3D-QSAR CoMFA and CoMSIA Study on Benzodipyrazoles as Cyclin Dependent Kinase 2 Inhibitors

Nigus Dessalew[a,*] and Sanjeev Kumar Singh[b]

[a]*Department of Pharmaceutical Chemistry, School of Pharmacy, Addis Ababa University, P.O.Box 1176, Addis Ababa, Ethiopia;* [b]*Bioinformatics Centre, School of Biotechnology, Madurai Kamaraj University, Madurai-625 021, Tamil Nadu, India*

Abstract: Cyclin dependent kinase 2 (CDK2) has appeared as an important drug target over the years with a multitude of therapeutic potentials. With the intention of designing compounds with enhanced inhibitory potencies against CDK2, the 3D-QSAR CoMFA and CoMSIA study on benzodipyrazoles series is presented here. The developed models showed a strong correlative and predictive capability having a cross validated correlation co-efficient of (r^2_{cv}) 0.699 for CoMFA and 0.794 for CoMSIA models. A very good conventional and predicted correlation co-efficients were also obtained: CoMFA (r^2_{ncv}, r^2_{pred}: 0.883, 0.754), CoMSIA (0.937, 0.815). The models were found to be statistically robust and are expected to be of an aid to design and/or prioritize drug likes for synthesis.

Key Words: 3D-QSAR, CoMFA, CoMSIA, CDK2, benzodipyrazoles.

1. INTRODUCTION

Protein phosphorylation and dephosphorylation are important processes in the control of protein functions. Phosphorylation occurs on serine, threonine and tyrosine residues and is catalyzed by protein kinases whose number transcends 800 in the human genome. Because of the importance of protein phosphorylation as a main post-translational mechanism used by cells to regulate enzymes and other proteins and the association of many maladies with its aberrations [1], kinases have increasingly become important targets and the hunt for kinase inhibitors has been intensified and attracted a great attention in drug discovery over the years [2-7].

Cyclin dependent kinases (CDKs) have been characterized extensively in the past two decades. CDKs are serine/threoinine kinases that play a crucial role in the regulation of the complex processes of the cell division cycle, apoptosis, transcription and differentiation [8,9]. Thus far family members of this class has been identified to include over nine CDKs . CDKs are inactive as monomers and activation requires binding to the corresponding cyclins, a diverse family of proteins whose levels oscillate during the cell cycle and phosphorylation by CDK-activating kinase (CAK) on a specific threonine residue. The basic cell cycle is divided into four phases: namely G_1, S, G_2, and M. Specific CDKs operate in the distinct phases of cell cycle. Cyclin dependent kinase 2 (CDK2) can form complexes with both cyclins A and E, and it is required for the G1/S transition and S phase progression [10], and centrosome duplication [11]. Because of the involvement of CDKs in the crucial processes of the cell, inhibitors of this kinase have got a wide spectrum of potential therapeutic applications ranging from protozoan infections (malaria, leishmania, trypanosomiasis), viral infections (HCMV, HSV, HIV, HPV), reproduction disorders, cardiovascular diseases (atherosclerosis, restenosis, cardiac hypertrophy), gulomerulonephritis, cancers to nervous system disorders (Alzheimer's disease, stroke, amyotrophic disease, drug abuse) [12].

Since its introduction in 1988, CoMFA [13] has emerged as one of the most powerful tools in ligand based drug design strategies [14-16]. It has combination of reasonable molecular description, statistical analysis and graphical display of results. Molecular structures are described with molecular interaction energies as steric and electrostatic fields surrounding the molecules, the statistics is computed by partial least square regression analysis and the output is displayed as contours superimposed on the molecules. The CoMFA methodology assumes that a suitable sampling of steric and electrostatic fields around a set of aligned molecules provides all the information necessary for understanding their biological properties. Apart from the CoMFA fields, CoMSIA [17] has additional field values for hydrophobic and hydrogen bonding interactions. These additional fields are of an aid in providing a further accounting of drug receptor interactions especially so in the contour displays.

In recent years hosts of ligands that possess diverse structural scaffolds such as staurosporins [18], flavonoids [19], indigoids [20], paullones [21], purines [22] etc. have been reported to inhibit CDK2.

Fig. (**1**) shows some examples of from these inhibitors. One of the main bottlenecks in developing a kinase inhibitor drug is the difficulty to attain selectivity and high affinity which mainly arises from the diverse nature of the substrates these enzymes have and the common mechanism the various kinases share among themselves. Although a variety of compounds have been known thus far, the non-selectivity and/or low affinity problem appears to impede the discovery and development into a therapeutically useful drugs for disorders involving CDK2. With this in mind we developed the QSAR models of these ATP competitive CDK2 inhibitors in the anticipation of getting a QSAR model that would account for the observed quantitative differences in the inhibitory activities seen for this series and to capitalize upon the insight

*Address correspondence to this author at the Department of Pharmaceutical Chemistry, School of Pharmacy, Addis Ababa University, P.O.Box 1176, Addis Ababa, Ethiopia;
E-mail: dnigus@phar.aau.edu.et / nigusd96@yahoo.com

1573-4064/08 $55.00+.00

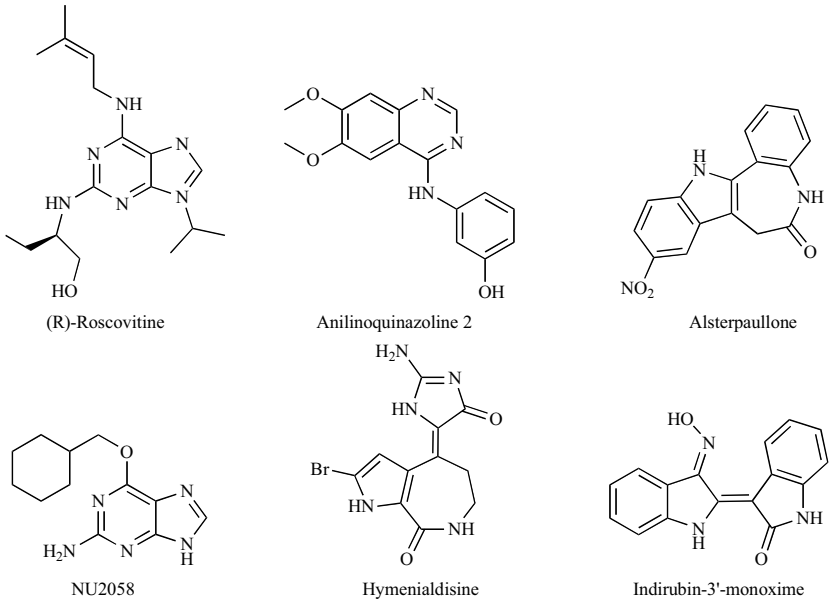


Fig. (1). Some promising examples of CDK2 inhibitors.

from the structural requirements to design ligands with pronounced inhibitory potency. This paper reports both the CoMFA and CoMSIA study results on benzodipyrazole series as CDK2 inhibitors.

2. COMPUTATIONAL DETAILS

2.1. Dataset for Analysis

The *in vitro* biological activity data reported as IC_{50} for inhibition of CDK2 by the benzodipyrazole series [23] was used for the current study. In the reference report 72 compounds were reported with their corresponding inhibitory activities expressed in IC_{50} values for kinase inhibition. Out of these, one molecule was found not to have bioactivity in exact numerical form for the inhibition of the enzyme and was subsequently removed from the analysis. Following this 71 molecules were left for the current study. As the log-dose response curve is linear about its middle region and to ensure symmetrical distribution of the biological data, the reported IC_{50} values were converted into the corresponding pIC_{50} using the formula.

$$pIC_{50} = -\log IC_{50}$$

2.2. Molecular Modeling

All molecular modeling studies were performed using the molecular modeling package SYBYL6.9 [24] installed on a Silicon Graphics Fuel Work station. As the crystal structure of the complex of CDK2 with these inhibitors is not available, the most active molecule was subjected to systematic conformational search with 10^0 increments followed by geometry optimization using MOPAC [25] interfaced with SYBYL. To further obtain a binding mode, the lowest energy conformer was docked to the active site of CDK2 using the FlexX docking algorithim [26]. The conformer with the best FlexX score was used as a template to construct the rest of the molecules which were further optimized by MOPAC. The AM1 Hamiltonian was used during energy minimization and Mulliken partial charges were computed.

2.3. Molecular Alignment

One of the fundamental assumptions wherein 3D-QSAR studies are based is that a geometric parallelism should exist between the modeled structures and that of the bioactive conformation. The spatial alignment of compounds under study is thus one of the most sensitive and determining factors in obtaining a robust and meaningful model. In the present study the MOPAC geometry optimized structures were aligned on the template by the ALIGN DATABASE command in SYBYL using the maximum substructure that is common to all. Fig. (**2**) shows the aligned molecules.

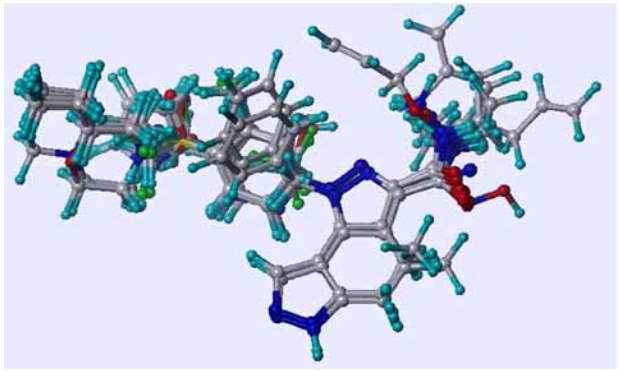


Fig. (2). Alignment of Molecules.

2.4. CoMFA and CoMSIA Interaction Energies

The steric and electrostatic CoMFA potential fields were calculated at each lattice intersection of a regularly spaced grid of 2.0Å. The grid box dimensions were determined automatically in such a way that the region boundaries were extended beyond 4Å in each direction from the co-ordinates of each molecule. The van dar Waals potential and Columbic terms, which represent steric and electrostatic fields, respectively, were calculated using the standard Tripos force fields.

A distance dependent dielectric constant of 1.00 was used. An sp^3 hybridized carbon atom with +1 charge served as probe atom to calculate steric and electrostatic fields. The steric and electrostatic contributions were truncated to +30.0 kcal/mol and electrostatic contributions were ignored at the lattice intersections with maximal steric interactions. In a similar vein, the steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor potentials were calculated at each lattice intersection of a regularly spaced grid of 2.0Å. A probe atom with radius of 1.0Å and +1 charge with hydrophobicity of +1 were used to calculate fields. All these descriptors contributions were truncated at 0.3 Kcal/mol. When the atoms of the molecules approach the probe very nearer, there will be a problem of sudden rise in energy. The GAUSSIAN type distance dependant function used by CoMSIA to calculate such properties are meant to overcome this problem.

2.5. Partial Least Square (PLS) Analysis

To quantify the relationship between the structural parameters (CoMFA and CoMSIA interaction energies) and the biological activities, the PLS [27] algorithm was used. The cross-validation analysis was performed using leave-one-out (LOO) method wherein one compound is removed from the dataset and its activity is predicted using the model derived from the rest of the dataset. The cross-validated r^2 that resulted in optimum number of components and lowest standard error of prediction was taken. To speed up the analysis and reduce noise, a minimum column filtering value (σ) of 2.00 Kcal/mol was used for the cross-validation. Final analysis (non-cross-validation) was performed to calculate conventional r^2 using the optimum number of component obtained from the initial leave one out cross validation analysis. To further assess the robustness and statistical confidence of the obtained models, bootstrapping analysis for 100 runs was performed.

2.6. Predictive Correlation Coefficient (r^2_{pred})

To further validate the derived models, biological activities of fifteen test set molecules were predicted using models derived from the training set. Predictive r^2 value was calculated using the formula:

$$r^2_{\text{predictive}} = (\text{SD} - \text{PRESS})/\text{SD}$$

where SD is the sum of squared deviation between the biological activities of the test set molecules and the mean activity of the training set molecules and PRESS is the sum of squared deviations between the actual and predicted activities of the test set molecules.

3. RESULTS AND DISCUSSION

The 3D-QSAR CoMFA and CoMSIA studies were carried out using 71 compounds from benzodipyrazole derivatives which are reported as CDK2 inhibitors. This was partitioned into a training set of 56 and a test set of 15 compounds at random with bias given to structural diversity in both the training set and the test set and so as to form the standard 4:1 training set to test set ratio for a QSAR study. Apart from segregating the molecules into test and training sets, all the procedures of geometry optimization, charge computation, alignments were done in identical way. Despite the ambiguity of drug-receptor interactions in general, a statistically robust models were obtained both from the CoMFA and CoMSIA study.

The CoMFA and CoMSIA PLS analysis is summarized in Table **1**. The cross-validated correlation co-efficient is used as a measure goodness of prediction whereas the non

Table 1. PLS Result Summary

Statistical Parameter	CoMFA	CoMSIA
R^2_{cv}	0.699	0.794
ONC	5	6
SEE	0.315	0.234
R^2	0.883	0.937
F_{value}	75.543	121.059
R^2_{bs}	0.932	0.958
SD	0.026	0.013
R^2_{pred}	0.754	0.815
Fraction of Field Contribution		
Steric	0.462	0.104
Electrostatic	0.538	0.229
Hydrophobic	-	0.314
Donor	-	0.243
Acceptor	-	0.109

R^2_{cv} =cross-validated correlation co-efficient; ONC = optimum number of components as determined by the PLS leave one out cross-validation study; SEE = standard error of estimate; R^2 = conventional correlation co-efficient; R^2_{pred} = predictive correlation co-efficient; R^2_{bs} = correlation coefficient after 100 runs of Bootstrapping: SD = standard deviation for 100 runs of bootstrapping;

cross-validated conventional correlation co-efficient indicates goodness of fit of a QSAR model. As can be seen from the body of the table, cross-validated correlations co-efficient of 0.699 and 0.794 were obtained respectively from CoMFA and CoMSIA PLS analysis. As a value of above 0.4 is usually considered significant, the r^2_{cv} obtained in both cases indicate a good internal predictive ability of both models. The models developed also exhibited a conventional correlation co-efficient of 0.883 for CoMFA and 0.937 for the CoMSIA study using optimum number of components of 5 for CoMFA and 6 for CoMSIA. In order to assess the external predictive ability of the models a randomly selected 15 compounds from the series were set aside during the model development to constitute the test sets. The predicted pIC_{50} obtained from this set of compounds, along with their actual pIC_{50} and the mean of the actual pIC_{50} value of the training set molecules, was used to compute the predictive correlation co-efficient which in turn served as measure of the external predictive ability of the QSAR models. In both cases a good figure was obtained for this parameter: 0.754 for CoMFA and 0.815 for CoMSIA study. Yet another way to further evaluate the statistical validity of a QSAR model is to carry-out a bootstrapping analysis [28]. Such analysis was done for 100 runs The higher r^2_{bs} value obtained after 100 runs of bootstrapping (0.932 for CoMFA and 0.958 for CoMSIA) further supports the statistical validity of the developed model and absence of chance correlation. When comparing the results, we found the CoMSIA method to provide a better quality model as compared to CoMFA. This is clearly noticeable from all the statistical parameters displayed in Table **1**. That CoMSIA did better than CoMFA is further supported by the lower residuals obtained both for the test set and training set molecules. This appears to come from the fact that the additional CoMSIA fields (hydrophobic, hydrogen bond donor and acceptor) are contributing predominantly to the ligand-receptor interaction for this series. This is seen from their fraction of field contribution and it is in agreement with the experimental fact that CDKs generally interact mainly *via* hydrophobic and 2-3 hydrogen bonds [29, 30].

The plots of actual versus predicted pIC_{50} is shown in Fig. (**3**) for CoMFA and Fig. (**4**) shows those of CoMSIA plots. The histogramms of residuals of test set molecules is shown in Fig. (**5**). Table **2** shows the structures and the corresponding actual and predicted pIC_{50} values for the training set inhibitors. Shown in Table **3** are the structures and the corresponding actual and predicted pIC_{50} values for the test set molecules.

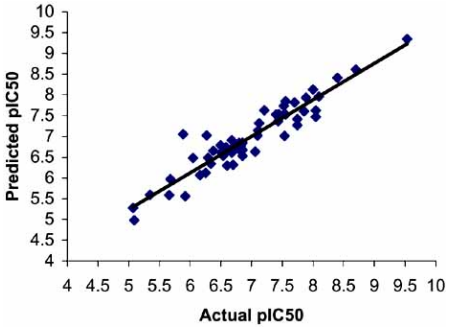


Fig. (3). Plot of actual versus predicted pIC_{50} values for the CoMFA model.

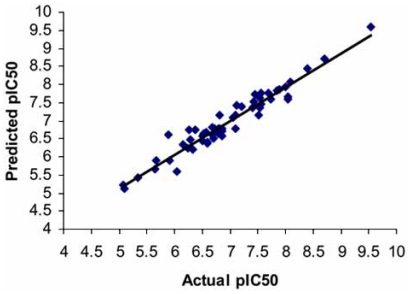


Fig. (4). Plot of Actual versus predicted pIC_{50} values for the CoMSIA model.

3.1. Contour Analysis

The QSAR produced by CoMFA and CoMSIA models is usually portrayed as three-dimensional coefficient contour maps [13, 17]. In general, the contour maps surround all lattice points where the QSAR is found to strongly associate changes in the molecular field values (which actually stand for structural changes) with changes in binding affinity or any other measure of biological property. More specifically, the polyhedra produced surround lattice points where the scalar products of the associated QSAR coefficient and the standard deviation of all values in the corresponding column of the data matrix are higher or lower than a user-specified value. In the standard SYBYL [22] setting, steric interactions are represented by green and yellow colored contours while electrostatic interactions are displayed as red and blue contours. Green contours stand for points where sterically bulkier groups are anticipated to increase the biological activity whereas the yellow contours are used to underscore the points where bulkier groups could lower the biological property. The electrostatic red plots show where the presence of a negative charge is expected to enhance the activity whereas the blue contours indicate where introducing or keeping positive charge is expected to better the observed activity. In addition to this, the hydrophobic interactions are displayed as white and yellow contours. White contours imply hydrophobically disfavored regions whereas the yellow maps indicate regions where increasing hydrophobicity is expected to enhance the biological property. The hydrogen bond donor field values are represented with cyan for preferred regions and purple for hydrogen bond disfavored regions. The hydrogen bond acceptor plots are portrayed as magneta and red

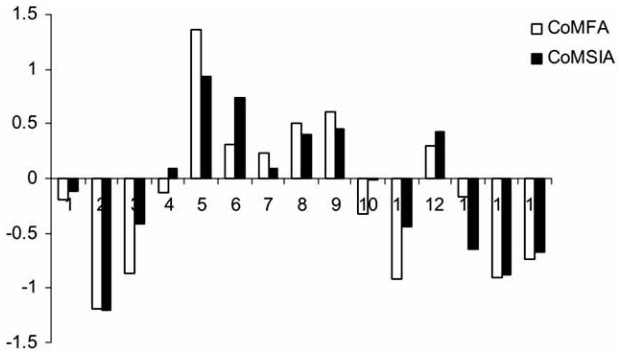
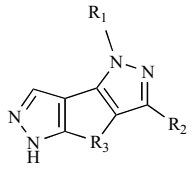


Fig. (5). Histograms of Residuals for Test set molecules.

Table 2. Structures and Actual Versus Predicted pIC_{50} Values for the Training set Compounds

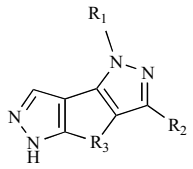


Sr. No.	Substituents			Actual	Predicted pIC_{50}		Residual	
	R1	R2	R3	pIC_{50}	CoMFA	CoMSIA	CoMFA	CoMSIA
1	-Ph-4-SO_2NH_2	-$CONH_2$	-CH_2-CH_2-	8.398	8.411	8.439	-0.013	-0.041
2	-Ph-4-SO_2NHMe	,,	,,	7.398	7.527	7.365	-0.129	0.033
3	-Ph-4-SO_2NMe_2	,,	,,	6.268	7.025	6.741	-0.757	-0.473
4	-Ph-4-SO_2NHBu	,,	,,	6.678	6.611	6.606	0.067	0.072
5	-Ph-4-SO_2Me	,,	,,	7.523	7.734	7.144	-0.211	0.379
6	-H	,,	,,	6.337	6.343	6.203	-0.006	0.134
7	-Ph	,,	,,	6.824	6.647	6.789	0.177	0.035
8	-Ph-4-Me	,,	,,	6.854	6.676	6.700	0.178	0.154
9	-Ph-4-OMe	,,	,,	6.854	6.533	6.563	0.321	0.291
10	-Ph-4-Cl	,,	,,	6.796	6.694	6.650	0.102	0.146
11	-Ph-4-F	,,	,,	6.569	6.622	6.677	-0.053	-0.108
12	-Ph-4-CF_3	,,	,,	7.097	7.020	6.780	0.077	0.317
13	-Ph-4-OCF_3	,,	,,	6.292	6.487	6.474	-0.195	-0.182
14	-Ph-4-CN	,,	,,	6.721	6.772	6.783	-0.051	-0.062
15	-Ph-4-N (CH_2-CH_2)$_2$O	,,	,,	6.252	6.123	6.252	0.129	0.000
16	-Ph-4-(2-imidazolo)	,,	,,	6.602	6.297	6.386	0.305	0.216
17	-Ph-3-Me	,,	,,	6.496	6.786	6.592	-0.290	-0.096
18	-Ph-3-F	,,	,,	6.377	6.654	6.742	-0.277	-0.365
19	-2-Pyridyl	,,	,,	6.585	6.741	6.413	-0.156	0.172
20	-Bn	,,	,,	5.658	5.587	5.651	0.071	0.007
21	-Me	,,	,,	6.538	6.537	6.640	0.001	-0.102
22	-CH_2-CF_3	,,	,,	7.699	7.819	7.748	-0.120	-0.049
23	-Ph-4-SO_2NH_2	,,	-CH =CH-	9.530	9.345	9.584	0.185	-0.054
24	-Ph-4-SO_2NMe_2	,,	,,	8.000	8.131	7.940	-0.131	0.060
25	-Ph-4-SO_2N (CH_2-CH_2)$_2$N-Me	,,	,,	6.824	6.765	7.138	0.059	-0.314
26	-H	,,	,,	7.108	7.145	7.168	-0.037	-0.06
27	-Ph	,,	,,	7.854	7.604	7.840	0.250	0.014
28	-Ph-4-OMe	,,	,,	8.046	7.473	7.589	0.573	0.457
29	-Ph-4-F	,,	,,	7.444	7.538	7.737	-0.094	-0.293
30	-Ph-4-CF_3	,,	,,	7.886	7.934	7.867	-0.048	0.019
31	-Ph-4-OCF_3	,,	,,	7.432	7.367	7.522	0.065	-0.090
32	-Ph-4-N (CH_2-CH_2)$_2$O	,,	,,	7.745	7.272	7.604	0.473	0.141
33	-Ph-4-(2-imidazolo)	,,	,,	7.125	7.319	7.430	-0.194	-0.305
34	-Ph-3-Me	,,	,,	8.046	7.626	7.653	0.420	0.393

(Table 2. Contd….)

Sr. No.	Substituents			Actual	Predicted pIC_{50}		Residual	
	R1	R2	R3	pIC_{50}	CoMFA	CoMSIA	CoMFA	CoMSIA
35	-Ph-3-Cl	,,	,,	7.537	7.546	7.618	-0.009	-0.081
36	-Ph-3-F	,,	,,	7.553	7.528	7.758	0.025	0.205
37	-2-Pyridyl	,,	,,	8.097	7.962	8.066	0.135	0.031
38	-3-Pyridyl	,,	,,	7.208	7.630	7.404	-0.422	-0.196
39	-Bn	,,	,,	6.854	6.849	6.772	0.005	0.082
40	-Me	,,	,,	7.745	7.419	7.590	0.326	0.155
41	$-CH_2-CF_3$	,,	,,	8.699	8.610	8.712	0.089	-0.013
42	$-CH_2-CH_2$-OH	,,	,,	7.553	7.848	7.454	-0.295	0.099
43	-Ph-4-OMe	-COOEt	$-CH_2-CH_2$	5.091	4.981	5.129	0.110	-0.038
44	,,	-COOH	,,	5.678	5.973	5.912	-0.295	-0.234
45	,,	-CONHOH	,,	6.161	6.068	6.350	0.093	-0.189
46	,,	$CONHOCH_2CH=CH_2$	,,	5.347	5.593	5.436	-0.246	-0.089
47	-Ph-4-SO_2NHBu	-COOH	,,	5.921	5.558	5.904	0.363	0.017
48	-Ph-4-Me	-COOEt	,,	5.071	5.278	5.236	-0.207	-0.165
49	,,	-COMe	,,	6.046	6.483	5.583	-0.437	0.463
50	-Ph-4-OMe	-COOH	-CH =CH-	6.678	6.910	6.798	-0.232	-0.120
51	,,	$-CONHNH_2$	,,	5.886	7.054	6.597	-1.168	-0.711
52	,,	-CONHOH	,,	7.537	7.013	7.356	0.524	0.181
53	,,	-C=N (OH) NH2	,,	6.509	6.610	6.455	-0.101	0.054
54	-Ph-4-SO_2NHBu	-CONHMe	,,	6.796	6.845	6.763	-0.049	0.033
55	,,	$CONHOCH_2CH=CH_2$	,,	6.699	6.318	6.498	0.381	0.201
56	-Ph-4-Me	-CONH	$-CH_2-C(CH_3)_2$	7.060	6.634	7.081	0.426	-0.021

Table 3. Structures and Actual Versus Predicted pIC_{50} Values for the Test set Compounds



Sr. No.	Substituents			Actual	Predicted pIC_{50}		Residuals	
	R1	R2	R3	pIC_{50}	CoMFA	CoMSIA	CoMFA	CoMSIA
57	-Ph-3-Cl	$-CONH_2$	$-CH_2-CH_2-$	6.444	6.634	6.560	-0.190	-0.116
58	-3-Pyridyl	,,	,,	5.796	6.992	7.003	-1.196	-1.207
59	$-CH_2-CH_2$-OH	,,	,,	5.939	6.778	6.352	-0.861	-0.413
60	-Ph-4-SO_2NHMe	,,	-CH =CH-	8.699	8.827	8.609	-0.128	0.090
61	-Ph-4-SO_2NHBu	,,	,,	8.699	7.335	7.765	1.364	0.934
62	-Ph-4-SO_2Me	,,	,,	9.000	8.686	8.269	0.314	0.731
63	-Ph-4-Me	,,	,,	7.824	7.589	7.736	0.235	0.088
64	-Ph-4-Cl	,,	,,	8.097	7.589	7.695	0.508	0.402
65	-Ph-4-CN	,,	,,	8.301	7.688	7.842	0.613	0.459

(Table 3. Contd….)

Sr. No.	Substituents			Actual	Predicted pIC_{50}		Residuals	
	R1	R2	R3	pIC_{50}	CoMFA	CoMSIA	CoMFA	CoMSIA
66	-Ph-4-OMe	$-CONHNH_2$	$-CH_2-CH_2-$	5.678	5.998	5.685	-0.320	-0.007
67	$-Ph-4-SO_2NHBu$	-COOEt	,,	5.000	5.924	5.441	-0.924	-0.441
68	,,	-CONHMe	,,	6.222	5.928	5.796	0.294	0.426
69	,,	$-CONHOCH_2CH=CH_2$	,,	5.545	5.719	6.190	-0.174	-0.645
70	-Ph-4-OMe	-COOEt	-CH=CH-	5.125	6.030	6.009	-0.905	-0.884
71	,,	-CN	,,	5.757	6.491	6.423	-0.734	-0.666

contours. The preferred regions for hydrogen acceptors are shown in the magneta plot whereas the red contour stands for areas not preferred for hydrogen bond acceptor.

The CoMFA Electrostatic contour map is shown in Fig. (**6**). The electrostatic contours show one medium sized and two small polyhedra around the sulphate group. This indicates where electron withdrawing groups are expected to enhance binding potency. In general compounds that have such groups in this region have a better experimentally determined inhibitory potency as compared to those that lack this functionality. This is exemplified in the higher inhibitory concentrations by compounds (**1**), (**2**) and (**5**) as against compounds (**8**) and (**9**). Apart from this, compounds that contain N attached to an electron withdrawing group are seen to own a better activity. This comes from the fact that in such compounds the nitrogen is oriented towards the extended blue polyhedron indicating its favorability to improve binding.

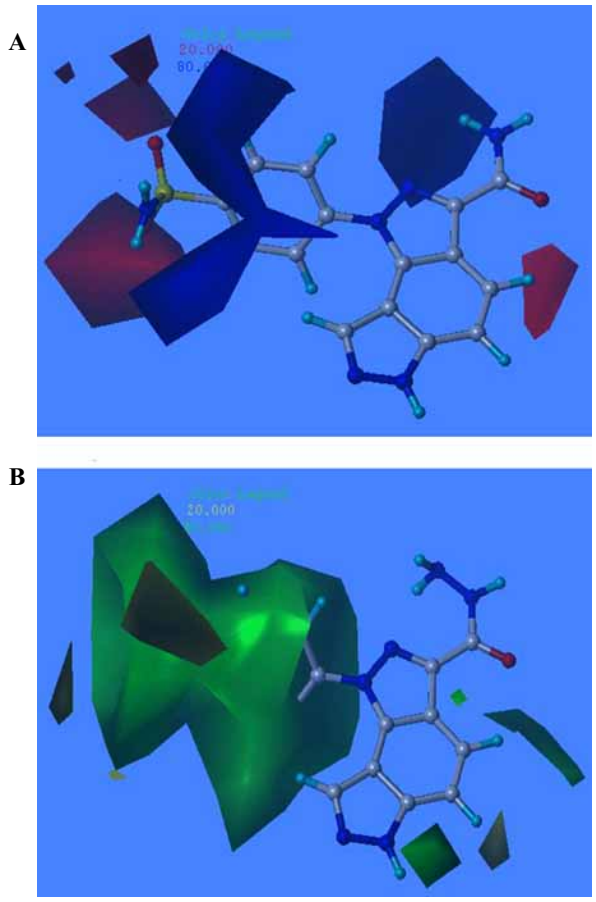


Fig. (6). CoMFA STDEVXCOEFF Electrostatic **(A)** and steric **(B)** contour maps. Compound 23 is shown in the background in both contours.

The CoMFA steric contours displayed in Fig. (**6B**) shows big green polyhedron around the phenyl substituent attached at R_1. This appears to explain why compounds (**8**) and (**9**) are better active than compound (**6**) and (**7**). The weak activity of compound (**4**) compared to compounds (**1**) and (**2**) comes from the fact that its n-butyl substituent is extended to the sterically forbidden yellow zone. This unfavorable interaction by the butyl attachment also appears to explain why compound (**46**) is weaker than compound (**45**) and why compound (**69**) is less active than compound (**68**)**.** The differences in the activity of compounds (**43**), (**9**) and (**44**) which contain respectively an ester, an amide and a carboxylic acid at R_2 appears to come from the fact that whereas the carbonyl group is oriented towards the red polyhedron in all cases, the N of the amide group is oriented to the blue plot accounting for its better activity. Compounds (**43**) and (**44**) have their oxygen oriented to the blue contours and thus

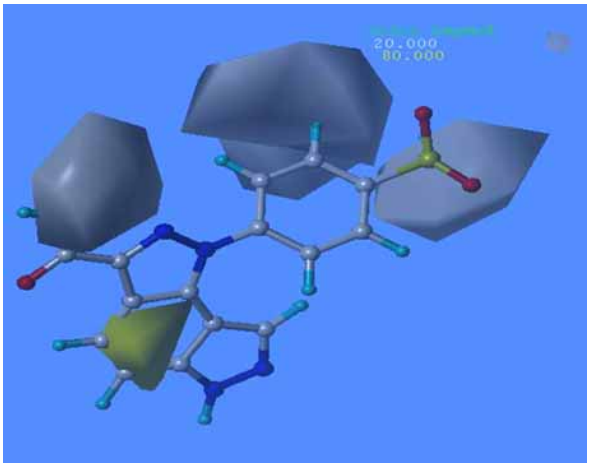


Fig. (7). STDEVXCOEFF hydrophobic contour map: Compound 23 is shown in the background.

making unfavorable interactions and this appear to contribute to their lower activities. The same fact seems to explain the lower activities of compounds (**50**) and (**70**) compared to compound (**28**) and compounds (**67**) and (**47**) as compared to compound (**4**).

The CoMSIA hydrophobic contour is displayed in Fig. (**7**). The plot shows two white polyhedrons enclosing the amino group of both $CONH_2$ and SO_2NH_2 groups. These regions indicate where hydrophobic groups are expected to reduce binding to CDK2 where as hydrophilic groups are anticipated to improve the binding affinity to CDK2. This fact is what explains why compound (**1**) more active than compound (**2**) which in turn is more active than compound (**4**) which itself is more active than compound (**3**). These compounds vary only by the substitution at the amino of the SO_2NH_2 group. Those that incorporate more hydrophobic groups at this region are less potent as expected than those that have hydrophilic attachments. The same fact explains why compound (**23**) is more active than compound (**24**) which in turn is more active than compound (**25**). In the same vein compounds with hydrophilic groups around the white polyhedron enclosing the NH_2 of $CONH_2$ are found to have a better activity in keeping with the contour display. This is seen in the poorer activity of compound (**43**) as compared to compounds (**9**), (**44**) and (**45**). The yellow hydrophobically preferred region is shown engulfing the R_3 group. Almost all compounds that have a CH=CH group at R_3 are found to be more active than those that have CH_2-CH_2 moiety which is in harmony with the displayed contour.

The hydrogen bond contours are shown in Fig. (**8A**) and Fig. (**8B**). The plots show a hydrogen bond acceptor favored magenta plot adjacent to the SO_2NH_2 group. This shows the need to have a hydrogen bond acceptor group to improve activity. A similar magenta plot is seen around the $CONH_2$ group. Also present is two red polyhedrons near the carbonyl and SO_2 groups indicating the need not to have acceptors in these regions to improve activity. The hydrogen bond donor plots depicted in Fig (**8A**) show two big purple polyhedrons surrounding the amino of $CONH_2$ and SO_2NH_2 groups which demands acceptor groups to increase activity along with a small cyan plot within the purple around the $CONH_2$. The QSAR study in both cases was able to reproduce the experimentally known facts indicating the potential utility of the developed models. Several important structural clues were gleaned from the study. We are currently employing the important clues to design more selective and potent inhibitors for this pharmaceutically important kinase.

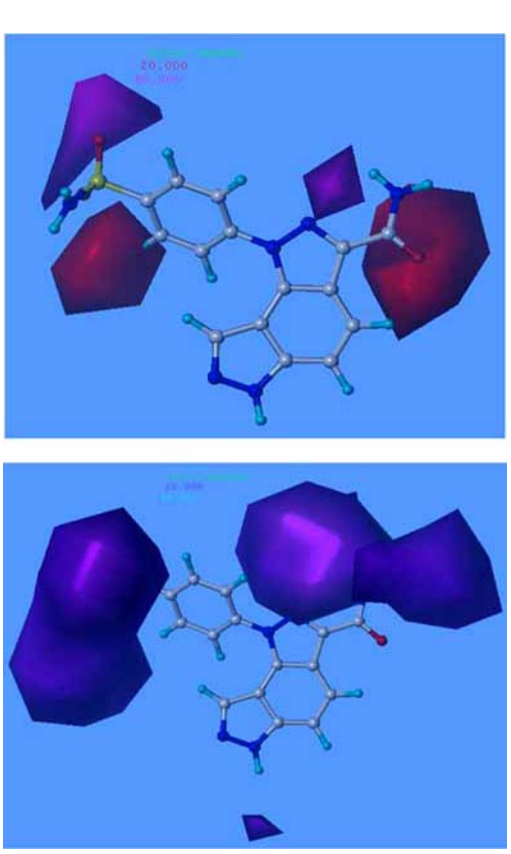


Fig. (8). STDEVXCOEFF Hydrogen bond acceptor **(A)** and Hydrogen bond donor **(B)** maps. Compound 23 is shown in the background in both cases.

CONCLUSION

A set of 71 compounds belonging to benzodipyrazole derivatives have been employed to develop correlation with the biological activities against CDK2. The validity of the correlation has been established using the PLS statistical parameters, the predictive power of the models and the contour map analysis. 3D-QSDAR study using both CoMFA and CoMSIA methodologies have been successfully employed to develop statistically robust models with good correlative and predictive ability for the inhibition of CDK2 by benzodipyarzole series. The analyses of contour plots for both models revealed useful insights about the structural requirements for the observed biological inhibitory activities. The developed models could be of great help in the rational design of compounds with an enhanced activity or to guide the prioritization of potential drug candidates for the purpose of synthesis and screening of chemical databases.

REFERENCES

[1] Refry, S.; Featherstone, J. *Nat. Rev. Drug Discov.*, **2002**, *1*, 175.
[2] Cohen, P. *Nat. Rev. Drug Discov.,* **2002**, *1*, 309.
[3] Noble, M.E.; Endicott, J.A.; Johnson, L.N. *Science*, **2004**, *303*, 1800.
[4] Adams, J.L.; Lee, D. *Curr. Opin. Drug Discovery Dev.,* **1999**, *2*, 96.
[5] Garcia-Echeverria, C.; Traxler, P.; Evans, D.B. *Med. Res. Rev.*, **2000**, *20,* 28.
[6] Sridhar, R.; Hanson-Painton, O.; Cooper, D.R. *Pharm. Res.,* **2000**, *17,* 1345.
[7] Dumas, J. *Exp. Opin. Ther. Patents,* **2001**, *11,*405.
[8] Rao, N.R. *Curr. Opin. Oncol*., **1996**, *8,* 516.
[9] Lingfei, K.; Pingzhang, Y.; Jianhua, G.; Yaowu, Z. *Cancer Lett.* **1998**, *130*, 93.
[10] Donnellan, R. Chetty, R. *FASEB J.,* **1999**, *13*,773.
[11] Meraldi, P.; Lukas, J.; Fry, A.M.; Bartek, J.; Nigg, E.A. *Nat. Cell Biol.* **1999**, *1,* 88.
[12] Marie, K.; Paul, G.; Laurent, M. *Trends Pharmacol. Sci.*, **2002**, *23,* 417.

[13] Cramer, R.D.III.; Patterson, D.E.; Bunce, J.D.*J. Am. Chem. Soc.,* 1**988**, *110*, 5959.
[14] Singh, S. K.; Dessalew, N.; Bharatam, P.V. *Med. Chem.,* **2007**, *3*, 75.
[15] Dessalew, N.; Patel, D.S.; Bharatam, P.V. *J. Mol. Graph. Model.,* **2007**, *25*, 885.
[16] Singh, S. K.; Dessalew, N.; Bharatam, P.V. *Eur. J. Med. Chem*., **2006**, *41*, 1310.
[17] Klebe, G.; Abraham, U.; Meitzner, T. *J. Med. Chem.* **1994**, *37,* 4130.
[18] Meijer, L.; Thunnissen, A.; White, A.; Granier, M.; Kiolic, M.; Tsai, L.H.; Walter, J.; Cleverly, K.E.; Salinas, P.C.; Wu, Y. Z.; Biernat, J.; Mandelkow, E. M.; Kim, S.; Pettit, G.R. *Chem. Biol.,* **2000**, *9,* 1143
[19] Sedlacek, H.H.; Czech, J.; Naik, R.; Kuar, G.; Worland, P.; Losiewiccz, M.; Parker, B.; Smith, A.; Senderowicz, A.; Sausville, E. *Int. J. Oncol.,* **1996**, *9,* 1143.
[20] Hoessel, R.; Leclerc, S.; Endicott, J.; Noble, M.; Lawrie, A.; Tunnah, P.; Leost, M.; Damiens, E.; Marie, D.; Marko, D.; Niederberger, E.; Tang, W.; Eisenbrand, G.; Meijer, L. *Nat. Cell Biol*., **1999**, *1*, 60.
[21] Zaharevitz, D.; Gussio, R.; Leost, M.; Senderowicz, A.; Lahusen, T.; Kunick, C.; Meijer, L.; Sausville, E.A. *Cancer Res.,* **1999**, *59*, 2566.
[22] Vesely, J.; Havlicek, L.; Strnad, M.; Blow, J.; Donella-Deana, A.; Pinna, L.; Letham, D.; Kato, J.; Detivaud, L.; Leclerc, S. *Eur. J. Biochem.,* **1994**, *224,* 771.
[23] Roberto, D.A.; Alberto, B.; Suzanne, M.; Gabriella, B.M.; Alexander, C.; Antonella, E.; Aurelio, M.; Paolo, P.; Fulva, R.; Marcellino, T.; Micheal, L.V.; Anna, V. *Bioorg. Med. Chem. Lett*., **2005**, *15*, 1315.
[24] SYBYL6.9: Tripos Inc., 1699 South Hanley Rd., St. Louis, MO 63144 USA.
[25] Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. *J. Am. Chem. Soc*., **1985**, *107*, 3902.
[26] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.,* **1996**, *261*, 470.
[27] Wold, S.; Johansson, A.; Cochi, M. In *QSAR in Drug Design: Theory, Methods and Applications;* Hugo Kubyini, Ed.; ESCOM, Lieden, **1993**, pp. 523-550 .
[28] Cramer, R.D. III.; Bunce, J.D.; Patterson, D.E. *Quant. Struct. Act. Relat.,* **1988**, *7*, 18.
[29] Dessalew, N.; Bharatam, P.V. *Eur. J. Med. Chem*., **2007**, *42*, 1014.
[30] Meijer, L.; Flajolet, M.; Greengard, P. *Trends Pharmacol. Sci*., **2004**, *25*, 471.

Received: 14 December, 2007 **Revised: 24 April, 2008** **Accepted: 25 April, 2008**