International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
DOI: 10.22192/ijarbs Coden: IJARQG(USA) Volume 5, Issue 4 - 2018

Research Article

DOI: http://dx.doi.org/10.22192/ijarbs.2018.05.014

Comparative QSAR study on selective CHK1 inhibitors on pyrazolo [1, 5-a] pyrimidines as an anticancer molecule

Jajati Pramanik[1], Indrajit Giri[2], Swapnanil Ray[3], Pallab Dasgupta[3], Milan Kumar Maiti[*1]

[1,2,3]BCDA College of Pharmacy and Technology
(Department of Medicinal Chemistry & Pharmacognosy Lab), Barasat, Kolkata-700127, West Bengal, India.

*Corresponding author: **Milan Kumar Maiti**
Assistant Professor, Department of Medicinal Chemistry, BCDA College of Pharmacy and Technology, Barasat, Kolkata-700127, West Bengal, India.
E-mail: *c*

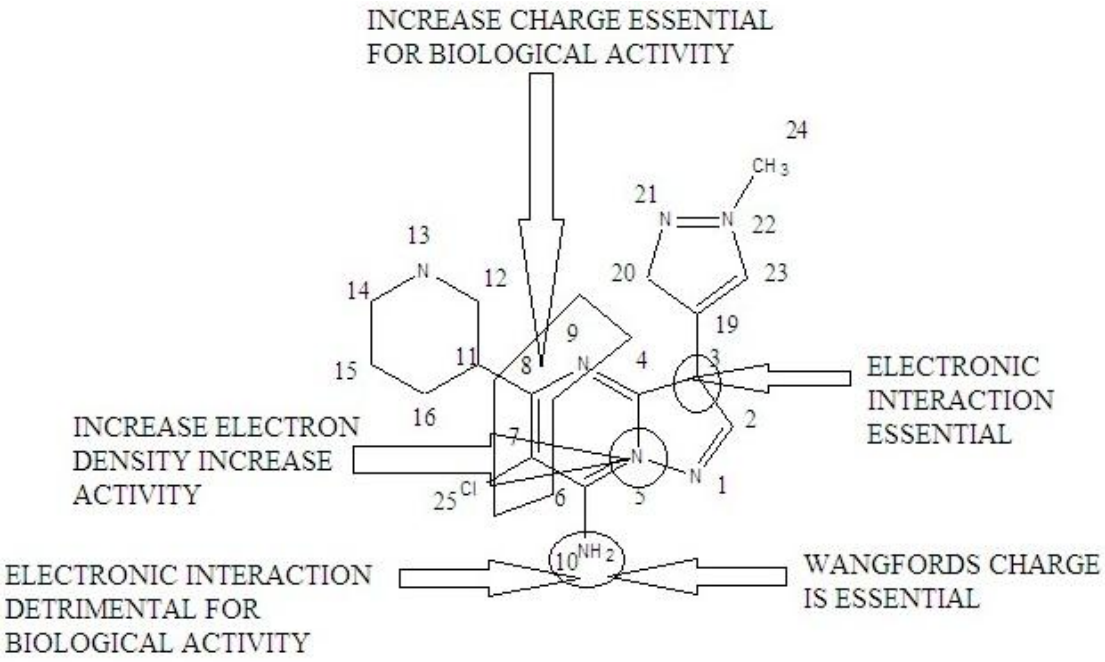
Abstract

Check point kinase 1 is a serine/threonine kinase that controls the cellular damage. Against the DNA damaging agent, CHK1 is activated by phosphorilation. Inhibition of CHK1 initiates cell cycle arrest, resulting in genomic instability & ultimate progression into mitosis & cell-death. The inhibition of CHK1 creates a 'synthetic lethal' response by which aberrant cells can't be replicated which should impede the progression of cancer. Due to the fact that inhibition of CHK1 represents a targeted approaches to enhance the cytotoxicity of DNA damaging agents towards tumour cells while having a lesser effect on normal cells, it has been an attractive target in the oncological field. QSAR study reveals that wangford and electrostatic potential charge is important for that data set.

Keywords: CHK1, DNA damages, Oncology, Mitosis & cell death, Synthetic Lethal response, Cytotoxicity, Wangford charge, Electrostatic potential charge.

1. Introduction

Check point kinase 1 is a serine/threonine kinase that controls the cellular damage[1]. In response to DNA damaging agent, CHK1 is activated by phosphorilation to arrest cells to various cell cycle check points (G1, S, G2) in order to initiate DNA repair process. Inhibition of CHK1 abrogayes cell cycle arrest resulting in genomic instability & ultimate progression into mitosis & cell death[2]. The inhibition of CHK1 creates a 'synthetic lethal' response by which aberrant cells can't be replicated which should impede the progression of cancer. In contrast, normal cells still arrest at G1 check point, via p53, to repair the DNA damage caused by the agents. Due to the fact that inhibition of CHK1 represents a targeted approaches to enhance the cytotoxicity of DNA damaging agents towards tumour cells while having a lesser effect on normal cell, it has been an attractive target in the oncological field[3] .



Increasing charges at atom numbers 7, 8 and 9 may be conducive and at atom number3 and 5 may be detrimental for oncogenic activity. Increasing value of the electrostatic potential charges at atom numbers 3 and 5 may be advantageous thus, electron releasing group at these positions beneficial for enhancing the cytotoxicity of DNA damaging agents. At atom number 10 the nitrogen atom containing lone pair electron may be detrimental but electron withdrawing group as well as increase positive charge may be conducive for better therapeutic activity. The increase the number of multiple bonds of whole molecules is inauspicious for biological activity. More over the statistical analysis give an insight about structural necessities to improve biological activities. The analysis implies that increase the values of van der Waals volume of whole molecules along with enhance charge at atom number 8 (Q8) is conducive for biological activity. Effects of polarizability is limited by Radial Distribution Function i.e., with in short path length of molecular path counts matrix, the effects of polarizability may be better moreover getter will be the biological activity.

Serine/threonine-protein kinase Chk1 is an enzyme that in humans is encoded by the CHEK1 gene[4] CHK1 is a kinase that phosphorylates cdc25, an important phosphatase in cell cycle control, particularly for entry into mitosis. Cdc25, when phosphorylated on serine 216 by chk1 becomes bound by an adaptor protein in the cytoplasm. Therefore it is inhibited from removing the inhibiting phosphate from MPF (mitotic/maturation promoting factor) added by Wee1. Consequently, a cell is prevented from entering into mitosis. **Wee1** is a nuclear kinase belonging to the Ser/Thr family of protein kinases in the fission yeast Schizosaccharomyces pombe (S. pombe). It has a molecular mass of 96 kDa and it is a key regulator of cell cycle progression. It influences cell size by inhibiting the entry into mitosis, through inhibiting Cdc2. It has homologues in many other organisms, including mammals[5]. A **cyclin-dependent kinase inhibitor protein** is a protein which inhibits cyclin-dependent kinase. Several function as tumor suppressor genes. Cell cycle progression is negatively controlled by cyclin-dependent kinases inhibitors (called CDIs, CKIs or CDKIs). CDIs are involved in cell cycle arrest at the G_1 phase[6].

2. DNA damage

DNA damage, due to environmental factors and normal metabolic processes inside the cell, occurs at a rate of 1,000 to 1,000,000 molecular lesions per cell per day. DNA is, however, super coiled and wound around "packaging" proteins called histones (in eukaryotes), and both superstructures are vulnerable to the effects of DNA damage[7].

Sources of damage – DNA damage can be subdivided into two main types:

1. endogenous damage such as attack by reactive oxygen species produced from normal metabolic byproducts (spontaneous mutation), especially the process of oxidative deamination
2. exogenous damage caused by external agents such as

    - ultraviolet [UV 200-300 nm] radiation from the sun
    - other radiation frequencies, including x-rays and gamma rays
    - hydrolysis or thermal disruption
    - certain plant toxins
    - human-made mutagenic chemicals, especially aromatic compounds that act as DNA intercalating agents
    - cancer chemotherapy and radiotherapy
    - viruses[8]

The replication of damaged DNA before cell division can lead to the incorporation of wrong bases opposite damaged ones. Daughter cells that inherit these wrong bases carry mutations from which the original DNA sequence is unrecoverable (except in the rare case of a back mutation, for example, through gene conversion).

Spontaneous damage can include the loss of a base, deamination, sugar ring puckering and tautomeric shift[9].

2.1 DNA repair mechanism

Damage to DNA alters the spatial configuration of the helix, and such alterations can be detected by the cell. Once damage is localized, specific DNA repair molecules bind at or near the site of damage, inducing other molecules to bind and form a complex that enables the actual repair to take place.

This is an expensive process because each MGMT molecule can be used only once; that is, the reaction is stoichiometric rather than catalytic[10].

A generalized response to methylating agents in bacteria is known as the adaptive response and confers a level of resistance to alkylating agents upon sustained exposure by upregulation of alkylation repair enzymes[11].

Cells are known to eliminate three types of damage to their DNA by chemically reversing it. These mechanisms do not require a template, since the types of damage they counteract can occur in only one of the four bases. Such direct reversal mechanisms are specific to the type of damage incurred and do not involve breakage of the phosphodiester backbone. The formation of pyrimidine dimers upon irradiation with UV light results in an abnormal covalent bond between adjacent pyrimidine bases. The photoreactivation process directly reverses this damage by the action of the enzyme photolyase, whose activation is obligately dependent on energy absorbed from blue/UV light (300–500 nm wavelength) to promote catalysis.[12] Another type of damage, methylation of guanine bases, is directly reversed by the protein methyl guanine methyl transferase (MGMT), the bacterial equivalent of which is called ogt. The third type of DNA damage reversed by cells is certain methylation of the bases cytosine and adenine.

When only one of the two strands of a double helix has a defect, the other strand can be used as a template to guide the correction of the damaged strand. In order to repair damage to one of the two paired molecules of DNA, there exist a number of excision repair mechanisms that remove the damaged nucleotide and replace it with an undamaged nucleotide complementary to that found in the undamaged DNA strand[12].

Double-strand breaks, in which both strands in the double helix are severed, are particularly hazardous to the cell because they can lead to genome rearrangements. Three mechanisms exist to repair double-strand breaks (DSBs): non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), and homologous recombination.[13] It was noted that double-strand breaks and a "cross-linkage joining both strands at the same point is irreparable because neither strand can then serve as a template for repair. The cell will die in the next mitosis or in some rare instances, mutate"[13]. DNA ligase helps in repairing chromosomal damage, is an enzyme that joins broken nucleotides together by catalyzing the formation of an internucleotide ester bond between the phosphate backbone and the deoxyribose nucleotides. In NHEJ, DNA Ligase IV, a specialized DNA ligase that forms a complex with the cofactor XRCC4, directly joins the two ends[14]

3. Role of p53 in preventing cancer

p53 (also known as **protein 53** or **tumor protein 53**), is a tumor suppressor protein that in humans is encoded by the *TP53* gene[15]. p53 is important in multicellular organisms, where it regulates the cell cycle and, thus, functions as a tumor suppressor that is involved in preventing cancer. As such, p53 has been described as "the guardian of the genome", the "guardian angel gene", and the "master watchman", referring to its role in conserving stability by preventing genome mutation. The name p53 is in reference to its apparent molecular mass: It runs as a 53-kilodalton (kDa) protein on SDS-PAGE.

4. QSAR study

A number of small molecule CHK1 inhibitors have been described recently and several check point kinase inhibitor, have recently entered the clinic in combination with various DNA damaging agent. Due to the therapeutic value of a CHK1 inhibitor as a chemopotentiator[16]. efforts were directed towards the identification of additional, novel CHK1 inhibitors. It was rationalized that proper substitution around the **pyrazolo[1,5-a]pyrimidine** core might improve the potency & selectively of this series for CHK1 previously **pyrazolo[1,5-a]pyrimidine** core ,which was orally bioavailable & found to be efficacious in a mouse tumor xenograft model.

To find the structural requirements for more active CHK1 inhibitors, QSAR study was performed on thirty-seven **pyrazolo[1,5-a]pyrimidine** moiety containing derivatives as a part of a composite program of rational drug design, discovery and development[17-26] . The general structure of these compounds with arbitrary numbering is shown in **Figure 1**.

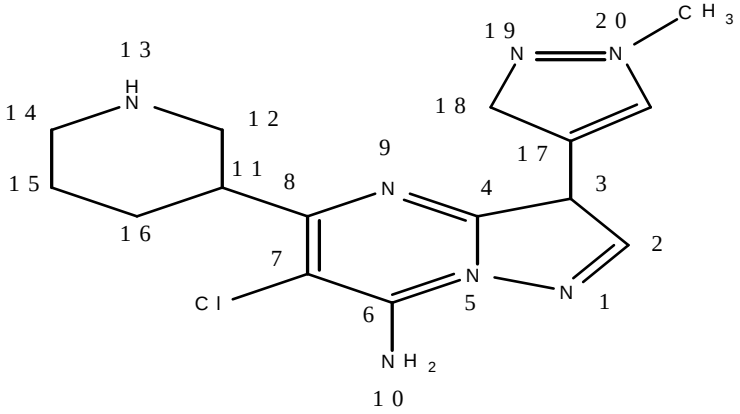


Figure 1. General structure of the data set containing compounds with arbitrary numbering

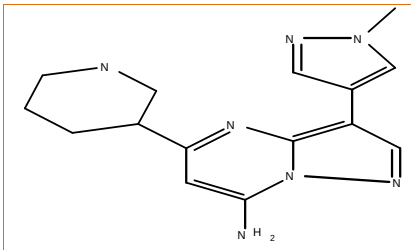
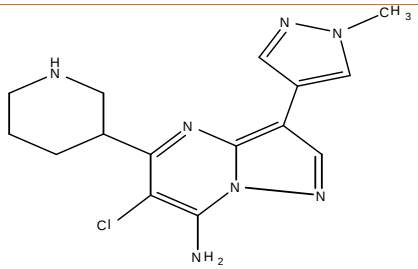
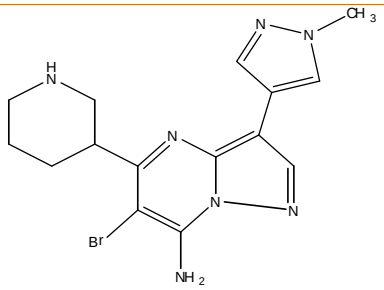
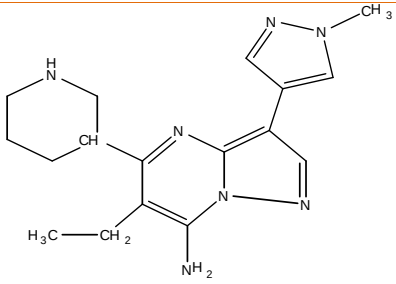
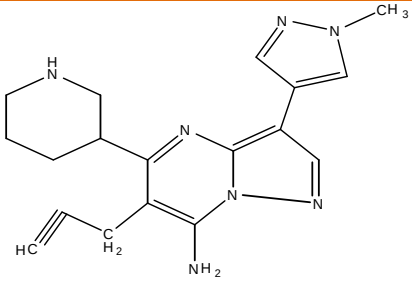
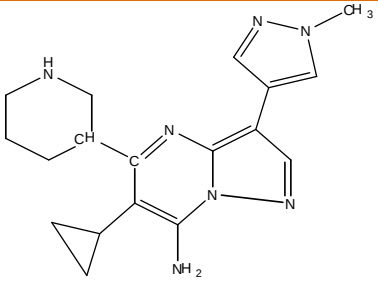
4.1 Dataset and parameters

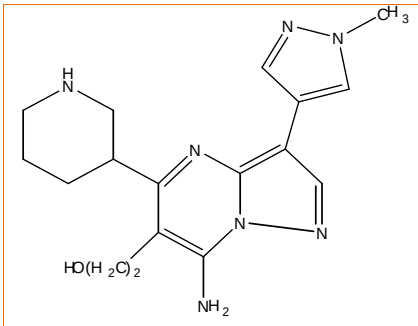
QSAR study was performed on thirty-seven **pyrazolo[1,5-a]pyrimidine** moiety containing derivatives that were synthesized and biologically evaluated previously by Labroli et al[27] . The functional inhibitory activity determined by assay of inhibition of CHK1 were expressed as IC_{50}, were considered as the biological activity parameters for the QSAR study. Here, the negative logarithm of inhibitory activity of these compounds (pIC_{50}) was used to develop QSAR models to obtain linear relationship. pIC_{50} values of all these compounds are shown in **Table 1 and their structure are given in Table 2**.

Table 1. Compounds and their pIC_{50} values

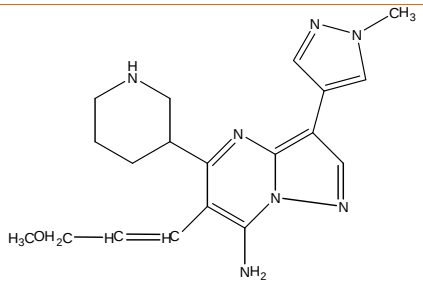
Cpd[a]	pIC_{50}	$pMIC_{50}$	Cpd[a]	pIC_{50}	$pMIC_{50}$
CPD1	0.060	7.221849	**CPD16**	0.100	7.000000
CPD2	0.003	8.522879	**CPD17**	0.017	7.769551
CPD3	0.003	8.522879	**CPD18**	0.025	7.602060
CPD4	0.031	7.508638	**CPD19**	0.008	8.096910
CPD5	0.045	7.346787	**CPD20**	4.000	5.397940
CPD6	0.017	7.769551	**CPD21**	2.900	5.537602
CPD7	0.026	7.585027	**CPD22**	0.035	7.455932
CPD8	0.052	7.283997	**CPD23**	0.130	6.886057
CPD9	0.100	7.000000	**CPD24**	0.006	8.221849
CPD10	0.560	6.251812	**CPD25**	0.031	7.508638
CPD11	0.086	7.065502	**CPD26**	0.034	7.468521
CPD12	0.160	6.79588	**CPD27**	0.087	7.060481
CPD13	0.340	6.468521	**CPD28**	0.007	8.154902
CPD14	0.510	6.292430	**CPD29**	0.007	8.154902
CPD15	0.230	6.638272	**CPD30**	0.005	8.301030

Table 2. Compounds with their structure

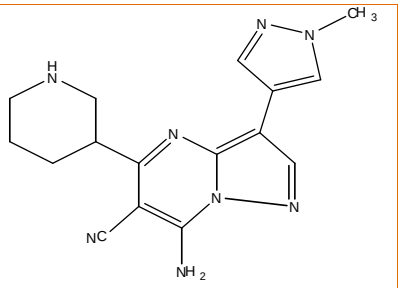
 CPD1	 CPD2	 CPD3
 C PD4	 CPD5	 CPD6



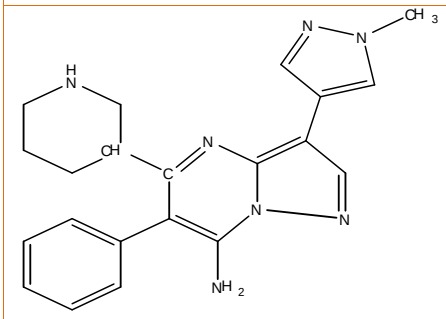
CPD7



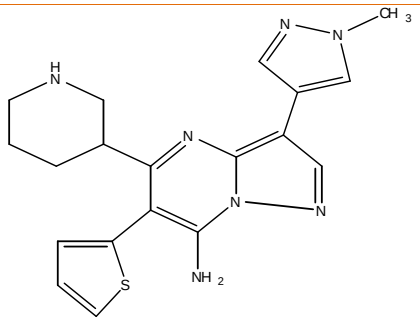
CPD8



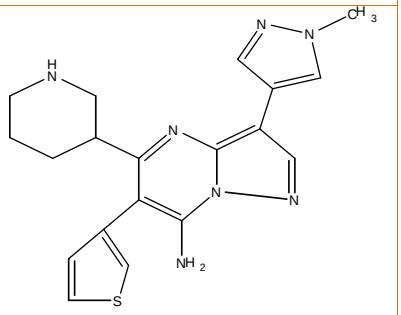
CPD9



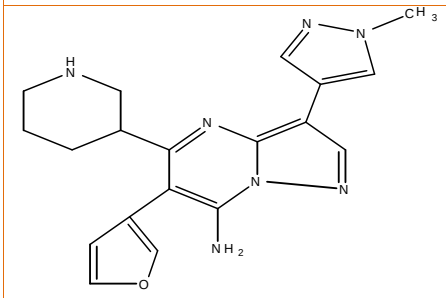
CPD10



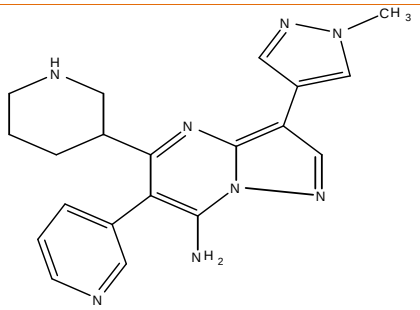
CPD11



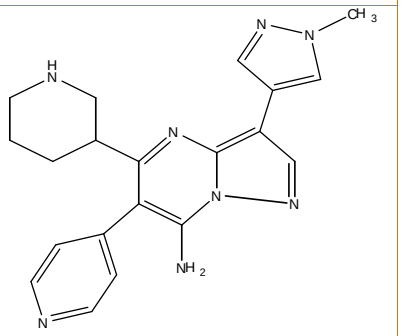
CPD12



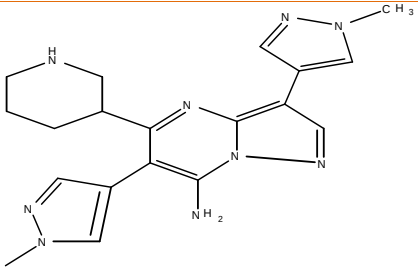
CPD13



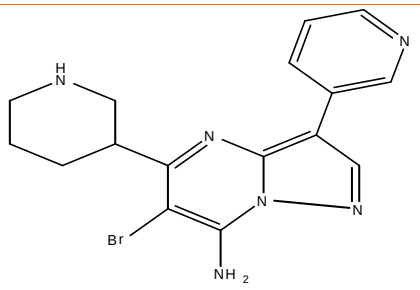
CPD14



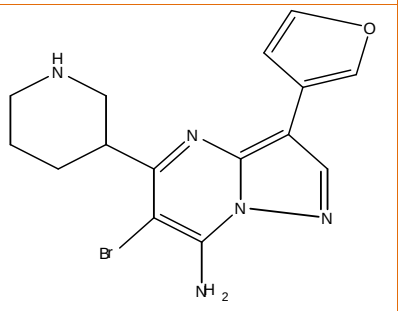
CPD15



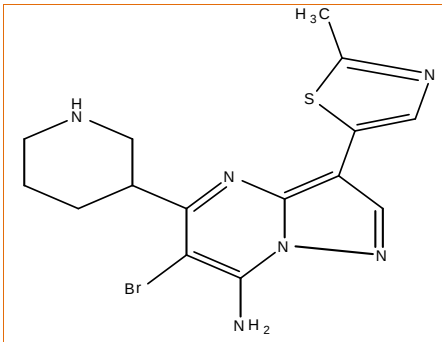
CPD16



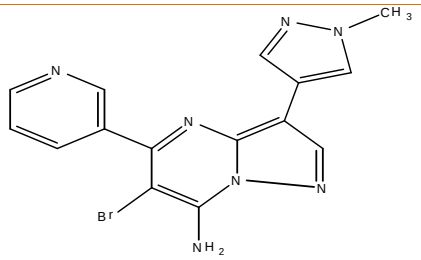
CPD17



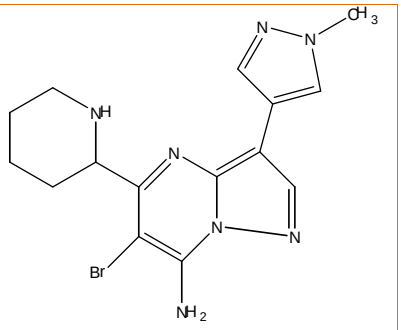
CPD18



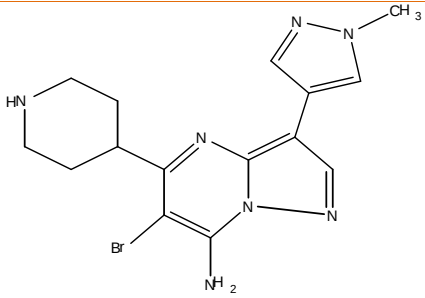
CPD19



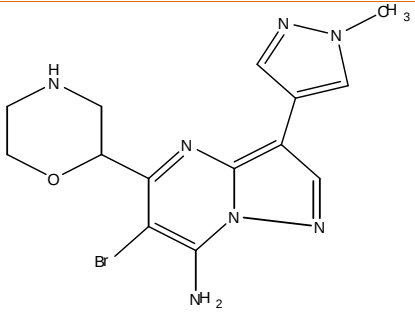
CPD20



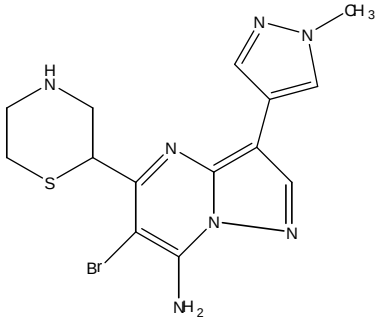
CPD21



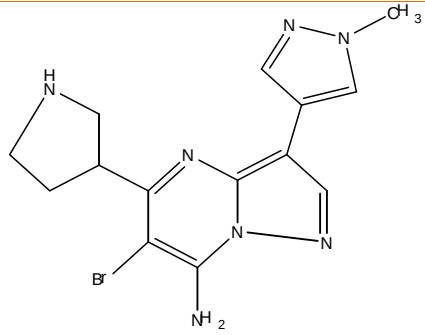
CPD22



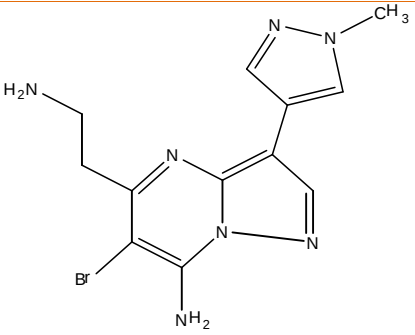
CPD23



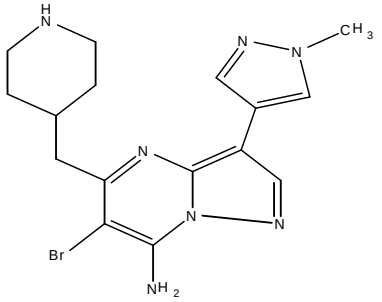
CPD24



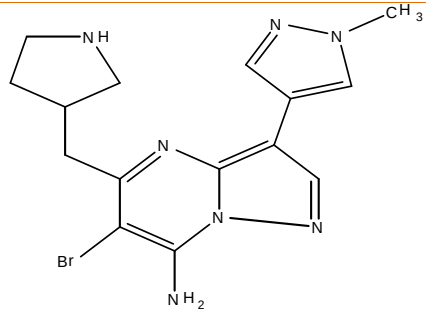
CPD25



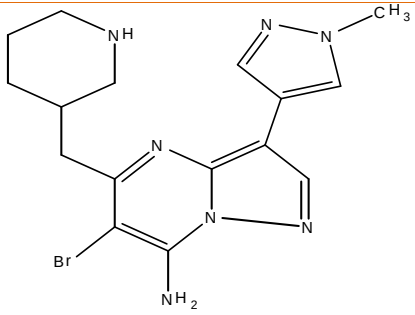
CPD26



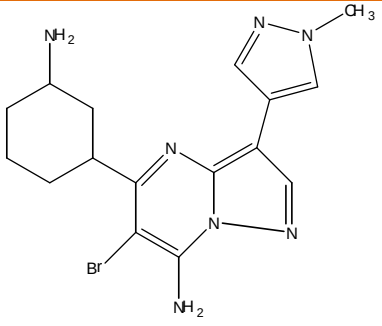
CPD27



CPD28



CPD29



CPD30

Three major types of parameters that are considered as significant are electronic, hydrophobic and steric parameters. The descriptors that were used as independent variables include different informational content indices, electronic descriptors like Wang-Ford charges, frontier electron density, and constitutional descriptors. Moreover different atom based and whole molecular quantum chemical descriptors were also used. Prior to calculation, atoms of these molecules were numbered arbitrarily keeping the serial number of atoms identical for all molecules. The arbitrary numbering is shown in Figure 1.

Different descriptors were calculated by using different type softwares in order to develop QSAR models. Different atom based and whole molecular quantum chemical descriptors were calculated by using Chem. 3D Pro package. The *2D* structure of all molecules was drawn separately in Chem draw ultra ver 5.0[28] .To create the 3-D models, all these structures were transformed to Chem 3D ver 5.0. Energy minimization of individual structure was done under MOPAC module using RHF (restricted Hartee-Fock: closed shell) wave function[29]. These energy minimized geometry were subjected to calculate Wang-Ford charges. Calculation of different quantum chemical descriptors were done using computer program Hyperchem Release 7.0 Pro. Package[30]. Dragon software[31] was utilized for the calculation of different topological, geometrical and constitutional descriptors.

4.2 Statistical procedure

Statistical methods like cluster analysis, stepwise multiple linear regression (sMLR)[32], factor analysis-multiple linear regression (FA-MLR)[33] and principle component regression analysis (PCRA)[34-35] were used separately for development of QSAR models.

4.3 Cluster analysis

For designing the test set and the training set, k-means cluster analysis (k-MCA) were performed to splits the whole dataset into four clusters. From each cluster, near about 25 % members are selected randomly for designing the test set. Test set is formed with six compounds and the remaining twenty four compounds were treated as the training set. The QSAR models were developed depending on training set.

4.4 Correlation Analysis

Correlation analysis was done with the response variable and independent parameters of the training set. Intercorellated independent parameters were not considered for multiple linear regressions and eliminated stepwise depending on their individual correlation with the biological activity.

4.5 Stepwise Regression

In stepwise regression, multiple linear regression (MLR) analysis was carried out via step by step. A primary stepping criteria based on F values was imposed on the system and predictor variables were added or removed according to this stepping criteria. The search was terminated when the stepping was no longer possible. After excluding the intercorrelated parameters, different combinations of parameters having factor loading more than 0.70 were subjected to MLR technique. Statistical qualities of MLR equations were judged by correlation coefficient (R), adjusted R^2 (R_A^2), variance ratio (F), probability factor related to F-ratio (p) and standard error of estimate (SEE).

4.6 Factor analysis- Multiple linear regression (FA-MLR)

Factor analysis was performed as a data-preprocessing step to select descriptors for QSAR equations. In FA-MLR, factors were extracted by principle component method to explain the data matrix to a higher extent. Followed by, it was rotated by VARIMAX rotation. The purpose of factor analysis was to symbolize the multidimensional data matrix in low dimensional space with minimum loss of information. The endeavor of VARIMAX rotation was to maximize the variance of the new variable as well as to reduce the variance around the new variable. Factor pattern suggested whether the biological activity could be explained satisfactorily within the parameter space or not.

4.7. Principal component regression analysis (PCRA)

In PCRA factor scores were extracted after factor analysis on the data matrix containing all independent variables. Factor analysis was performed after VARIMAX rotation. Regression analysis was performed by considering the factor scores as

interpreter variables. These factor scores were used as independent parameters for developing QSAR equations. As factor score contains information on behalf of different descriptors, the probability of loss of information is less.

4.8 Validation of the QSAR models

For the internal validation, Leave-One-Out (LOO-) cross validation method was used to validate the predictive powers of all MLR, FA-MLR, and PCRA models. The predictive residual sum of square (PRESS), total sum of square (SSY), cross-validated R^2 (R^2_{cv}), deviation error of prediction (SDEP) and standard error of PRESS (S_{PRESS}) were considered for the validation of QSAR models.

4.9 Prediction of test set compounds

On the basis of the developed QSAR models on training set , the activity of the test set compounds were predicted and R^2_{Pred} values for the test set compounds were calculated. The external predictivity of the QSAR models is also judged by the similar factors along with another factor R^2_{Pred}. R^2_{Pred} is an external validated coefficient obtained from training set of compounds and used to define the compound present in the test set by using following expression:

$$R^2_{Pred} = 1 - \frac{\sum (Y_{observed} - Y_{prediction})^2}{\sum (Y_{observed} - Y_{train})^2}$$

5. Results

5.1 Correlation Analysis

Correlation analysis was carried out with the response variable and independent parameters of the training set. Intercorrelated independent parameters were not considered for multiple linear regression analysis and eliminated stepwise depending on their individual correlation with the biological activity. The calculated values of selected independent parameters to develop QSAR models are listed in **Table 3**.

Table 3: Descriptors used to develop QSAR models

Cpd[a]	*Q8*	Q_9	*NBM*	*NAB*	*MATS8p*	*RDF090v*	*RDF050p*	*RDF090p*
1	0.466	-0.645	12.000	12.000	-0.109	1.969	8.966	2.471
2	0.912	-0.760	12.000	12.000	-0.167	6.008	9.629	6.772
3	0.967	-0.709	12.000	12.000	-0.171	1.222	13.249	1.561
4	0.514	-0.631	12.000	12.000	-0.096	1.929	13.133	2.133
5	0.644	-0.667	13.000	12.000	-0.134	3.425	12.953	3.686
6	0.475	-0.551	12.000	12.000	-0.061	5.060	13.927	5.675
7	0.489	-0.619	12.000	12.000	-0.097	3.593	13.391	3.899
8	0.775	-0.666	14.000	14.000	-0.107	5.243	14.116	6.362
9	0.136	-0.202	13.000	12.000	-0.155	1.358	10.198	1.730
10	0.550	-0.560	19.000	19.000	-0.077	5.780	16.257	6.509
11	0.727	-0.670	16.000	16.000	-0.065	6.223	13.988	7.298
12	0.657	-0.703	16.000	16.000	-0.093	6.498	15.442	7.159
13	0.610	-0.695	16.000	16.000	-0.069	6.969	13.105	7.608
14	0.785	-0.696	19.000	19.000	-0.074	6.734	16.517	7.489
15	0.757	-0.653	19.000	19.000	-0.076	8.002	16.435	8.494
16	0.013	-0.614	16.000	16.000	-0.078	3.395	15.170	3.966
17	0.742	-0.708	15.000	15.000	-0.099	3.851	13.781	4.309
18	0.649	-0.715	12.000	12.000	-0.119	1.279	10.457	1.787
19	0.602	-0.692	12.000	11.000	-0.162	2.227	14.347	2.874
20	0.717	-0.697	19.000	19.000	-0.232	3.271	6.941	3.697
21	-0.819	-0.700	12.000	12.000	-0.106	2.870	9.772	3.597
22	0.995	-0.782	12.000	12.000	-0.106	2.870	9.772	3.597

23	0.791	-0.676	12.000	12.000	-0.105	1.827	9.197	2.253
24	1.161	-0.828	12.000	12.000	-0.172	1.942	9.902	2.373
25	0.599	-0.555	12.000	12.000	-0.174	0.856	7.837	1.090
26	0.748	-0.717	12.000	12.000	-0.121	1.760	8.188	2.185
27	0.938	-0.727	12.000	12.000	-0.126	5.631	10.710	6.355
28	0.540	-0.514	12.000	12.000	-0.066	2.594	11.055	3.204
29	0.586	-0.410	12.000	12.000	-0.164	4.329	12.267	4.811
30	1.058	-0.749	12.000	12.000	-0.154	2.312	11.871	2.814

5.2 *Validation of* QSAR models

Leave-one-out (LOO) cross validation method was applied to validate the QSAR models developed. Predictive powers of these equations were justified by this method. Predicted residual sum of square (PRESS), total sum of squares (SSY), cross-validated R^2 (R^2_{CV}), standard deviation of error of prediction (SDEP) and standard error of PRESS (S_{PRESS}) were considered for validation of these models.

6. Stepwise Regression

In stepwise method, using stepping criteria on the basis of F value (F = 1.00 for inclusion; F = 0.00 for exclusion), following best equation (Eq.1) was derived with four variables.

pIC_{50} = 7.590(±0.471) – 0.183(±0.034) nBM + 1.149(±0.153) Q8 + 0.192(±0.032) RDF050v – 0.184(±0.048) RDF090v

Eq. (1)

n= 24; R= 0.943; R^2= 0.889; R^2_A = 0.866; F (4, 19) = 38.069; p<0.0000; S.E.E. = 0.291; SSY= 14.476; PRESS= 2.842; R^2_{cv}= 0.803; SDEP= 0.118; S_{PRESS}= 0.149; R^2_{Pred} =0.755.

Where n is the number of data points i.e., the number of pyrazolo [1,5-a] pyrimidine ring containing compounds present in the data sets. The 95% confidence intervals of the regression coefficient are shown in parentheses. Eq. (1) explains 94.30% and predicts 88.90% variances of the inhibition of CHK1 inhibitors activity. Values within the parenthesis are the standard error of corresponding parameters. Eq. 1 suggests significances of the number of multiple bonds (nBM), charge at atom number 8 (Q8), Radial Distribution Function - 050 weighted by van der Waals volume (RDF050v) and Radial Distribution Function - 090 weighted by van der Waals volume (RDF090v), in the biological activity. The positive coefficients of those the above mentioned parameters implies that higher values of these parameters may correspond to the superior inhibitory activity as far as the better CHK1 inhibitors activity is concerned. The negative coefficient of the descriptors nBM is implies that increase the number of multiple bonds of whole molecules is inauspicious for biological activity. The negative coefficients of the descriptors RDF050v and RDF090v are implies that increaser the values of these descriptors as well as increase van der Waals volume of whole molecules advantageous for biological activity. Eq. 1 also recommend that enhance charge at atom number 8 (Q8) is conducive for biological activity.

The Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 1 is shown in following Figure 2.

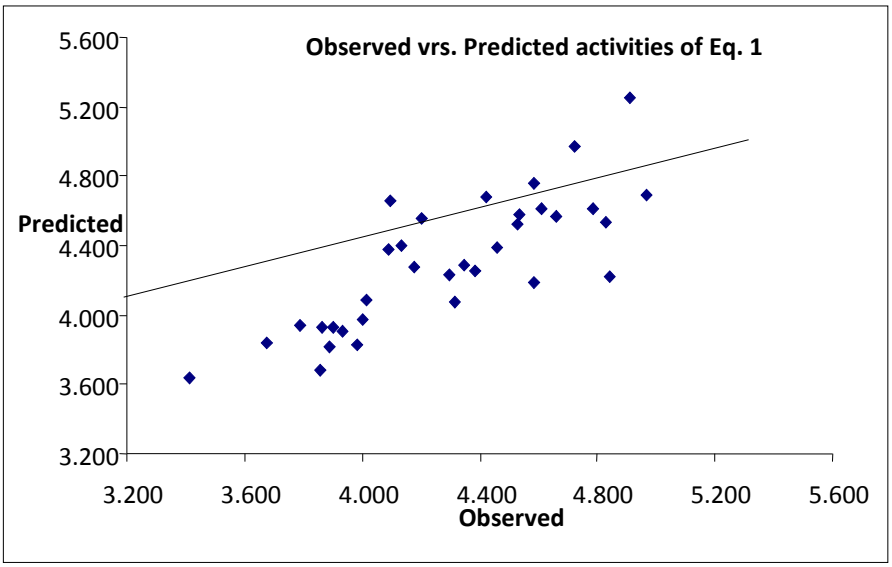


Figure 2. Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 1

Attempts were made to develop another QSAR models for the CHK-1 inhibitory activity for whole data set by increasing number of descriptors. By forward stepwise selection method on the basis of F value (F=1.2for inclusion; F=0.1 for exclusion), Eq. 2 was derived.

pIC_{50} = 7.185(±0.405) – 0.155(±0.028) nBM + 1.073(±0.126) Q8 + 0.192(±0.024) RDF050p – 0.139(±0.024) RDF090v – 0.171(±0.064) RDF100p Eq. (2)

n= 24; R= 0.964; R^2= 0.930; R^2_A = 0.910; F (5, 18) =47.694; p<0.00000; S.E.E. = 0.237; SSY= 14.467; PRESS= 1.952; R^2_{cv}= 0.865; SDEP= 0.0.081; S_{PRESS}= 0.108; R^2_{Pred} =0.594.

Where n is the number of data points i.e., the number of pyrazolo[1,5-a]pyrimidine compounds present in earlier equation. Eq. (2) explains 93.00% and predicts 86.50% variances of the inhibitory activity of CHK1. Eq. 2 also suggests importances of the atom base descriptors i.e charge at atom number8 (Q_3) and other new whole molecular descriptors i.e., Radial Distribution Function - 050 weighted by polarizability (RDF050p) and Radial Distribution Function - 100 weighted by polarizability (RDF100p), in the biological activity. The positive coefficients of those the above mentioned parameters implies that elevated values of these parameters may keep up a correspondence to the higher inhibitory activity. Effects of polarizability is limited by Radial Distribution Function i.e., with in short path length of molecular path counts matrix, the effects of polarizability may be better moreover getter will be the biological activity. The equation is valuable to extract information for further works but not authenticated in term of QSAR or QSPR study as the number of data point is less in this particular set. The Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 2 is shown in Figure 3.

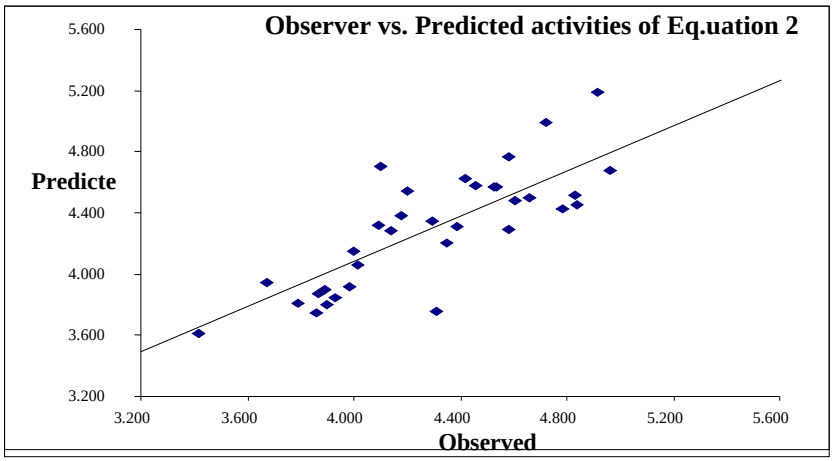


Figure 3. Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 2

6.1 Factor analysis- Multiple linear regression (FA-MLR)

Factor analysis was performed after VARIMAX rotation and different combinations of parameters having factor loading of more than 0.7 (excluding very poorly loaded factors) were subjected to multiple linear regression. Intercorrelation less than 0.5 in between variables were taken in consideration for the development of equations. The results evolved with Eq. (3) with five descriptors after regression analysis:

pIC_{50} = 8.138(±0.494) – 0.144(±0.031) nAB + 1.255(±0.151) Q8 + 0.141(±0.028) RDF050p – 0.216(±0.047) RDF090p + 4.242(±1.920) MATS8p Eq. (3)

n= 24; R= 0.946; R^2= 0.912; R^2_A = 0.887; F (5, 18) = 37.182; p<0.0000; S.E.E. = 0.266; SSY= 14.476; PRESS= 2.196; R^2_{cv}= 0.848; SDEP= 0.091; S_{PRESS}= 0.122; R^2_{Pred} =0.631.

It was observed that 10 factors can explain the data matrix to the extent of 94.60% and predicts only 84.80% for training sets compound of the data set. The pIC_{50} is highly loaded with factor 2, 4 and 1, moderately loaded with factor 3, 6, 7 and poorly loaded with factor 5, and 8. The factor loading of these parameters are shown in Table 3. The newly developed descriptor in that equation shows importance's towards antagonistic activity be fond of number of aromatic bonds(nAB); Radial Distribution Function - 050 / weighted by polarizability (RDF050p); Radial Distribution Function - 090 weighted by polarizability (RDF090p); and Moran autocorrelation of lag 8 weighted by polarizability (MATS8p) are also observed. The negative coefficients all of these descriptors are implies that increaser the values of these descriptors are detrimental for biological activity. Here, also the effects of polarizability has a threshold limit that are guided by Radial Distribution Function i.e., with in short path length of molecular path counts matrix, the effects of polarizability may be better moreover getter will be the biological activity for whole molecule. The negative coefficient of the descriptors nAB is implies that increase the number of multiple bonds in aromatic ring of whole molecules is inauspicious for biological activity. The Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 3 is shown in Figure 4.

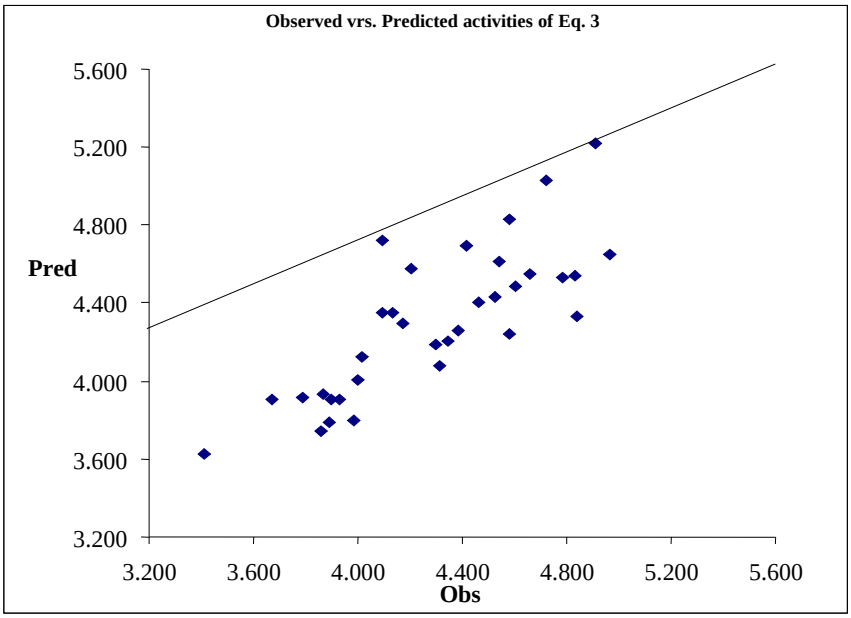


Figure 4. Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 3

6.2 Principal component regression analysis (PCRA)

In PCRA factor scores were achieve from factor analysis on the data matrix containing all independent variables. Regression analysis ware carried out considering factor scores as predictor variables. Here, ten factor scores were extracted by principle component method and then rotated by VARIMAX rotation. These factor scores were used as sole independent parameters for developing QSAR equations. As factor score contains information for the different descriptors, the chance for loss of information is less. Using forward selection technique the following equation was developed:

pIC_{50}= 7.217 (±0.131) +0.320(±0.125) f_9 – 0.326 (± 0.137) f_1 – 0.281(±0.134) f_2

Eq (4)

n = 24; R= 0.658; R^2= 0.434; R^2_A= 0.348; F (3, 20) = 5.107; p<0.0000; S.E.E = 0.640; R^2_{cv}= 0.163; SSY = 14.676; PRESS = 12.119; SDEP=0.505; S_{PRESS}= 0.638; R^2_{Pred} =0.638.

Eq. (4) explains 56.50% and predicts 53.20% variances of the biological activity and shows the importance of factor 1, 2 and 9. This equation also shows that the factor 9(f_9) is highly significant for that data set and factor loading is moderate among eight factors. The Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 4 is shown in Figure 5.

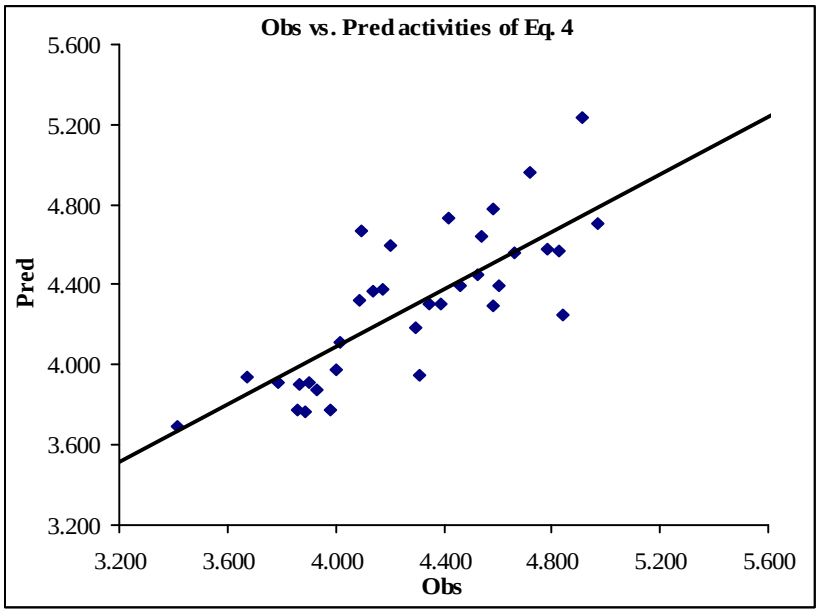


Figure 5. Observed (Obs) vs. LOO-predicted (Pred) activities of Eq. 4

All coefficients of parameters and intercepts in all equations are of 95% confidence intervals as supported by their t- and p- values. t- and p- values of equations 1-4 are shown in Table 4.

Table 4. The *t*- and *p*- values of equations 1-4

	Eq.1 t-VALUE	p-VALUE		Eq.2 t-VALUE	p-VALUE
Intercept	16.11040	0.000000	Intercept	17.71977	0.000000
nBM	-5.38423	0.000034	nBM	-5.44949	0.000035
Q8	7.49965	0.000000	C8	8.51496	0.000000
RDF050v	5.94801	0.000010	RDF050p	7.95208	0.000000
RDF090v	-3.78603	0.001249	RDF090v	-3.20048	0.004957
			RDF100p	-2.67982	0.015293
	Eq.3			Eq4	
Intercept	16.46792	0.000000	Intercept	54.91975	0.000000
Nab	-4.57062	0.000237	NewVar9	2.56240	0.018574
C8	8.28302	0.000000	NewVar1	-2.38117	0.027309
RDF050p	5.02823	0.000087	NewVar2	2.09478	0.049128
RDF090p	-4.56992	0.000237			
MATS8p	2.20862	0.040408			

The correlation matrix among the response variable and those selected descriptors is shown in **Table 5.**

Table 5. correlation matrix among the response variable

	pMIC50	C7	C8	nBM	nAB	MATS8p	RDF090v	RDF050p	RDF090p
pMIC50	1.00	0.08	0.54	-0.62	-0.62	-0.08	-0.45	0.09	-0.45
C7		1.00	-0.25	0.11	0.08	0.22	-0.08	0.28	-0.09
C8			1.00	-0.01	0.01	-0.24	0.08	0.02	0.07
nBM				1.00	0.99	0.18	0.69	0.42	0.68
nAB					1.00	0.21	0.70	0.39	0.68
MATS8p						1.00	0.55	0.56	0.56
RDF090v							1.00	0.59	1.00
RDF050p								1.00	0.58
RDF090p									1.00

7. Conclusion

The QSAR study disclose the structural requirements of on pyrazolo [1,5-a] pyrimidines for oncogenic activity. It is apparent from the QSAR models that the attachment of electron donating group at C-3(Fig. 1) may be advantageous for anticancer activity. The study reveals that the increasing charges at atom numbers 7, 8, 9 and 10 may be conducive but the increase in the charge at the atom number 5 may be damaging for oncogenic activity.

The study also reveals that the topological structure of the molecules depends on atomic mass, van der waal,s volume as well as on polarizibilities. At atom number 10 the nitrogen atom containing lone pair electron may be detrimental but electron withdrawing group as well as increase positive charge may be conducive for much better therapeutic activity. The increase the number of multiple bonds of whole molecules is inauspicious for biological activity. The increaser the values of van der Waals volume of whole molecules is advantageous for biological activity and enhance charge at atom number 8 (Q8) is conducive for the same. Effects of polarizability is limited by Radial Distribution Function i.e., with in short path length of molecular path counts matrix, the effects of polarizability may be better that lead to a getter the biological activity concerned. The increase the number of multiple bonds in aromatic ring of whole molecules is inauspicious for biological activity.

The important atoms and substituents of pyrazolo [1,5-a] pyrimidines derivatives for anticancer activity. The result was supported by the energy minimized geometry as well as the 3D isosurface electrostatic potential map of the most active compound (shown in Figure 6) obtained during the energy minimizations by AM1 calculations.

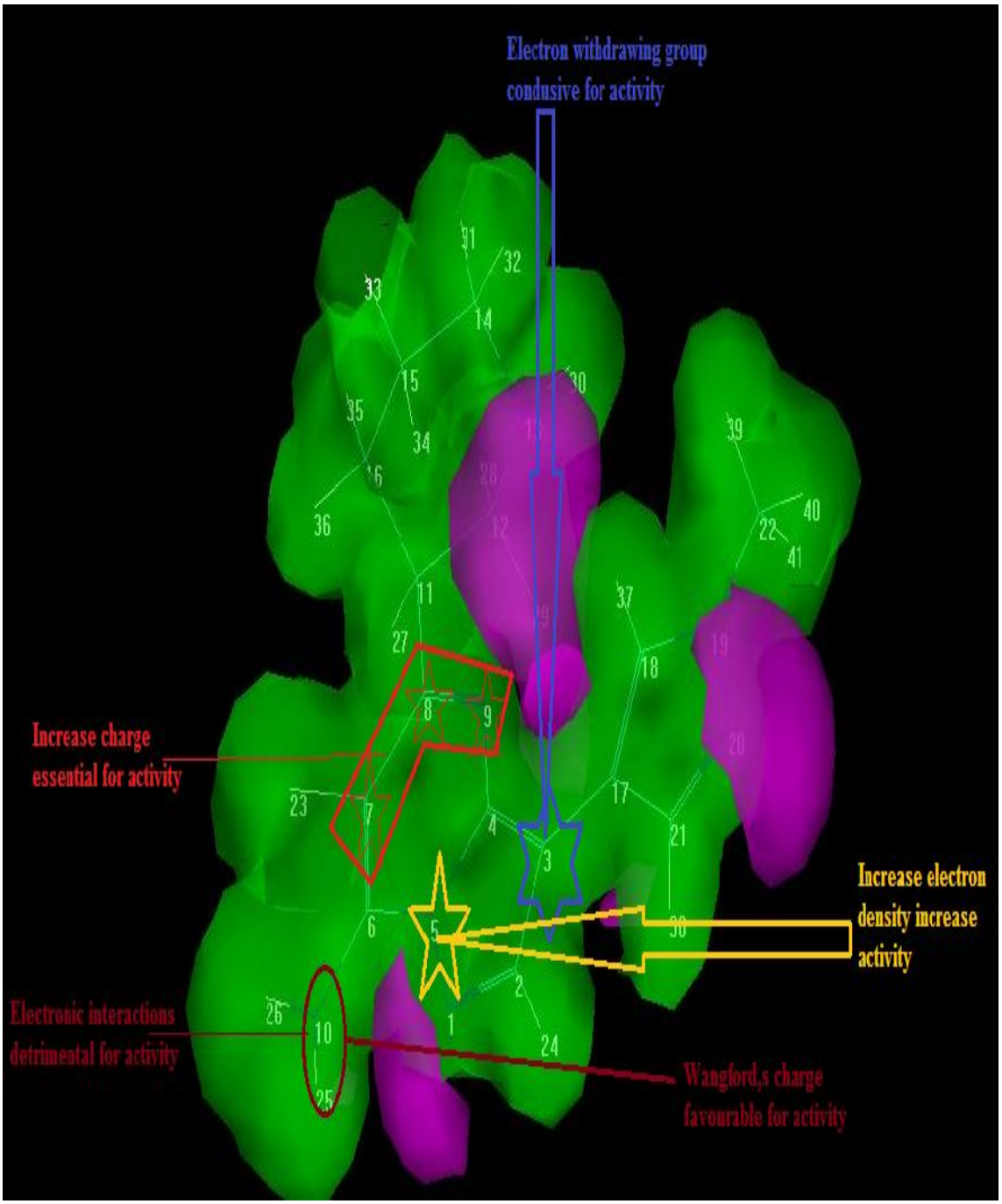


Figure 6. 3D isosurface electrostatic potential map of the most active compound (compound 3)

8. Acknowledgments

Authors are thankful to BCDA College of pharmacy and Technology (Department of medicinal chemistry & Pharmacognosy), Barasat, Kolkata-700127, West Bengal, India, for its technical support to perform this research work.

9. References

1. Russo AA, Jeffrey PD, Patten AK, Massagué J, Pavletich NP (July 1996). "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex". *Nature* **382** (6589): 325–31
2. Hoshino R, Chatani Y, Yamori T, Tsuruo T, Oka H, Yoshida O, Shimada Y, Ari-i S, Wada H, Fujimoto J, Kohno M (January 1999). "Constitutive activation of the 41-/43-kDa mitogen-activated protein kinase signaling pathway in human tumors". *Oncogene* **18** (3): 813–22
3. Peng CY, Graves PR, Thoma RS, *et al.* (1997). "Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216.". *Science* **277** (5331): 1501–5.
4. Ouyang B, Li W, Pan H, *et al.* (1999). "The physical association and phosphorylation of Cdc25C protein phosphatase by Prk.". *Oncogene* **18** (44): 6029–36.
5. Sanchez Y, Wong C, Thoma RS, Richman R, Wu Z, Piwnica-Worms H, Elledge SJ (Sep 1997). "Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25". *Science* **277** (5331): 1497–501.
6. Flaggs G, Plug AW, Dunks KM, Mundt KE, Ford JC, Quiggle MR, Taylor EM, Westphal CH, Ashley T, Hoekstra MF, Carr AM (Feb 1998). "Atm-dependent interactions of a mammalian chk1 homolog with meiotic chromosomes". *Curr Biol* **7** (12): 977–86.
7. Giaccia AJ, Kastan MB (1998). "The complexity of p53 modulation: emerging patterns from divergent signals.". *Genes Dev.* **12** (19): 2973–83.
8. Kastan MB, Lim DS (2001). "The many substrates and functions of ATM.". *Nat. Rev. Mol. Cell Biol.* **1** (3): 179–86.
9. De Boer, J; Andressoo, JO; De Wit, J; Huijmans, J; Beems, RB; Van Steeg, H; Weeda, G; Van Der Horst, GT et al. (2002). "Premature aging in mice deficient in DNA repair and transcription". *Science* **296** (5571): 1276–9.
10. Cromie, GA; Connelly, JC; Leach, DR (2001). "Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans". *Mol Cell.* **8** (6): 1163–74.
11. Storey A, Thomas M, Kalita A, Harwood C, Gardiol D, Mantovani F, Breuer J, Leigh IM, Matlashewski G, Banks L (May 1998). "Role of a p53 polymorphism in the development of human papilloma virus-associated cancer". *Nature* **393** (6682): 229–34.
12. Bell S, Klein C, Müller L, Hansen S, Buchner J (2002). "p53 contains large unstructured regions in its native state". *J. Mol. Biol.* **322** (5): 917–27.
13. O'Brien, PJ. (2006). "Catalytic promiscuity and the divergent evolution of DNA repair enzymes". *Chem Rev* **106** (2): 720–52.
14. Piskacek S, Gregor M, Nemethova M, Grabner M, Kovarik P, Piskacek M (June 2007). "Nine-amino-acid transactivation domain: establishment and prediction utilities". *Genomics* **89** (6): 756–68.
15. Boddy MN, Furnari B, Mondesert O, Russell P (May 1998). "Replication checkpoint enforced by kinases Cds1 and Chk1". *Science* **280** (5365): 909–12.
16. Debnath B, Samanta S, Gayen S, Ghosh B, Basu A, Jha T, Internet Electron. J. Mol. Des. 4 (2005) 393-412, www.biochempress.com.
17. Samanta S, Debnath B, Gayen S, Ghosh B, Basu A, Jha T, IL Farmaco 10 (2005) 818-825.
18. Samanta S, Alam Sk. M, Panda P, Jha T, Internet Electron. J. Mol. Des. 5 (2006) 503-514. www.biochempress.com.
19. Samanta S, Gayen S, Ghosh B, Panda P, Srikanth K. Jha T, Int. J. Appl. Chem. 2 (2006) 169-180.
20. Samanta S, Debnath B, Basu A, Gayen S, Srikanth, Jha T, Eur. J. Med. Chem. 41 (2006) 1190-1195.
21. Debnath B, Gayen S, Samanta S, Basu A, Ghosh B, Jha T, Ind. J. Chem. 45 (2006) 93-99.
22. Basu A, Gayen S, Samanta S, Panda P, Srikanth K, Jha T, Can. J. Chem. 84 (2006) 458-463.
23. Labroli M, Paruch K, Dwyer M. P, Alvarez C, Guzi T. J, Bio. Org. Med. Chem. 21 (2011) 471-474.
24. Chem 3d Pro version 5.0 and Chem Draw Ultra version 5.0 are programs of Cambridge Soft Corporation, USA
25. A.R Leach, Molecular modeling principles and applications, second ed., Prentice Hall, 2001.
26. Hyper Chem Professional' software – version 7.0 of Hypercube Inc. www.hyper.com
27. DRAGON web version 2.1 is a QSAR software developed by Milano Chemometrics and QSAR Research Group, Dipartimento di Scienze dell' Ambiente e del Territorio Universitadegli Studio di Milano – Bicocca.
28. Adhikari N, Maiti M.K, Jha T, Eur. J. Med. Chem. 45 (2010)1119-1127.
29. Adhikari N, Maiti M.K, T. Jha, Bioorg. Med. Chem. Lett. 20 (2010) 4021-4026.
30. Franke R, Theoretical drug design methods, 1st ed., Elsevier, Amsterdam, 1984.

31. Leonard J.T, Roy K, Bioorg. Med. Chem. 14 (2006) 1039–1045.
32. Sendecor G.W, Cochran W.G, Multiple Regressions in Statistical Methods, Oxford, New Delhi, India.
33. Halder A.K, Adhikari N, Jha T, Bioorg. Med. Chem. Lett. 19 (2009) 1737-1739.
34. Martens H, Naes T, Multivariate Calibration. Chichester etc.: Wiley, 1989.
35. Finar I.L, Organic Chemistry, 6th ed., Vol I, Pearson Education, Singapore, 2002.

Access this Article in Online	
Quick Response Code	Website: www.ijarbs.com Subject: Medicinal Chemistry
DOI:10.22192/ijarbs.2018.05.04.014	

How to cite this article:
Jajati Pramanik, Indrajit Giri, Swapnanil Ray, Pallab Dasgupta, Milan Kumar Maiti. (2018). Comparative QSAR study on selective CHK1 inhibitors on pyrazolo [1, 5-a] pyrimidines as an anticancer molecule. Int. J. Adv. Res. Biol. Sci. 5(4): 123-140.
DOI: http://dx.doi.org/10.22192/ijarbs.2018.05.04.014