Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 41 (2006) 1310–1319
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://france.elsevier.com/direct/ejmech

Original article

3D-QSAR CoMFA study on indenopyrazole derivatives as cyclin dependent kinase 4 (CDK4) and cyclin dependent kinase 2 (CDK2) inhibitors

S.K. Singh[a,*], N. Dessalew[b], P.V. Bharatam[b]

[a] *Pharmacoinformatics Division, National Institute of Pharmaceutical Education and Research (NIPER), Sector 67, S.A.S. Nagar, Mohali 160062, Punjab, India*
[b] *Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research (NIPER), Sector 67, S.A.S. Nagar, Mohali 160062, Punjab, India*

Received in revised form 10 March 2006; accepted 23 June 2006
Available online 04 August 2006

Abstract

Cyclin dependent kinases (CDKs) have appeared as an important drug targets over the years with diverse therapeutic potentials. With the objective of designing new chemical entities with enhanced inhibitory potencies against CDK 2 (CDK2) and CDK 4 (CDK4), the 3D-QSAR CoMFA study carried out on indenopyrazole derivatives as inhibitors of these kinases is presented here. The developed model showed a strong correlative and predictive capability having a cross validated correlation co-efficient of 0.747 for CDK4 and 0.755 for CDK2 inhibitions. The conventional and predictive correlation co-efficients were, respectively, found to be 0.913 and 0.760 for CDK4, 0.941 and 0.765 for CDK2. The models could be employed to design ligands with enhanced inhibitory potencies and/or to predict the potencies of analogues to guide synthesis.

Keywords: CoMFA; CDK2; CDK4; Indenopyrazoles

1. Introduction

Protein phosphorylation is an important process in the control of protein functions. Biological phosphorylation occurs on serine, threonine and tyrosine residues and is catalyzed by protein kinases whose number transcends 800 in the human genome. Given the importance of protein phosphorylation as a main post-translational mechanism used by cells to regulate enzymes and other proteins and the fact that protein hyperphosphorylation is either the cause or consequence of many maladies [1], kinases have increasingly become important targets and the hunt for kinase inhibitors has attracted a great deal of attention in drug discovery over the years [2–7].

Cyclin dependent kinases (CDKs) have been characterized extensively in the past two decades. CDKs belong to the family of serine/threoinine kinases that play a key role in the regulation of the complex processes of the cell division cycle, apoptosis, transcription, and differentiation [8,9]. Thus far, members of this class is identified to include over nine CDKs. CDKs are inactive as monomers and activation requires binding to the corresponding regulatory proteins (cyclins) 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 the cell cycle. CDK 2 (CDK2) is required to complete G1 and to trigger the S phase. CDK4 is required to integrate extracellular signals and directs the cell cycle engine according the cell's environment. Because of their involvement in the key regulatory processes of the cell cycle and deregulation of CDKs in various disorders, CDK inhibitors are known to have a wide spectrum of applications ranging from protozoan infections (malaria, leishmania, trypanosomiasis), viral infections (HCMV, HSV, HIV, HPV), reproduction disorders, cardiovascular diseases (atheroscelorosis, restenosis, cardiac hypertrophy), gulomerulonephritis, cancers to nervous system diseases (Alzheimer's disease, stroke, amyotrophic disease, drug abuse) [10].

Since its introduction in 1988, comparative molecular field analysis (CoMFA) [11] has emerged as one of the most powerful tools in ligand based drug design strategies [12]. It has a

* Corresponding author. Tel.: +91 052 2459 141; mobile: +91 944 2449 150.
E-mail address: sanjeevsky@hotmail.com (S.K. Singh).

0223-5234/$ - see front matter © 2006 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2006.06.010

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 (PLS) 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.

Although a number of CDK2 and CDK4 inhibitors are reported thus far including staurosporins [13], flavonoids [14], indigoids [15], paullones [16] and purines [17], none of them have progressed into a clinically useful drug. Fig. 1 shows some inhibitors of these kinases. One of the main bottlenecks hampering the developing a kinase inhibitor drug is the difficulty to attain selectivity. This appears to stem from the diverse nature of the kinase substrates and the common mechanism these enzymes share among themselves. CoMFA is generally employed to enhance the binding affinity. CoMFA has recently been used to design selective GSK-3 inhibitors by Lescot et al. [18]. CoMFA along with docking study has also been employed to examine the structure–activity and structure–selectivity correlation of cyclic guanine derivatives as phosphodiesterase-5 inhibitors by Yang et al. [19]. Iskander et al. [20] have used CoMFA to optimize the pharmacophore model for 5-HT4 agonists. The B-DNA recognition of minor groove binders has been studied employing CoMFA by Oliveira et al. [21]. These all attest the usefulness of such a methodology in understanding the pharmacological properties of a given series. The CoMFA methodology especially when used in a comparative investigation for compounds acting on more than one target is valuable in pinpointing the structural basis of the observed quantitative differences in their pharmacotoxicological properties. Such insights are of an aid to design a new entity with a biased selectivity to the required receptor.

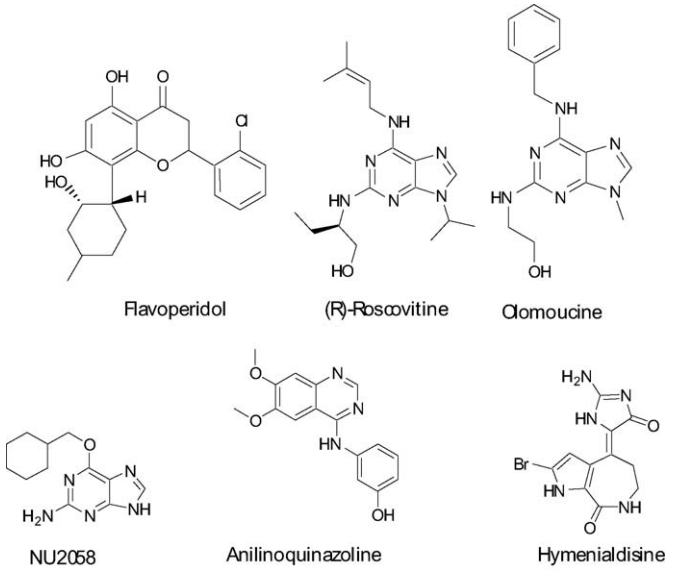


Fig. 1. Some examples of CDK inhibitors.

With this in mind we developed the QSAR CoMFA models of these ATP competitive CDK2 and CDK4 inhibitors in the anticipation of getting a model that would account for the biological activity seen in this series and to capitalize upon the insights to design ligands with pronounced inhibitory potency. This paper reports the CoMFA study results on indenopyrazole derivatives with the intention of designing potential leads with a higher inhibitory and discriminatory activity against these enzymes.

2. Computational details

2.1. Dataset for analysis

The in vitro biological activity data reported as IC_{50} for inhibition of CDK2 and CDK4 by the indenopyrazole derivatives [22–26] was used for the current study. All the molecules were obtained from sources by the same research group reported at different times. The in vitro assays employed cell lysates from insect cells expressing CDK2 and CDK4 and their corresponding cyclins. The compounds were evaluated for their inhibitory activity using ^{32}P labeled ATP. Those molecules which do not have biological activity for inhibition of the enzyme under study in exact numerical form were excluded from the analysis. As biological data are generally skewed, the reported IC_{50} values were converted into the corresponding pIC_{50} using the following formula:

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

2.2. Molecular modeling

All molecular modeling studies were performed using the molecular modeling package SYBYL6.9 [27] 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 (compound **108**) was docked into the active sites of CDK2 and CDK4 using the FlexX docking algorithm [28]. The conformer with the highest total FlexX score was taken. The maximum substructure of **108** that is common to each molecule in the series was used as a template to build their 3D structures. The AM1 Hamiltonian was used during energy minimization for all molecules and MOPAC [29] partial charges were computed.

2.3. Molecular alignment

One of the fundamental assumptions wherein 3D-QSAR studies are based is that a geometric similarity 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 [11]. In the present study the MOPAC geometry optimized structures were aligned on the template(**108**) employing the divide and conquer strategy [30] as follows: for compounds in Table 2 the indenopyr-

Table 1
CoMFA PLS result summary

QSAR parameters	CDK4	CDK2
r^2_{cv}	0.747	0.755
ONC	5	6
S.E.E.	0.315	0.191
r^2	0.913	0.941
F_{value}	169.929	212.287
r^2_{LFO}	0.740	0.775
r^2_{bs}	0.941	0.961
S.D.	0.112	0.071
r^2_{pred}	0.760	0.765
Fraction of field contributions		
Steric	0.554	0.527
Electrostatic	0.446	0.473

r^2_{cv} = cross validated correlation co-efficient; ONC = optimum number of components as determined by the PLS leave one out cross-validation study; S.E.E. = standard error of estimate; r^2 = conventional correlation co-efficient; r^2_{LFO} = non-cross validated correlation co-efficient for the leave-five out analysis; r^2_{pred} = predictive correlation co-efficient; r^2_{bs} = correlation co-efficient after 100 runs of Bootstrapping: S.D. = standard deviation for 100 runs of bootstrapping.

azole ring and the attached amide moiety of the most active molecule obtained from the docking experiment was used as a template for alignment as it is the common maximum substructure in this group; for compounds **94**–**110** the indenopyrazole ring and the attached morpholinocarbamate of **108** was used as the template whereas the indenopyrazole ring alone was used to align compounds **111**–**114**, **116** and **117**. Each of these class of molecules was separately aligned by the ALIGN DATABASE command available in SYBYL using the maximum substructure common to the respective classes with the template. Compound **115** was aligned on the template using the ATOM-FIT command available in SYBYL6.9. Finally all aligned molecules were combined for the molecular field generation. Fig. 2 shows the alignment of the molecules.

2.4. CoMFA interaction energies

The steric and electrostatic COMFA potential fields were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. The grid box dimensions were determined automatically in such a way that the region boundaries were extended beyond 4 Å 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 field. 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.

2.5. PLS analysis

To quantify the relationship between the structural parameters (CoMFA interaction energies) and the biological activities, the PLS [31] 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. Equal weights for CoMFA were assigned to steric and electrostatic fields using CoMFA_STD scaling option. 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 components obtained from the leave one out cross-validation analysis. To further assess the robustness and statistical confidence of the obtained models, leave-five-out study and bootstrapping analysis [32] for 100 runs was performed.

2.6. Predictive correlation co-efficient (r^2_{pred})

The predictive ability of the 3D-QSAR models were determined from a set of thirty compounds that were excluded during model development. The optimization and alignment of these test sets were the same as that of the training set compounds as described above, and their activities were predicted using the model produced by the training set. The predictive correlation co-efficient (r^2_{pred}), based on the test set molecules, is computed using

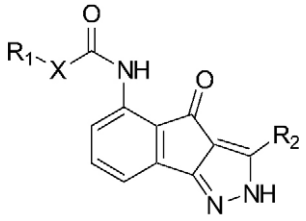
$$r^2_{pred} = (SD - PRES)/SD$$

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

3. Results and discussion

The 3D-QSAR CoMFA studies were carried out using indenopyrazole derivatives which are reported as CDK2 and CDK4 inhibitors by the same group over different times. Apart from molecules which do not have bioactivity in exact numerical form for the inhibition, those which lack bioactivity for both enzymes were removed from the analysis. This was done to aid in the comparative investigation about the structural requirements for interaction with the respective kinases. Following this 119 molecules were left for the current study. During the processes of model development and validation, we found two molecules not to fit to either the training set or test set for both enzyme inhibition. Outliers generally exist when they posses a unique scaffold and hence act on a different receptor or when they act on a different binding site of the same receptor or because of the limitations on the quality of the biological data. But the structures of these molecules are not that unique to claim that they bind differently. These too were removed and 117 molecules were remained for our study. This was partitioned into a training set of 87 and a test set of 30 compounds at random with bias given to both chemical and biological

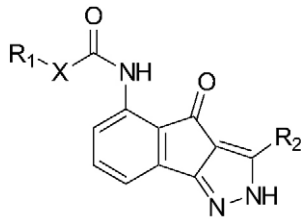
Table 2
Structures and actual versus predicted pIC_{50} of compounds **1**–**93**



	Substituents			CDK4			CDK2		
Serial number	R_1	X	R_2	Actual pIC_{50}	Predicted pIC_{50}	Residual	Actual pIC_{50}	Predicted pIC_{50}	Residual
1	$(CH_3)2$	CH	–Ph-4-OMe	4.456	4.407	0.049	5.854	5.653	0.201
2	$(CH_3)3$	C	–Ph-4-OMe	4.347	4.651	–0.304	5.678	6.004	–0.326
3	$(Me)_2N$-	CH_2	–Ph-4-OMe	6.046	6.265	–0.219	7.356	7.237	0.119
4t	Morpholine-4-yl	CH_2	–Ph-4-OMe	6.710	6.873	–0.163	7.678	7.516	0.162
5	Piperazin-1-yl	CH_2	–Ph-4-OMe	6.947	6.720	0.227	7.481	7.342	0.139
6	Ethyl NH–	CH_2	–Ph-4-OMe	6.038	6.791	–0.753	6.886	7.261	–0.375
7	N-methyl piperazine	CH_2	–Ph-4-OMe	6.903	7.065	–0.162	7.921	7.789	0.132
8	4-Aminomethylpiperidine	CH_2	–Ph-4-OMe	7.699	7.329	0.370	7.921	7.874	0.047
9t	4-Amidopiperidine	CH_2	–Ph-4-OMe	7.119	7.177	–0.058	8.097	7.824	0.273
10	4-Hydroxylmethylpiperidine	CH_2	–Ph-4-OMe	7.086	7.196	–0.11	7.959	7.826	0.133
11	4-Amidopiperazine	CH_2	–Ph-4-OMe	7.194	7.181	0.013	8.097	8.269	–0.172
12	4-Amidinopiperazine	CH_2	–Ph-4-OMe	7.585	7.337	0.248	8.155	8.029	0.126
13	H	CH_2	–Ph-4-OMe	6.347	5.760	0.587	6.569	6.327	0.242
14t	Benzyl	NH	–Ph-4-OMe	6.512	7.341	–0.829	7.022	7.211	–0.189
15	Phenyl	NH	–Ph-4-OMe	6.057	6.735	–0.678	7.495	7.444	0.051
16	n-butyl	NH	–Ph-4-OMe	6.108	6.698	–0.59	7.076	7.125	–0.049
17t	$(Me)_2N$	NH	–Ph-4-OMe	7.678	7.373	0.305	8.301	7.549	0.752
18	4-methylpiperazine	NH	–Ph-4-OMe	8.046	8.172	–0.126	7.921	7.789	0.132
19	Morpholine-4-yl	NH	–Ph-4-OMe	7.921	7.827	0.094	7.745	7.691	0.054
20	Piperidin-1-yl	NH	–Ph-4-OMe	7.921	7.893	0.028	7.657	7.742	–0.085
21	Pyrrolidine-1-yl	NH	–Ph-4-OMe	7.921	7.542	0.379	7.824	7.617	0.207
22	H	CH_2	–Ph	6.076	6.102	–0.026	6.620	6.749	–0.129
23	H	CH_2	–Ph-4-Me	6.310	6.258	0.052	6.796	6.791	0.005
24t	H	CH_2	–Ph-4-Et	6.194	6.090	0.104	6.553	6.563	–0.01
25	H	CH_2	–Ph-4-n-Pr	5.721	6.161	–0.44	6.310	6.778	–0.468
26	H	CH_2	–Ph-4-OH	5.638	5.866	–0.228	6.276	6.625	–0.349
27t	–Ph-4-NH2	CH_2	–Ph-4-OMe	6.319	5.955	0.364	7.420	6.849	0.571
28t	H	CH_2	–Ph-4-NMe_2	6.509	6.416	0.093	6.495	6.662	–0.167
29	H	CH_2	–Ph-4piperidino	6.167	6.196	–0.029	6.046	6.245	–0.199
30t	H	CH_2	–Ph-4-morpholino	6.065	6.252	–0.187	6.432	6.423	0.009
31	H	CH_2	–Ph-4-SMe	6.468	6.107	0.361	6.921	6.647	0.274
32	Morpholino	CH_2	–Ph-4-NMe_2	6.959	7.125	–0.166	7.444	7.319	0.125
33	4-(OH)piperidine-1-yl	CH_2	–Ph-4-NMe_2	7.301	7.422	–0.121	7.481	7.472	0.009
34	4-(Aminomethyl) piperidin-1-yl	CH_2	–Ph-4-NMe_2	8.155	7.680	0.475	7.824	7.653	0.171
35t	N-methylpiperazin-1-yl	CH_2	–Ph-4-NMe_2	7.260	7.465	–0.205	7.538	7.529	0.009
36	Morpholino	CH_2	–Ph-4-morpholino	6.921	7.173	–0.252	7.432	7.327	0.105
37	4-(OH) piperidine-1-yl	CH_2	–Ph-4-morpholino	6.959	7.394	–0.435	7.070	7.431	–0.361
38t	4-(Aminomethyl) piperidin-1-yl	CH_2	–Ph-4-morpholino	7.745	7.631	0.114	7.585	7.568	0.017
39	H	NH	3-thienyl	7.174	7.209	–0.035	8.046	7.925	0.121
40t	N-methylpiperazin-1-yl	CH_2	–Ph-4-morpholino	6.721	6.734	–0.013	7.377	7.362	0.015
41t	4-(Aminomethyl) piperidin-1-yl	CH_2	Et	5.886	6.696	–0.81	6.620	7.454	–0.834
42	4-(Aminomethyl) piperidin-1-yl	CH_2	Cyclopropyl	6.366	6.653	–0.287	7.260	7.367	–0.107
43	4-(Aminomethyl) piperidin-1-yl	CH_2	Cyclohexane	6.444	6.264	0.18	7.018	7.158	–0.14
44t	H	NH	Cyclopropyl	5.854	6.395	–0.541	6.854	7.332	–0.478
45	H	CH_2	4-Pyridyl	5.886	5.852	0.034	7.119	6.697	0.422
46	H	CH_2	2-Thienyl	5.347	6.015	–0.668	6.745	6.782	–0.037
47	H	NH	2-Thienyl	6.824	6.478	0.346	7.959	7.555	0.404
48	H	NH	2-Thienyl,3-OMe	6.456	6.397	0.059	7.769	7.580	0.189
49t	H	NH	2-Thienyl,5-Me	7.276	6.544	0.732	7.886	7.401	0.485
50	H	NH	2-Furanyl	6.076	6.151	–0.075	7.065	7.175	–0.11
51t	H	NH	2-Thienyl,5-CO_2Et	6.523	6.634	–0.111	6.886	7.161	–0.275
52	H	NH	3-Thienyl,5-Cl	7.244	7.112	0.132	7.886	7.862	0.024

(continued)

Table 2 *(continued)*



	Substituents			CDK4			CDK2		
Serial number	R_1	X	R_2	Actual pIC_{50}	Predicted pIC_{50}	Residual	Actual pIC_{50}	Predicted pIC_{50}	Residual
53	H	NH	3-Pyrrolyl,1-Me	6.886	6.183	0.703	7.585	7.722	–0.137
54	Dimethylamino	NH	2-Thienyl	6.468	6.513	–0.045	7.444	7.510	–0.066
55t	Dimethylamino	NH	5-(OMe) thien-2-yl	7.076	6.689	0.387	7.495	7.519	–0.024
56	Dimethylamino	NH	5-(Me) thien-2-yl	7.091	6.876	0.215	7.602	7.574	0.028
57	Dimethylamino	NH	5-(CO_2EtMe) thien-2-yl	6.733	6.549	0.184	7.553	7.467	0.086
58	Dimethylamino	NH	3-Thienyl	6.971	7.079	–0.108	7.620	7.899	–0.279
59	Dimethylamino	NH	5-(Cl) thien-3-yl	7.678	6.953	0.725	8.155	7.91	0.245
60t	Dimethylamino	NH	2,5-(di-Me) thien-3-yl	6.288	6.946	–0.658	7.585	7.711	–0.126
61	Dimethylamino	NH	Furan-2-yl	6.197	6.679	–0.482	7.585	7.846	–0.261
62t	Dimethylamino	NH	2,4-(di-Me)thiazol-5-yl	6.824	6.971	–0.147	8.398	7.949	0.449
63	Morpholine-4-yl	NH	5-(Me) thien-2-yl	7.638	7.160	0.478	8.000	8.006	–0.006
64	Morpholine-4-yl	NH	5-(CO_2EtMe) thien-2-yl	7.523	7.192	0.331	7.509	7.491	0.018
65	Morpholine-4-yl	NH	5-(Cl) thien-3-yl	8.155	7.436	0.719	8.00	7.897	0.103
66	4-(Methyl)piperazin-1-yl	NH	5-(CO_2EtMe) thien-2-yl	7.523	7.097	0.426	7.444	7.521	–0.077
67t	4-(Aminomethyl) piperidin-1-yl	CH_2	Isopropyl	6.468	6.629	–0.161	7.319	7.441	–0.122
68	4-(Methyl)piperazin-1-yl	NH	2,5-(di-Me) thien-3-yl	7.046	7.304	–0.258	7.921	7.906	0.015
69	4-(Methyl)piperazin-1-yl	NH	2,4-(di-Me)thiazol-5-yl	7.337	7.348	–0.011	8.097	8.096	0.001
70	$(Me)_2$CHCONH–	NH	–Ph-4-OMe	8.187	7.770	0.417	7.721	7.724	–0.003
71	4-(OH)Ph(CH2)2CONH–-	NH	–Ph-4-OMe	8.046	7.890	0.156	7.886	7.875	0.011
72	4-(OMe)PhCONH–	NH	–Ph-4-OMe	7.959	8.068	–0.109	7.824	7.904	–0.08
73	3-(NO_2)PhCONH–	NH	–Ph-4-OMe	8.000	7.900	0.1	7.620	7.759	–0.139
74	3,4,5-(tri-OMe)PhCONH–	NH	–Ph-4-OMe	8.155	7.862	0.293	7.699	7.680	0.019
75	3-(Me)PhCONH–	NH	–Ph-4-OMe	8.155	8.156	–0.001	7.886	7.889	–0.003
76t	3,4-(di-OMe)PhCONH–	NH	–Ph-4-OMe	7.620	7.749	–0.129	7.387	7.650	–0.263
77	(4-OH,3-NH2) PhCONH–	NH	–Ph-4-OMe	8.155	8.128	0.027	7.620	7.897	–0.277
78	2,5-(di-Cl)PhCONH–	NH	–Ph-4-OMe	7.553	7.677	–0.124	7.444	7.611	–0.167
79	3,4-(di-OH)PhCONH–	NH	–Ph-4-OMe	8.301	8.113	0.188	8.000	7.955	0.045
80	3,5-(di-NH2)PhCONH–	NH	–Ph-4-OMe	7.796	8.044	–0.248	7.409	7.742	–0.333
81	MeOCONH–	NH	–Ph-4-OMe	8.097	7.671	0.426	7.796	7.690	0.106
82t	2-(OH)PhCONH–	NH	–Ph-4-OMe	7.222	7.739	–0.517	8.046	7.632	0.414
83	Naphthalen-2-yl CONH–	NH	–Ph-4-OMe	7.886	8.044	–0.158	7.456	7.457	–0.001
84	BnCONH–	NH	–Ph-4-OMe	7.569	7.869	–0.3	7.481	7.530	–0.049
85	PhCONH–	NH	–Ph-4-OMe	8.155	8.066	0.089	7.921	7.917	0.004
86	4-pyrridilCONH–	NH	–Ph-4-OMe	8.046	7.953	0.093	7.921	7.895	0.026
87	3-pyrridylCONH–	NH	–Ph-4-OMe	8.046	7.935	0.111	7.796	7.873	–0.077
88	MeCONH–	NH	–Ph-4-OMe	7.244	7.037	0.207	7.432	7.419	0.013
89t	4-(OH)PHCONH–	NH	–Ph-4-OMe	8.301	7.664	0.637	8.046	7.590	0.456
90	H_2NCOCONH–	NH	–Ph-4-OMe	6.620	6.822	–0.202	7.013	7.052	–0.039
91	3-(NH_2)PhCONH–	NH	–Ph-4-OMe	8.155	7.748	0.407	7.886	7.582	0.304
92t	2,4-(di-OH)PhCONH–	NH	–Ph-4-OMe	7.921	7.740	0.181	7.796	7.648	0.148
93	4-NH_2PhCONH–	NH	–Ph-4-OMe	8.000	7.770	0.23	8.000	7.700	0.300
OL1	H	NH	–Ph-4-OMe	7.180	5.343	1.837	8.155	6.274	1.881
OL2	4-Picolyly	NH	–Ph-4-OMe	5.939	7.677	–1.738	7.539	7.475	0.064

Compound number with "t" refers to those compounds included in the test set. OL1 and OL2 refer to the two outlier molecules.

diversity in both the training set and the test set molecules. Despite the ambiguity of drug–receptor interactions in general, a statistically robust models were obtained from the CoMFA study for both kinases.

The CoMFA PLS analysis is summarized in Table 1. The cross validated correlation co-efficient is used as a measure of goodness of prediction whereas the conventional correlation co-efficient indicates goodness of fit of a QSAR model. The F_{value} stands for the degree of statistical confidence on the developed model. As can be seen from the body of the table, a cross validated correlation co-efficient of 0.747 for CDK4 and 0.755 for CDK2 were obtained using, respectively, five and six optimum numbers of components. The r^2_{cv} obtained in both cases indicates a good internal predictive ability of

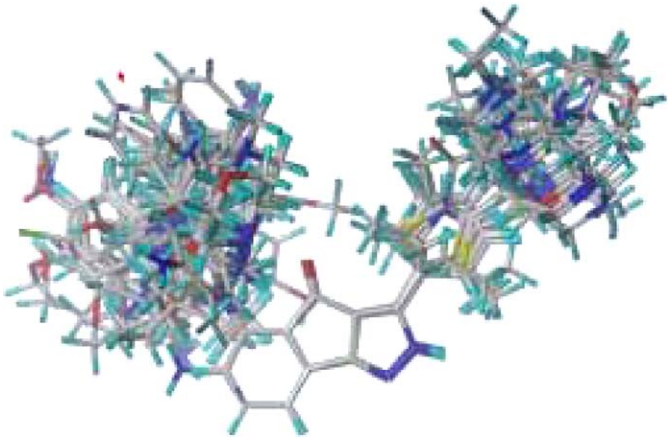


Fig. 2. Alignment of all molecules used for molecular field generation.

the models. The models developed also exhibited a good non-cross validated correlation co-efficient of 0.913 for CDK4 and 0.941 for CDK2. Test sets are generally used to evaluate the external predictive capabilities of QSAR models. For this purpose, a randomly selected 30 compounds from the series were set-aside during model development. In both cases a predictive correlation co-efficient of 0.760 for CDK4 and 0.765 for CDK2 were obtained indicating good forecasting capabilities of the models. Yet another way to further evaluate the usefulness of the developed models is to test for statistical validity. To this end, both bootstrapping and leave-five-out analyses were performed. In both cases a higher correlation co-efficient (0.941 for CDK4 and 0.961 for CDK2) was obtained after 100 runs of bootstrapping. The non-cross validated correlation co-efficient from the leave-five-out study were 0.740 for CDK4 and 0.775 for the CDK2 model. The figures obtained for the different statistical parameters of CoMFA strongly indicate the statistical validity and stability of the developed models. The contributions of steric to electrostatic fields were found to be about 55:45 for CDK4 and 53:47 for CDK2.

The plots of actual versus predicted pIC_{50} values is shown in Fig. 3 for CDK4 and Fig. 4 shows such plot for CDK2 inhibitions. The histograms of residuals of test set molecules is shown in Fig. 5. Table 2 shows the structures and the corresponding actual and predicted pIC_{50} Tables 3 and 4 values for

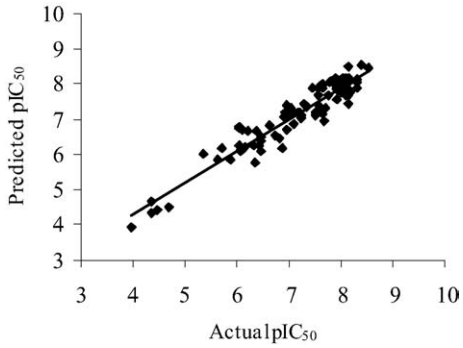


Fig. 3. Plot of actual versus predicted pIC_{50} values for the CDK4 model.

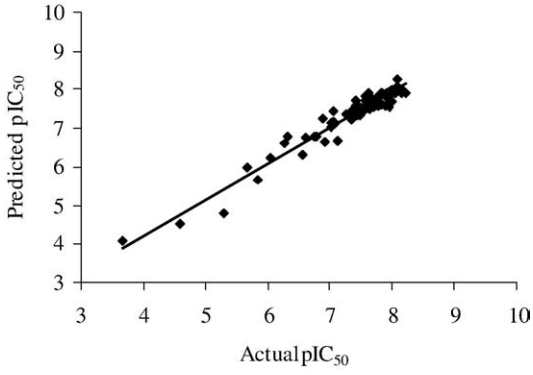


Fig. 4. Plot of actual versus predicted pIC_{50} values for the CDK2 model.

all the molecules. All the aligned molecules are shown in Fig. 2.

3.1. Contour analysis

The QSAR produced by a CoMFA model is usefully portrayed as three-dimensional co-efficient contour maps [11]. In general, the contour maps surround all lattice points where the QSAR is found to strongly associate changes in the molecular field values (which basically means changes in structure) 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

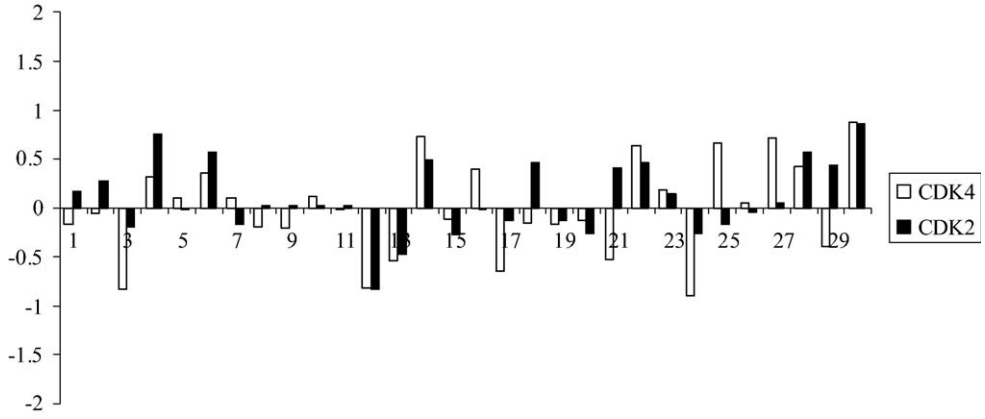
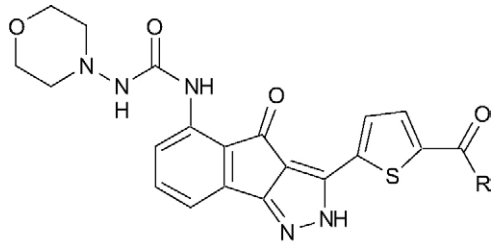


Fig. 5. Histograms of residuals (blue for CDK4 and red for CDK2) for test set molecules.

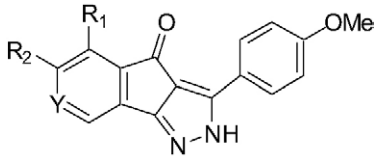
Table 3
Structures and actual versus predicted pIC_{50} of compounds **94–110**



Serial numbers	Substituents R	CDK4 Actual pIC_{50}	CDK4 Predicted pIC_{50}	CDK4 Residual	CDK2 Actual pIC_{50}	CDK2 Predicted pIC_{50}	CDK2 Residual
94	2-(Dimethylamino)ethylamino	7.824	8.071	–0.247	7.678	7.762	–0.084
95	2-(Pyrrolidin-1-yl)ethylamino	7.638	7.966	–0.328	7.602	7.574	0.028
96	2-(Piperidine-1-yl)ethylamino	7.656	7.987	–0.331	7.509	7.546	–0.037
97t	2-(Morpholin-4-yl)ethylamino	7.065	7.957	–0.892	7.292	7.551	–0.259
98	Piperidin-1-yl	7.656	7.869	–0.213	7.602	7.627	–0.025
19	3-(Dimethylamino) Piperidin-1-yl	7.886	8.154	–0.268	7.620	7.772	–0.152
100	4-(Dimethylamino) Piperidin-1-yl	8.097	8.159	–0.062	7.638	7.799	–0.161
101t	Piperazin-1-yl	8.398	7.729	0.669	7.420	7.584	–0.164
102	4-(Ethyl)piperazin-1-yl	8.301	8.048	0.253	7.721	7.544	0.177
103	3-(Amino)pyrrolidin-1-yl	7.456	7.865	–0.409	7.409	7.591	–0.182
104	3-(Methylamino)pyrrolidin-1-yl	8.301	7.866	0.435	7.959	7.941	0.018
105	3-(Dimethylamino)pyrrolidin-1-yl	7.921	7.917	0.004	7.356	7.441	–0.085
106	Azepan-1-yl	7.770	7.665	0.105	7.638	7.501	0.137
107t	4-(Methyl)piperazin-1-yl	8.000	7.958	0.042	7.495	7.550	–0.055
108	[1,4]Diazepan-1-yl	8.523	8.456	0.067	8.222	7.895	0.327
109	4-(Methyl)- [1,4]diazepan-1-yl	8.398	8.537	–0.139	7.824	7.919	–0.095
110	4-(Ethyl)- [1,4]diazepan-1-yl	8.155	8.504	–0.349	7.854	7.895	–0.041

Compound number with "t" refers to those compounds included in the test set.

Table 4
Structures and actual versus predicted pIC_{50} of compounds **110–117**



Serial numbers	Substituents R1	R2	Y	CDK4 Actual pIC_{50}	CDK4 Predicted pIC_{50}	CDK4 Residual	CDK2 Actual pIC_{50}	CDK2 Predicted pIC_{50}	CDK2 Residual
111	H	H	C	4.347	4.334	0.013	4.585	4.518	0.067
112t	Acetamide	H	C	6.337	5.625	0.712	6.292	6.247	0.045
113	H	NH_2	C	3.966	3.940	0.026	3.678	4.073	–0.395
114t	NH_2	H	C	4.638	4.215	0.423	5.076	4.502	0.574
115t	H	H	N	4.462	4.860	–0.398	5.328	4.885	0.443
s116	OH	H	C	4.699	4.504	0.195	5.284	4.798	0.486
117t	Formamido	H	C	6.699	5.813	0.886	7.097	6.241	0.856

Compound number with "t" refers to those compounds included in the test set.

co-efficient and the standard deviation of all values in the corresponding column of the data matrix are higher or lower than a user-specified value. According to the standard SYBYL 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 the Lennard–Jones potential has to be increased by appropriate groups to increase the biological activity whereas the yellow contours are used to underline the points where such a potential has to be decreased by suitable substituents to correlate with increased binding affinity. The electrostatic red plots show where the presence of a negative charge is expected to enhance the activity whereas the blue contours indicates regions where placing more positive charge is expected to correlate with increased binding affinity.

The electrostatic contour for CDK4 inhibitory model is displayed in Fig. 6. Compound **108** is displayed in all of the backgrounds to aid in visualization. The plot shows a big blue polyhedron near the azapine ring. This calls for an increase in positive charge near this region to improve activity. This fact

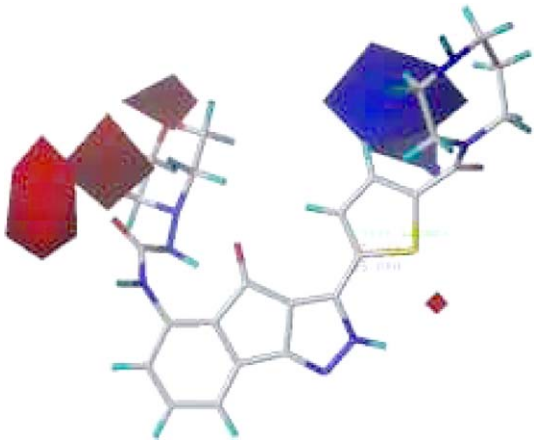


Fig. 6. CoMFA STDEVXCOEFF electrostatic contour map for CDK4 inhibition: red for negative charge favored region; blue for positive charge preferred region to improve binding affinity.

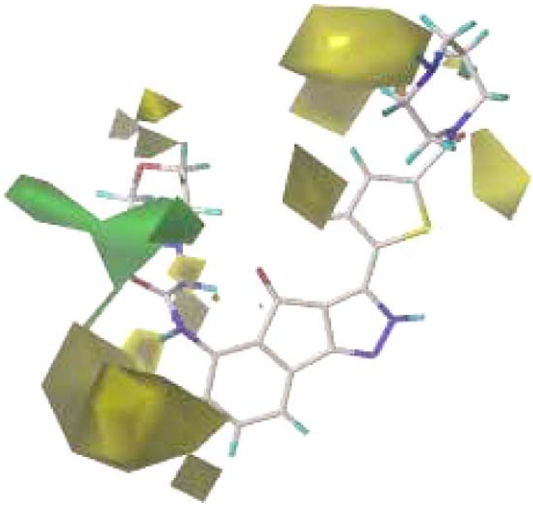


Fig. 7. CoMFA STDEVXCOEFF steric contour map for CDK4 inhibition: green indicates regions which prefer an increase in Lennard–Jones potential and the yellow map calls for a reduction of this potential to improve the affinity.

seems to explain why compound **98** is less active than compound **101**. The former has piperidine substituent whereas the latter has piperazine and hence making more favorable interaction at this plot that appears to account for its better inhibitory potency. Apart from this the better activities of compound **66** as compared to compound **57**, compound **68** as compared to **60** and **69** as against **62** appears to come from the same reason. In all this cases, better active molecules have a piperazine substituent at the blue plots and hence making favorable interaction that lead to their higher potencies. The above circumstances appear to arise due to the likelihood of the nitrogen getting charged under physiological conditions and hence increasing the positive charge. Furthermore, the higher inhibitory potency of **29** in comparison with **30** supports this observation. Compound **29** has a piperidine substituent whereas **30** has a morpholine ring. The oxygen of morpholine moiety is seen making untoward interaction with the blue contour and hence accounting for its lower activity. This is further substantiated by the better activity of compound **5j** which has dimethylamino group as compared to **22**(which has just H), **23** (Me), **24** (Et) and **25** (n-Pr). The contour also shows two medium sized red polyhedra near the carbonyl attached to the indene ring. This calls for an increase of negative charge around that region to improve activity. This is what actually explains for the better activity of compound **112** as compared to compounds **116** and **114**. In compound **112**, the carbonyl of the acetamide group is making the favorable interaction where such interaction is absent in cases of compounds **116** (OH) and **114** (NH_2).

The CoMFA steric contour map for CDK4 is shown in Fig. 7. This plot shows two medium sized yellow polyhedra near the diazepane ring and other two below these plots. This calls for reduction in Lennard–Jones potential on this area to improve affinity. This fact seems to explain the differences in activities of compound **109** and **110** compared to **108**. Compounds **109** and **110** have a methyl and ethyl attachment, respectively, that makes unfavorable interaction with the yellow plots in this region. Further the relatively lower activities of compound **105** relative to compound **104** appear to come from this unfavorable steric interaction. The dimethyl of **105** is oriented to the yellow plots and thus making more unfavorable steric interaction. Moreover, the lower activity of compound **60** in contrast to **58** comes from the reason that the methyl group of the former molecule is seen making unfavorable interactions with the yellow plot near the thienyl ring. That the four yellow polyhedra in this region impact binding affinity is further substantiated in the increasingly lower potencies of compounds **23**–**25**. This is because the n-Pr of **25** is making more negative interaction as compared to the ethyl group of **24** which in turn is making more negative interaction than the methyl of **23**.

The CoMFA electrostatic contour for CDK2 is displayed in Fig. 8. The major difference between the two electrostatic contours is the appearance of a blue plot near the morpholine ring and another in front the NH connecting the morpholine to the amide attached to the indene ring and the big reduction in the

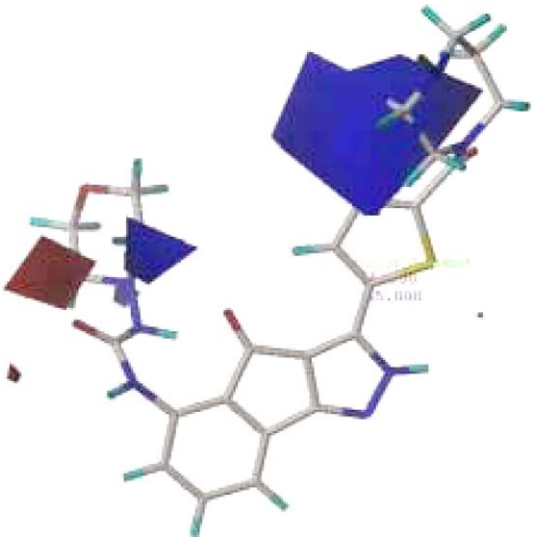


Fig. 8. CoMFA STDEVXCOEFF electrostatic contour map for CDK2 inhibition: red for negative charge favored regions; blue for positive charge preferred regions to improve binding affinity.

size of the two red plots near this area. These discrepancies could be employed to bias the selectivity to either receptor. The plot shows a big blue polyhedron engulfing parts of the diazepane ring where an increase in positive charge is expected to improve the binding affinity. Apart from the size difference, this plot is more or less the same as the one for CDK4 plot. And it does explain the higher activity of compound **108** as compared to compound **94** for in the former the two 'N' are interacting favorably in the blue plot. The same reason seems to account for the difference in activities of compound **109** versus compound **107**, compound **110** versus compound **102**. The piperazine ring in compounds **107** and **102** is seen to orient the carbonyl connecting it with thienyl ring to the blue plot while the same moiety is oriented away from the blue plot in the better active molecules. This apparently contributes for the decrease in the activity of piperazine containing compounds as compared to diazepane containing compounds. This blue plot is also important for the better activities of compound **45** as compared to compound **46**, compounds **32**–**35** as against compounds **36**–**39**, respectively. The lower active compounds have a morpholino group whose electronegative oxygen is negatively interacting with the blue plot near this region. In addition, a blue plot is seen enclosing the carbon number 2 of the morpholino group where increase of positive charge is expected to improve binding potency. The plot also show a red plot near the carbonyl of the amido group which calls for electronegative groups to enhance the affinity. This appears to explain why compounds that have amido carbonyl are more potent (compounds **117**, **13**, **112**) than those molecules that lack this (**111**, **116**, **113**, **114**).

The steric contour for CDK2 is displayed in Fig. 9. The plot shows a big sterically unfavorable yellow region near the azepane ring. This appears to expound why compound **32** is more active than compound **36**, compound **33** is more active than compound **37**, compound **38** is less active than compound **34**. In the less active compounds, carbons 2 and 4 of the morpholino ring are incorporated into the unfavorable yellow plot where as it is only a methyl group that is incorporated in the more active molecules. The plot also shows a medium sized yellow plot near the phenyl of the indene ring and a small one beneath it. These regions are generally expected to reduce the activity with an increase in steric bulkiness around this region. Another small sized yellow contour is seen near the 'N' of the morpholino ring. As a major difference in the steric contours between these kinases is the disappearance of two yellow plots on opposite side of the thienyl ring, the appearance of two small yellow plots above the morpholine ring and the absence and reduction of the yellow contours near the phenyl of the indene ring. These clues are expected to be of importance to design a selective inhibitor. The green plot of the two maps also shows some differences in position. Compounds that have bulkier groups are expected to interact favorably.

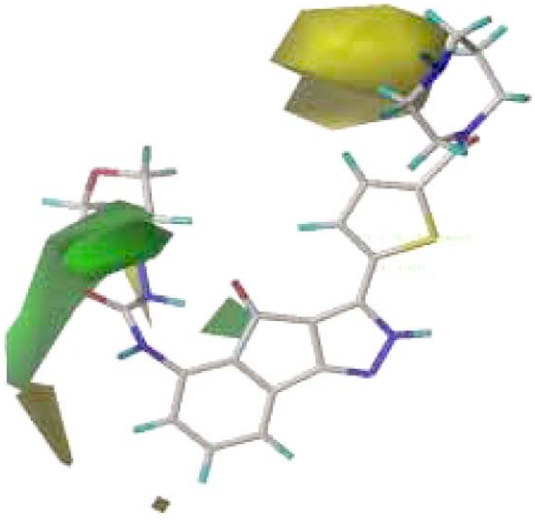


Fig. 9. CoMFA STDEVXCOEFF steric contour map for CDK2 inhibition: green indicates regions which prefer an increase in Lennard–Jones potential and the yellow map calls for a reduction of this potential to improve the affinity.

4. Conclusion

The CoMFA analysis was used to build statistically significant models with good correlative and predictive capability for the inhibition of CDK2 and CDK4 by 117 indenopyarzole derivatives. These models could be used to predict the inhibitory potencies of related structures. The analysis of contours for the CoMFA models has provided a clue about the structural requirement for the observed biological activity for the respective kinases: a more electropositive and less bulky substitution on the pyrazole ring are expected to improve the affinity in both kinases whereas subtituents that increase both the negative charge and the Lennard–Jones potential near the carbamate moiety are likely to improve the inhibitory potency. The differences in size and position of these and the yellow plot near the phenyl of the indene ring could be employed to bias the selectivity to either receptor. This comparative analysis of contours is expected to be of an aid in the design of compounds with an enhanced inhibitory activity and better selectivity to these kinases. We are using these clues to design new chemicals in our lab.

References

[1] S. Renfrey, J. Featherstone, Nat. Rev. Drug Discov. 1 (2002) 75–176.
[2] P. Cohen, Nat. Rev. Drug Discov. 1 (2002) 309–315.
[3] M.E. Noble, J.A. Endicott, L.N. Johnson, Science 303 (2004) 1800–1805.
[4] J.L. Adams, D. Lee, Curr. Opin. Drug Discov. Dev. 2 (1999) 96–109.
[5] C. Garcia-Echeverria, P. Traxler, D.B. Evans, Med. Res. Rev. 20 (2000) 28–57.
[6] R. Sridhar, O. Hanson-Painton, D.R. Cooper, Pharm. Res. 17 (2000) 1345–1353.
[7] J. Dumas, Exp. Opin. Ther. Patents 11 (2001) 405–429.
[8] N.R. Rao, Curr. Opin. Oncol. 8 (1996) 516–524.
[9] K. Lingfei, Y. Pingzhang, G. Jianhua, Z. Yaowu, Cancer Lett. 130 (1998) 93–101.
[10] M. Krockaert, P. Greengard, L. Meijer, Trends in Pharmacol. Sci 23 (2002) 417–425.
[11] R.D. Cramer III, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.

[12] G.R. Desiraju, B. Gopalakrishnan, R.K. Jetti, A. Nagaraju, J.D. Raveendra, A. Sharma, M.E. Sobhia, R. Thilagavathi, J. Med. Chem. 45 (2002) 4847–4857.
[13] L. Meijer, A. Thunnissen, A. White, M. Granier, M. Kiolic, L.H. Tsai, J. Walter, K.E. Cleverly, P.C. Salinas, Y.Z. Wu, J. Biernat, E.M. Mandelkow, S. Kim, G.R. Pettit, Chem. Biol. 7 (2000) 51–63.
[14] H.H. Sedlacek, J. Czech, R. Naik, G. Kuar, P. Worland, M. Losiewiccz, B. Parker, A. Smith, A. Senderowicz, E. Sausville, Int. J. Oncol. 9 (1996) 1143–1147.
[15] R. Hoessel, S. Leclerc, J. Endicott, M. Noble, A. Lawrie, P. Tunnah, M.E.D. Leost, D. Marie, D. Marko, E. Niederberger, W. Tang, G. Eisenbrand, L. Meijer, Nat. Cell Biol. 1 (1999) 60–67.
[16] D. Zaharevitz, R. Gussio, M. Leost, A. Senderowicz, T. Lahusen, C. Kunick, L. Meijer, E.A. Sausville, Cancer Res. 59 (1999) 2566–2569.
[17] J. Vesely, L. Havlicek, M. Strnad, J. Blow, A. Donella-Deana, L. Pinna, D. Letham, J. Kato, L. Detivaud, S. Leclerc, Eur, J. Biochem. (Tokyo) 224 (1994) 771–786.
[18] E. Lescot, R. Bureau, J.S.O. Santos, C. Rochais, V. Lisowski, J.C. Lancelot, S. Rault, J. Chem. Inf. Model. 48 (2005) 7103–7112.
[19] G.-F. Yang, H.T. Lu, Y. Xiong, C.-G. Zhan, Bioorg. Med. Chem. Lett. 14 (2006) 1462–1473.
[20] M.N. Iskander, L.M. Leung, T. Buley, F. Ayad, J.D. Iulio, Y.Y. Tan, I.M. Coupar, Eur. J. Med. Chem. 41 (2006) 16–26.
[21] A.M. Oliveira, F.B. Custódio, C.L. Donnici, C.A. Montanar, Eur. J. Med. Chem. 38 (2003) 141–155.
[22] A.N. David, E. Anne-Marie, V. Anup, A.B. Pamela, B. Micheal, R.B. Catherine, C. Sarah, M.C. Philip, D. Deborah, P.S. Steven, J. Med. Chem. 44 (2001) 1334–1336.
[23] A.N. David, E. Ana-Marie, V. Anup, A.B. Pamela, B. Micheal, R.B. Catherine, C. Sarah, D. Deborah, P.S. Steven, J. Med. Chem. 45 (2002) 5224–5232.
[24] W.Y. Eddy, H.C. Anne, V.D. Susan, J.C. David, A.N. David, B. Carrie, A.B. Pamela, R.B. Catherine, C. Sarah, H.G. Robert, M.S. Diane, M.S. Lisa, F.B. John, K.M. Jodi, M.S. Angela, C. Haiying, C. Chong-Hwan, P.S. Steven, L.T. George, J. Med. Chem. 45 (2002) 5233–5248.
[25] W.Y. Eddy, V.D. Susan, H.C. Anne, A.M. Jay, R.B. Catherine, A.B. Pamela, H.G. Robert, C. Sarah, K.M. Jodi, M.S. Angela, C. Haiying, C. Chong-Hwan, L.T. George, P.S. Steven, Bioorg. Med. Chem. Lett. 14 (2004) 343–346.
[26] A.N. David, V. Anup, D.D. Carolyn, Bioorg. Med. Chem. Lett. 14 (2004) 5489–5491.
[27] SYBYL6.9; Tripos Inc., 1699 South Hanley Rd., St. Louis, MO 63144, USA.
[28] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, J. Mol. Biol. 261 (1996) 470–489.
[29] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902–3909.
[30] E.A. Amin, W.J. Welsh, J. Med. Chem. 44 (2001) 3849–3855.
[31] S. Wold, A. Johansson, M. Cochi, in: H. Kubinyi (Ed.), 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Lieden, 1993, pp. 523–550.
[32] R.D. Cramer III, J.D. Bunce, D.E. Paterson, Quant. Struct. Act. Relat. 7 (1988) 18–25.