.62	6.61	6.80
26	6.00	5.98	5.99	5.94	6.07
27	6.36	6.63	6.69	6.88	6.92
28	6.72	6.63	6.69	6.88	6.92
29[d]	6.74	6.03	6.02	6.11	6.28
30	7.35	7.84	7.80	7.63	7.57
31	6.05	6.01	5.99	6.05	6.19
32	6.36	6.26	6.20	6.18	6.16
33[d]	6.82	6.36	6.28	6.31	6.27
34	6.00	5.92	5.93	6.05	6.30
35	7.43	7.25	7.25	7.08	7.20
36	8.82	8.33	8.25	8.21	8.07

[a]As in Table1; [b]On molar basis; [c]Taken from reference [6]; [d]Test set compounds.

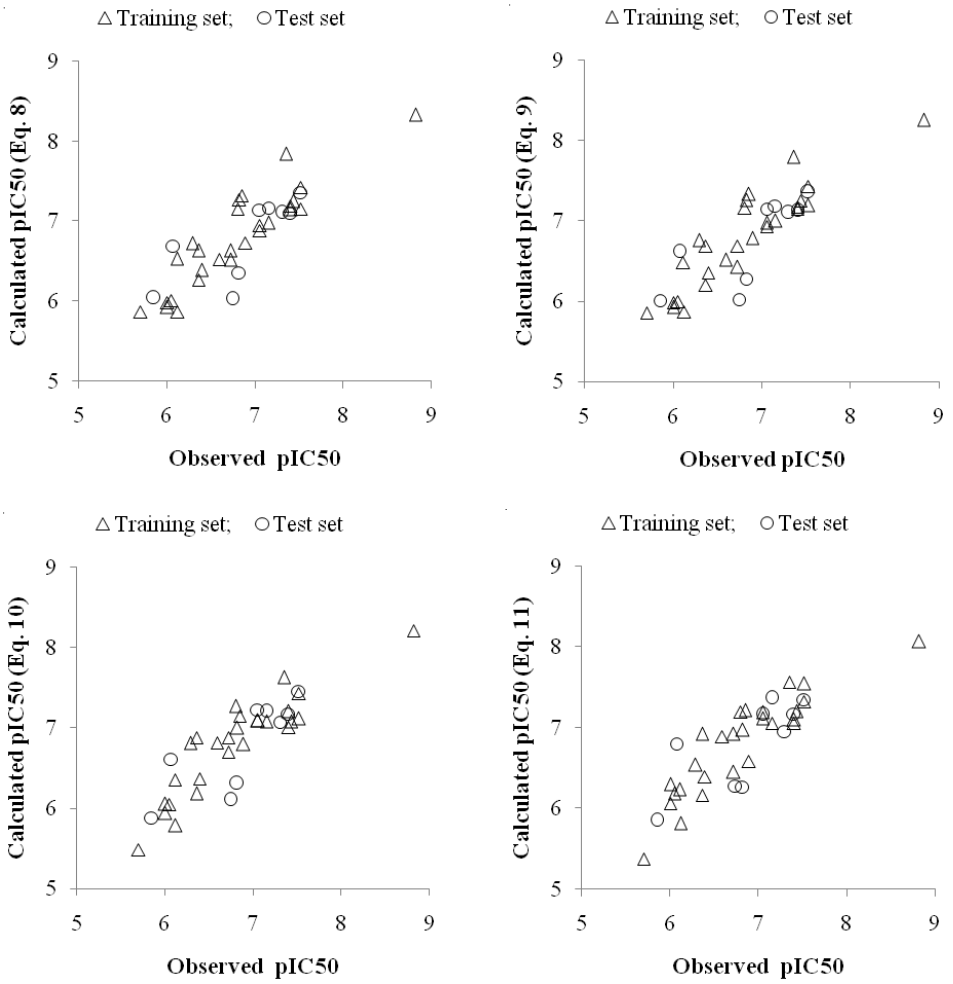


Fig. 1. Plot of observed *versus* calculated pIC_{50} values, from five descriptor CP-MLR models, for the training and test set compounds

nomenclature, the penultimate character, a number, indicates the number of consecutively connected edges considered in its computation and is called as the autocorrelation vector of lag n (corresponding to the number of edges in the unit fragment). The very last character of the descriptor's nomenclature indicates the physicochemical property considered in the weighting component - m for mass or v for volume or e for Sanderson electronegativity or p for polarizability-for its computation. It is evident from the sign of regression coefficients of the participating descriptor that descriptor GATS7e from this class has contributed negatively to the activity. Thus a lower value of descriptor GATS7e (Geary autocorrelation of lag 7 weighted by atomic Sanderson's electronegativities) will be in favour of activity. The CONST class descriptors are dimensionless or 0D descriptors and are independent

from molecular connectivity and conformations. The descriptor nR05 representing number of five membered rings in a molecular structure has shown negative influence on the activity. This suggests that presence of five membered rings in molecular structures would be deleterious to the activity.

Conclusion

The study has provided a rational approach for the development of new pyrrolotriazine derivatives as Met kinase inhibitors. The descriptors identified in CP-MLR analysis have highlighted the role of atomic Sanderson's electronegativity in Modified Burden eigenvalue (BELe4) and in respective lag of 2D-autocorrelation (GATS7e), 9th order self-returning walk count (SRW09), structural information content of 4th order neighborhood symmetry (SIC4) and aromatic ratio (ARR) to explain the Met kinase inhibitory activity of pyrrolotriazine derivatives. Certain structural fragments (C-005 and N-069) and number of five membered rings (nR05) in molecular structures have also shown prevalence to optimize the Met kinase inhibitory activity of titled compounds. These guidelines may be used to develop new Met kinase inhibitors based on pyrrolotriazine scaffold.

ACKNOWLEDGEMENT

The authors are thankful to their respective Institutions for providing necessary facilities to complete this study.

REFERENCES

1. (a) B. Peruzzi and D.P. Bottaro, *Clin. Cancer Res.*, **12**, 3657 (2006); (b) J.G. Christensen, J. Burrows and R. Salgia, *Cancer Lett.*, **225**, 1 (2005); (c) P.C. Ma, G. Maulik, J. Christensen and R. Salgia, *Cancer Metastasis Rev.*, **22**, 309 (2003).
2. C.S. Cooper, M. Park, D.G. Blair, M.A. Tainsky, K. Huebner, C.M. Croce and G.F. Vande Woude, *Nature*, **311**, 29 (1984).
3. (a) B.S. Knudsen and E. Magnus, *Adv. Cancer Res.*, **91**, 31 (2004); (b) P.A. Humphrey, X. Zhu, R. Zarnegar, P.E. Swanson, T.L. Ratliff, R.T. Vollmer and M.L. Day, *Am. J. Pathol.*, **147**, 386 (1995).
4. D.A.M. Heideman, P.J.F. Snijders, E. Bloemena, C.J.L.M. Meijer, G.J.A. Offerhaus, S.G.M. Meuwissen, W.R. Gerritsen and M.E. Craanen, *J. Pathol.*, **194**, 428 (2001).
5. L. Schmidt, K. Junker, N. Nakaigawa, T. Kinjerski, G. Weirich, M. Miller, I. Lubensky, H.P. Neumann, H. Brauch, J. Decker, C. Vocke, J.A. Brown, R. Jenkins, S. Richard, U. Bergerheim, B. Gerrard, M. Dean, W.M. Linehan and B. Zbar, *Oncogene*, **18**, 2343 (1999).
6. G.M. Schroeder, X.-T. Chen, D.K. Williams, D.S. Nirschl, Z.-W. Cai, W. Cai, D. Wei, J.S. Tokarski, Y. An J. Sack, Z. Chen, T. Huynh, W. Vaccaro, M. Poss, B. Wautlet, J. Gullo-Brown, K. Kellar, V. Manne, J.T. Hunt, T.W. Wong, L.J. Lombardo, J. Fargnoli and R.M. Borzilleri, *Bioorg. Med. Chem. Lett.*, **18**, 1945 (2008).
7. R.S. Bhide, Z.-W. Cai, Y.-Z. Zhang, L. Qian, D. Wei, S. Barbosa, L.J. Lombardo, R.M. Borzilleri, X. Zheng, L.I. Wu, J.C. Barrish, S.-H. Kim, K. Leavitt, A. Mathur, L. Leith, S. Chao, B. Wautlet, S. Mortillo, R. Jeyaseelan, D. Kukral, J.T. Hunt, A. Kamath, A. Fura, V. Vyas, P. Marathe, C. D'Arienzo, G. Derbin and J. Fargnoli, *J. Med. Chem.*, **49**, 2143 (2006).

8. (a) A.V. Gavai, P. Chen, D. Norris, B.E. Fink, H. Mastalerz, Y. Zhao, W.-C. Han, G. Zhang, W. Johnson, E. Ruediger, P. Dextraze, J.-P. Daris, S.-H. Kim, K. Leavitt, K. Kim, S. Lu, P. Zheng, A. Mathur, D. Vyas, J.S. Tokarski, C. Yu, S. Oppenheimer, H. Zhang, F. Lee, T.W. Wong and G.D. Vite, Abstracts of Papers, 233rd National Meeting of the American Chemical Society, Chicago, IL, March 25-29, 2007, American Chemical Society: Washington, DC, 2007, MEDI-017; (b) A.V. Gavai, W.-C. Han, P. Chen, E.H. Ruediger, H. Mastalerz, B.E. Fink and D.J. Norris, U.S. Patent US 7,064,203 (2006).
9. ChemDraw Ultra 6.0 and Chem3D Ultra, Cambridge Soft Corporation, Cambridge, USA.
10. R. Todeschini and V. Consonni, Dragon Software (Version 1.11-2001), Milano, Italy.
11. Y.S. Prabhakar, *QSAR Comb. Sci.*, **22**, 583 (2003).
12. S. Sharma, Y.S. Prabhakar, P. Singh and B.K. Sharma, *Eur. J. Med. Chem.*, **43**, 2354 (2008).
13. B.K. Sharma, P. Pilania and P. Singh, *J. Enzym. Inhib. Med. Chem.*, **24**, 607 (2009).
14. S. Sharma, B.K. Sharma, P. Pilania, P. Singh and Y.S. Prabhakar, *J. Enzym. Inhib. Med. Chem.*, **24**, 1024 (2009).
15. B.K. Sharma, P. Pilania, P. Singh and Y.S. Prabhakar, *QSAR Environ. Res.*, **21**, 169 (2010).
16. B.K. Sharma, P. Pilania, K. Sarbhai, P. Singh and Y.S. Prabhakar, *Mol. Divers*, **14**, 371 (2010).
17. H. Akaike, in eds.: B.N. Petrov and F. Csaki, Information Theory and An Extension of the Minimum Likelihood Principle, Second International Symposium on Information Theory, Akademiai Kiado, Budapest, pp. 267-281 (1973).
18. H. Akaike, *IEEE Trans Automat. Contr.*, **AC-19**, 716 (1974).
19. H. Kubinyi, *Quant. Struct.-Act. Relat*., **13**, 285 (1994).
20. H. Kubinyi, *Quant. Struct.-Act. Relat.*, **13**, 393 (1994).
21. J. Friedman, In: Technical Report No. 102, Laboratory for Computational Statistics, Stanford University, Stanford (1990).
22. S.-S. So and M. Karplus, *J. Med. Chem.*, **40**, 4347 (1997).
23. Y.S. Prabhakar, V.R. Solomon, R.K. Rawal, M.K. Gupta and S.B. Katti, *QSAR Comb. Sci.,* **23**, 234 (2004).

(*Received*: 22 May 2010; *Accepted*: 5 August 2010) AJC-8962