

Molecular-docking-guided 3D-QSAR studies of substituted isoquinoline-1,3-(2*H*,4*H*)-diones as cyclin-dependent kinase 4 (CDK4) inhibitors

Xiao-Yun Lu · Ya-Dong Chen · Ni-yue Sun ·
Yong-Jun Jiang · Qi-Dong You

Received: 15 February 2009 / Accepted: 3 May 2009 / Published online: 20 June 2009
© Springer-Verlag 2009

Abstract The cyclin-dependent kinases (CDKs) are critical regulators of cell cycle progression, and are involved in uncontrolled cell proliferation—a hallmark of cancer. This suggests that small molecular inhibitors of CDKs might be attractive as prospective antitumor agents. To explore the relationship between the structures of substituted isoquinoline-1,3-(2*H*,4*H*)-diones and their inhibition of CDK4, 3D-QSAR studies were performed on a dataset of 48 compounds. The bioactive conformation of template compound **34** was obtained by performing molecular docking into the ATP binding site of the homology model of CDK4 and ranking by highest consensus score, which was then used to build and align the rest of the molecules in the series. The constructed comparative molecular similarity indices analysis (CoMSIA) produces significantly better results than comparative molecular field analysis (CoMFA), with $r_{cv}^2 = 0.707$ and $r^2 = 0.988$. The contours analysis provides useful information about the structural requirements for substituted isoquinoline-1,3-(2*H*,4*H*)-diones for CDK4 inhibitory activity.

Keywords Cyclin-dependent kinase 4 (CDK4) · 3D-QSAR · Molecular docking · Homology modeling · Isoquinoline-1,3-(2*H*,4*H*)-diones

X.-Y. Lu · Y.-D. Chen · N.-y. Sun · Q.-D. You (✉)
Department of Medicinal Chemistry,
China Pharmaceutical University,
24 Tongjiaxiang,
Nanjing 210009, People's Republic of China
e-mail: youqidong@gmail.com

Y.-J. Jiang
Ningbo Institute of Technology, Zhejiang University,
Ningbo 315104, People's Republic of China

Introduction

The cell cycle is a frequent target of genetic alterations in cancer because of its central role in the control of cell growth and proliferation. Cyclin-dependent kinases (CDKs) are a family of serine-threonine kinases that play key roles as regulators of cell progression through forming complexes with cyclins [1, 2]. The basic cell cycle is divided into four phases, namely G1, S, G2, and M. Specific CDKs operate in distinct phases of the cell cycle. 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 to the cell's environment. Abnormal CDK control of the cell cycle has been linked strongly to the molecular pathology of cancer. Thus, inhibitors of CDK-cyclin complexes might have a broad range of therapeutic applications in cancer as well as other diseases [3]. Furthermore, CDK4/6-cyclinD complexes, have very recently been demonstrated as effective breast cancer targets [4].

A huge variety of small molecules have been reported as CDK inhibitors (CDKIs). These include various scaffolds, e.g., R-roscovitine (CYC202, Seliciclib) [5], flavopiridol [6], 7-hydroxystaurosporine (UCN-01) [7], pyridopyrimidine (PD-0332991) [8] and the aminothiazole compound (BMS-387032) [9], etc., but so far none of them have progressed to being a clinically useful drug. One of the main bottlenecks hampering the development of a kinase inhibitor drug is the difficulty in attaining selectivity. This appears to stem from the diverse nature of the kinase substrates and the common mechanism these enzymes share among themselves. Nowadays, the synthesis of novel highly selective CDKIs as candidates for CDK-target therapy in cancer treatment is in high demand [10].

To the best of our knowledge, there have been only a few reports so far on three-dimensional quantitative structure activity relationships (3D-QSAR) and binding site analysis for CDKIs, especially for CDK4 inhibitors. Singh et al. developed comparative molecular field analysis (CoMFA) models for indenopyrazole derivatives as CDK2 and CDK4 inhibitors, with the intention of designing ligands with enhanced inhibitory potencies, and predicting the potencies of analogues to guide synthesis [11]. Dessalew et al. carried out a 3D-QSAR study on benzodipyrzoles as CDK2 inhibitors [12]. The same authors developed CoMFA models on a bisar-ylmaleimide series as GSK3, CDK2 and CDK4 inhibitors with the intention of optimizing and enhancing selectivity toward GSK3 [13]. More recently, Mascarenhas et al. carried out a QSAR study focusing on the essential physicochemical characteristics of CDK4 inhibitors based on combining ligand and structural approaches with a homology model of CDK4 [14]. Caballero et al. reported 2D autocorrelation and a 3D-QSAR analysis of pyrido[2,3-d]pyrimidin-7-ones as CDK4 inhibitors [15].

Recently, a series of new isoquinoline-1,3-(2*H*,4*H*)-diones derivatives has been designed and synthesized that can clearly inhibit CDK4 in a selective manner [16]. Thus, it would be very useful to investigate the QSAR as well as the mechanism of CDK4 inhibitory activity for this new series. In this study, 3D-QSAR studies guided by molecular docking with the ATP binding site of the homology model of CDK4 were performed using 48 substituted isoquinoline-1,3-(2*H*,4*H*)-diones derivatives. We believe this study provides useful information about the structural requirements of CDK4 selective inhibitors, and expect the results will aid in the design of new synthetic drugs.

Materials and methods

Data sets

To obtain a reliable and robust QSAR model, it is desirable to consider a dataset that covers reasonable chemical diversity and biological activity. Hence, a set of 48 compound biological data taken from one laboratory as reported by Tsou et al. [16] was chosen. The biological data were considered comparable and were divided into a training set and a test set as shown in Table 1. The training set was selected randomly and consists of 41 compounds; the test set is comprised of 7 compounds. The test set includes compounds representing all categories of activity of the training set, i.e., inactive, active, and more active compounds comprising all the structural features that are important for activity. The IC_{50} values were converted into pIC_{50} ($-\log IC_{50}$), which were then used for subsequent QSAR analysis as dependent variables.

Molecular modeling

The 3D structures of the compounds were constructed using the molecular modeling package Sybyl6.9 [17]. Partial atomic charges of all molecules were calculated by the Gasteiger-Hückel method in Sybyl6.9. As the crystal structure of CDK4 is not available, we performed docking analysis of the most potent compound **34** at the ATP binding site of the homology model of CDK4 using the docking program GOLD 3.1 [18]. The conformation of compound **34** with the highest consensus score was taken for further 3D-QSAR study, used as the template to construct and align the structures of the remaining compounds in the data set.

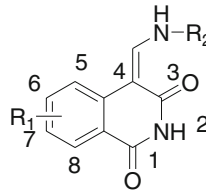
Molecular alignment

Molecular alignment is the most sensitive parameter in 3D-QSAR analysis. The quality and the predictive ability of the model are directly dependent on the alignment rule. In the present study, the docking-guided conformation of compound **34** was used as the basic skeleton to build the remaining compounds by modifying the required substitutions. Partial atomic charges were assigned to each atom, and energy minimization of each molecule was then performed using the Tripos standard force field while maintaining the conformation of the common structure of isoquinoline-1,3-(2*H*,4*H*)-diones. Compound **34** was used as a template to align the other 47 compounds from the series by common substructure alignment, using the ALIGN DATABASE command in Sybyl, ver. 6.9. The common substructure used for alignment, and the superimposed structure after alignment are presented in Fig. 1.

CoMFA and CoMSIA

The steric and electrostatic CoMFA potential fields were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. Energy cutoff values of 30 kcal mol⁻¹ were selected for both the electrostatic and steric fields. The optimal number of components was designated such that cross-validated r_{cv}^2 was maximal and the standard error of prediction was minimal.

Taking the same aligned molecules that were used for CoMFA, we performed a statistical evaluation of the comparative molecular similarity indices analysis (CoMSIA) analysis, using some of the same parameters as for CoMFA. Five similarity fields, namely, steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor fields were evaluated at each lattice intersection of a regularly spaced grid of 2.0 Å. In optimizing CoMSIA performance, the most important parameter is how to combine the five fields in the CoMSIA model. In order to choose the optimal

Table 1 Structures, pIC₅₀ values (experimental and predicted) and residuals of substituted isoquinoline-1,3-(2*H*,4*H*)-dione derivatives


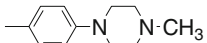
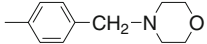
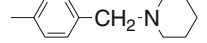
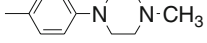
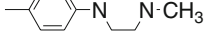
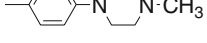
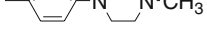
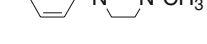
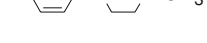
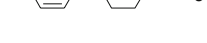

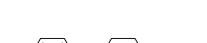
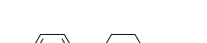
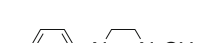
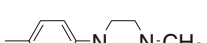

compd	R ₁	R ₂	IC ₅₀ (μM)	pIC ₅₀	CoMFA		CoMSIA	
					PA ^{a)}	Δ ^{b)}	PA ^{a)}	Δ ^{b)}
1	H		4.1	5.39	5.45	-0.06	5.48	-0.09
2	H		10	5.00	5.24	-0.24	5.00	0
3*	H		3.3	5.48	5.48	0	5.80	-0.32
4	7-Br		12.1	4.92	4.96	-0.04	4.91	0.01
5	6-Br		1.40	5.85	5.67	0.18	5.94	-0.09
6	6-NO ₂		11	4.96	5.64	-0.68	5.00	-0.04
7	6-I		0.48	6.32	5.84	0.48	6.11	0.21
8	6-C(O)N(CH ₃) ₂		41	4.39	4.52	-0.13	4.52	-0.13
9	6-pyrrolyl		0.14	6.85	6.84	0.01	6.86	-0.01
10	6-NHAc		11.0	4.96	5.12	-0.16	4.88	0.08
11	6-piperidinyl		1.62	5.79	6.48	-0.69	5.80	-0.01
12	6-phenyl		0.39	6.41	6.57	-0.16	6.44	-0.03
13*	6-(2-furyl)		0.33	6.48	6.10	0.38	6.48	0
14	6-(3-furyl)		0.22	6.66	6.47	0.19	6.65	0.01
15	6-(3-thienyl)		0.10	7.00	6.73	0.27	6.97	0.03
16	6-(4-CHO-phenyl)		0.13	6.89	6.82	0.07	6.91	-0.02

Table 1 (continued)

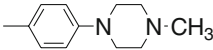
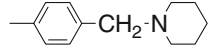
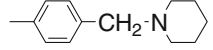
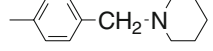
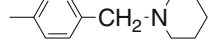
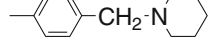
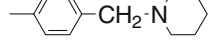
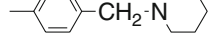
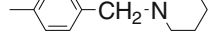
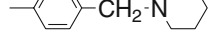
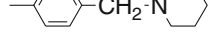
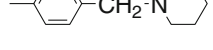
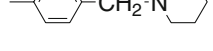
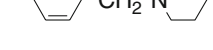
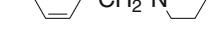
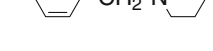
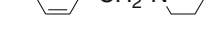




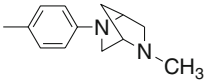
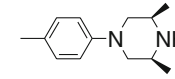
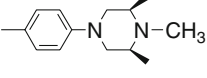
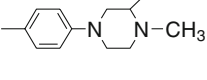
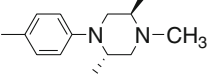
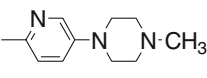
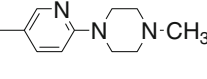
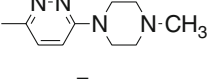
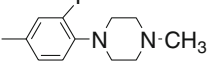
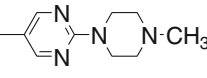
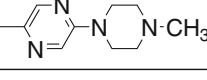
17	6-(C≡-phenyl)		2.30	5.64	6.32	-0.68	5.70	-0.06
18	6-Br		1.1	5.96	5.68	0.28	5.85	0.11
19	6-Cl		2.5	5.60	5.56	0.04	5.80	-0.2
20	6-OCH3		2	5.70	5.32	0.38	5.68	0.02
21	6-NO2		15.1	4.82	5.54	-0.72	4.91	-0.09
22*	6-I		0.3	6.52	5.64	0.88	6.01	0.51
23	6-C(O)N(CH3)2		21.2	4.67	4.52	0.15	4.48	0.19
24	6-Pyrrolyl		0.14	6.85	6.48	0.37	6.75	0.1
25	6-NHAc		3.5	5.45	5.26	0.19	5.39	0.06
26*	6-piperidinyl		1.0	6.00	6.40	-0.4	6.42	-0.42
27	6-morpholinyl		0.92	6.04	6.49	-0.45	5.96	0.08
28	6-NH-Ph		1.8	5.75	6.12	-0.37	5.82	-0.07
29	6-Phenyl		0.32	6.50	6.33	0.17	6.65	-0.15
30*	6-(3-furyl)		0.037	7.43	7.64	-0.21	7.07	0.36
31	6-(3-pyridyl)		0.05	7.30	7.24	0.06	7.23	0.07
32	6-(3-thienyl)		0.13	6.89	6.75	0.14	7.06	-0.17
33	6-(4-F-phenyl)		0.31	6.51	6.87	-0.36	6.46	0.05
34	6-(3-OH-phenyl)		0.027	7.57	7.49	0.08	7.61	-0.04
35	6-(4-OH-phenyl)		0.041	7.39	7.15	0.24	7.39	0
36	6-(4-OCH3-phenyl)		0.13	6.89	7.04	-0.15	6.78	0.11
37*	6-CN		15.8	4.80	5.55	-0.75	5.53	-0.73

Table 1 (continued)

38	6-I		0.32	6.49	6.12	0.37	6.44	0.05
39	6-I		1.0	6.00	5.61	0.39	5.95	0.05
40	6-I		2.8	5.55	5.58	-0.03	5.61	-0.06
41	6-I		2.7	5.57	5.73	-0.16	5.67	-0.1
42	6-I		3.0	5.52	5.57	-0.05	5.53	-0.01
43	6-(3-furyl)		0.13	6.89	6.72	0.17	6.93	-0.04
44*	6-(3-furyl)		0.11	6.96	6.54	0.42	6.34	0.62
45	6-(3-furyl)		0.82	6.09	5.43	0.66	6.01	0.08
46	6-(3-furyl)		0.08	7.10	6.52	0.58	7.05	0.05
47	6-(3-furyl)		0.25	6.60	6.52	0.08	6.62	-0.02
48	6-(3-furyl)		1.25	5.90	6.27	-0.37	5.79	0.11

*Test set molecules

^a Predicted activity;^b Residual of experimental and predicted activities.

result, we systemically altered the combination of fields and chose values that gave ideal non-cross-validation, standard errors of estimate and F values.

Partial least squares analysis

Partial least squares (PLS) [19, 20] methodology was used for all 3D-QSAR analysis. The cross-validation [21, 22] analysis was performed using the leave one out (LOO) method, in which one compound is removed from the dataset, and its activity is then predicted using the model derived from the rest of the dataset. The cross validated r^2 that resulted in the optimum number of components and the lowest standard error of prediction were considered for further analysis. To speed up the analysis and reduce noise, a minimum filter value, σ , of 2.00 kcal mol⁻¹ was used. A

final analysis was performed to calculate conventional r^2 using the optimum number of components obtained from the cross-validation analysis.

The predictive power of each of the 3D-QSAR models was determined from a set of seven molecules that were excluded during model development. The optimization, alignment and all other steps of these test set molecules were the same as that of the training set molecules described above, and their activities were predicted using the model produced by the training set. The predictive correlation (r_{pred}^2) based on the test set molecules, is computed using

$$r_{\text{pred}}^2 = (\text{SD} - \text{PRESS})/\text{SD}$$

Where SD is defined as the sum of the squared deviations between the biological activities of the test set and the mean

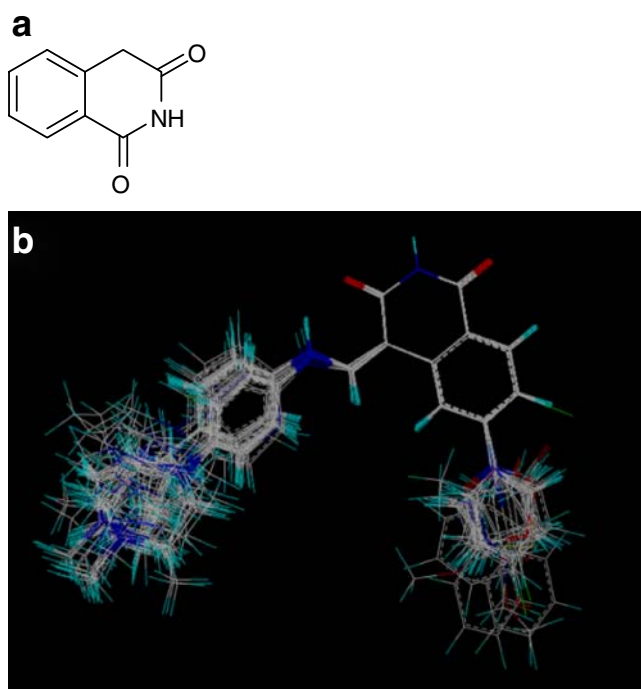


Fig. 1 **a** Common substructure used for alignment. **b** Superimposed structure after alignment of molecules

activity of the training set compounds, and PRESS is the sum of the squared deviation between the predicted and actual activity values for each molecule in the test set.

Homology modeling

Since the 3D structure of CDK4 has not been determined, structural information was obtained by homology modeling. According to one published study [23], despite the very high sequence identity between CDK4 and CDK6 (68%), we did not use the 3D structure of CDK6 as the homology template. Instead, the known crystal structure of CDK2 (PDB code: 1PKD) [24], with a sequence identity of 46%, was used as a template. The sequence of CDK4 was obtained from the SWISS-PROT protein sequence database, and the crystal structure of CDK2 was taken from the Protein Data Bank [24]. For modeling the 3D structure of CDK4, the Modeler program implemented in Insight II [25] was used. The model was finally minimized with CVFF force fields and the final structure was further checked using the programs PROCHECK and PROFILE-3D.

Molecular docking

The most potent compound, **34**, was docked into the ATP binding site of the homology model of CDK4, using GOLD 3.1, a powerful genetic algorithm (GA) method for conformational searches that is widely regarded as one of

the best docking programs [26]. The genetic operators were: 100 for the population size, 1.1 for the selection, 5 for the number of subpopulations, 100,000 for the maximum number of genetic applications, and 2 for the size of the niche used to increase population diversity. The weights were chosen so that crossover mutations were applied with equal probability (95/95 for the values) and migration was applied 5% of the time. The ChemScore function encoded in GOLD was applied to rank different binding poses. To identify the bioactive conformation of compound **34**, ten docked poses obtained from the GOLD program were transferred to Discovery studio (v 1.7) to apply other scores, namely LigScores, PLP scores, PMF, Jain, Ludi scores [27]. The conformation of compound **34** with highest consensus score was then further used for the 3D-QSAR study.

Results and discussion

CoMFA and CoMSIA results

CoMFA and CoMSIA 3D-QSAR models were derived from a training set of 41 molecules with IC_{50} values ranging from 0.027 μ M to 41 μ M. The stepwise development of CoMFA and CoMSIA models using different fields is presented in Table 2. The CoMFA model describing CDK4 inhibition used both steric and electrostatic fields and has a r_{cv}^2 value of 0.505 using three components, with a standard deviation (SEE=0.357) and a Fischer ratio ($F=56.62$). In comparison to CoMFA, CoMSIA methodology has the advantage of exploring more fields. The best CoMSIA model included steric, electrostatic and hydrogen bond donor fields (CoMSIA-SED) and has a r_{cv}^2 value of 0.707 using eight components, with a low standard deviation (SEE=0.101) and a high Fischer ratio ($F=319.65$). The predicted values for the 41 compounds in the training set and 7 compounds in the test set using CoMFA and CoMSIA-SED models are shown in Table 1. The correlations between the predicted and experimental values of all compounds are shown in Fig. 2. The predictability of the models is the most important criterion for assessment of both methods. The predictive power of the CoMSIA-SED model was evaluated as better than that of CoMFA, with r_{pred}^2 values of 0.761 and 0.618, respectively. The PLS statistics of both CoMFA and CoMSIA 3D-QSAR models indicate that CoMSIA-SED produced better results than CoMFA. The steric and electrostatic plots from the CoMSIA model are in accordance with the field distributions of CoMFA contour plots, and thus the contour plots of CoMSIA model only are displayed for analysis.

Table 2 Summary of three-dimensional quantitative structure activity relationships (3D-QSAR) analysis results obtained using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). r_{cv}^2 Cross-validated correlation coefficient, N number of components, r^2 conventional correlation coefficient, SEE

standard error of estimate, F -value F -test value, S steric field, E electrostatic field, H hydrophobic field, D hydrogen bond donor field, A hydrogen bond acceptor field, r_{pred}^2 predicted correlation coefficient for test set of compounds, SE standard error of test set

Parameter	CoMFA			CoMSIA							
	S	E	S,E	S	E	H	D	A	E,H	S,E,D	ALL
r_{cv}^2	0.382	0.318	0.505	0.095	0.706	0.355	0.034	0.253	0.619	0.707	0.615
N	7	2	3	9	6	3	2	1	4	8	9
r^2	0.948	0.682	0.821	0.864	0.966	0.661	0.234	0.389	0.892	0.988	0.993
SEE	0.205	0.470	0.357	0.340	0.162	0.492	0.730	0.644	0.281	0.101	0.075
F -value	85.29	40.74	56.62	21.93	162.78	24.07	5.81	24.81	74.68	319.65	517.07
Contributions											
Steric			0.495							0.166	0.131
Electrostatic			0.505						0.542	0.757	0.372
Hydrophobic									0.458		0.239
Donor										0.077	0.093
Acceptor											0.165
r_{cv}^2			0.618							0.761	
SE			1.486							0.780	

CoMSIA contour plots analysis

Steric, electrostatic and hydrogen bond donor contour plots obtained using the CoMSIA method are useful to explore protein–ligand interactions, as shown in Fig. 3. For simplicity, only interactions between the most active compound **34** and the contour plots are shown. Contour plots show the requirements of the basic scaffold of isoquinoline-1,3-(2*H*,4*H*)-diones for increasing CDK4 inhibitory activity.

The green contour observed near the *m*- and *p*-positions of the phenyl ring at the C-6 of the isoquinoline-1,3-(2*H*,4*H*)-dione (see Table 1 for atom numbering) indicates that some bulky substitutions at these positions are favorable for activity (Fig. 3a), which is known to be extended into a hydrophobic pocket created by the side chains of Lys142, Glu144 and Asn145. Compounds **16** and **33–36** show similar activities, as they possess CHO, F, OH and OCH₃ groups in that region. However, compounds **16**, **33**, and **36**, which are substituted at the *p*-position, are less potent than compound **34** with OH at the *m*-position, which is known to form a hydrogen bond with the backbone carbonyl of Asn145. This green contour may also indicate that R₁ groups at the 6-position of the scaffold of isoquinoline-1,3-(2*H*,4*H*)-diones with bulky substituents (6-I and 6-heteroaryl) are more potent than 6-Br and 6-Cl substituents. This explains why compounds **22**, **30–32** are more potent than compounds **18** and **19**. The small yellow contour near the phenyl ring at the C-6 of the isoquinoline-

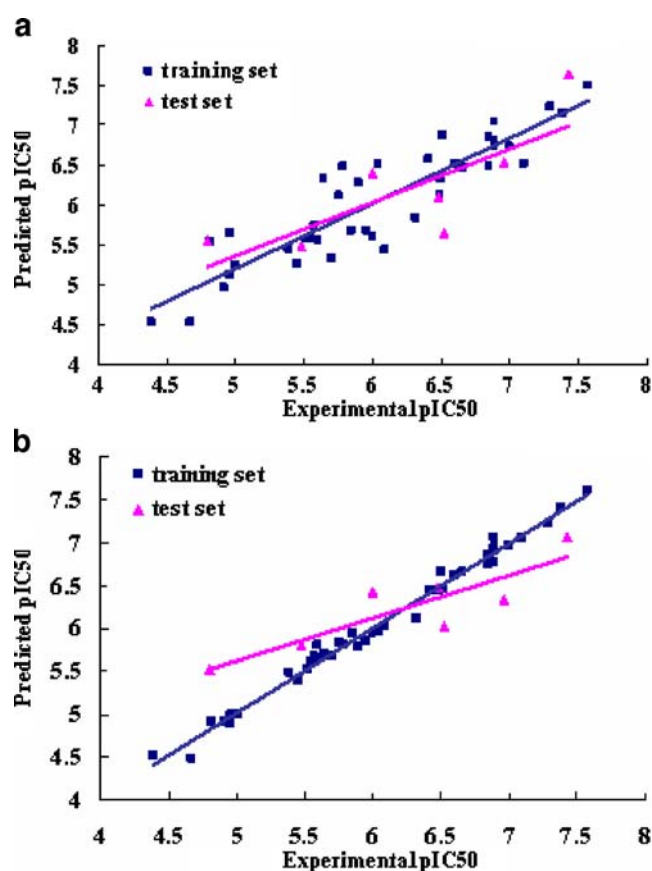


Fig. 2 Correlation between experimental pIC_{50} and predicted pIC_{50} of the training and test sets. **a** Comparative molecular field analysis (CoMFA) model. **b** Comparative molecular similarity indices analysis (CoMSIA)-SED model (CoMSIA model included steric, electrostatic and hydrogen bond donor fields)

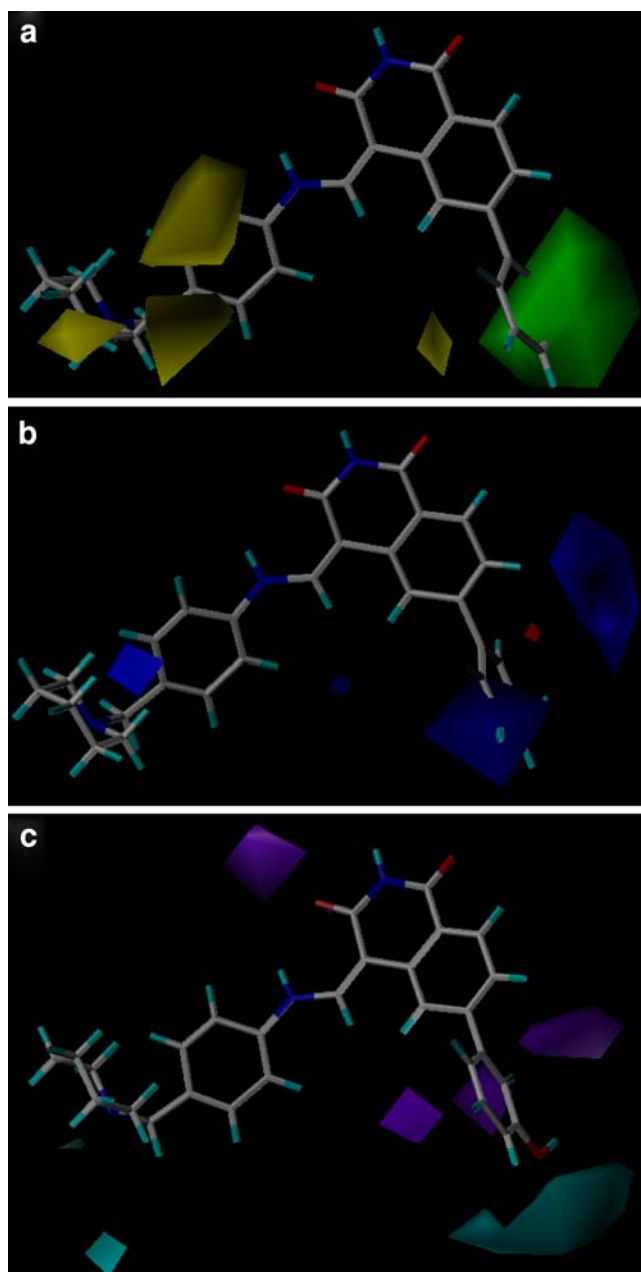


Fig. 3a–c Stereo view of CoMSIA contour plots for CDK4 inhibitors. **a** Steric fields: *green contours* indicate regions where bulky groups increase activity, while *yellow contours* indicate regions where bulky groups decrease activity. **b** Electrostatic fields: *blue contours* indicate regions where electropositive groups increase activity, while *red contours* indicate regions where electronegative groups increase activity. **c** Hydrogen bond donor fields: *cyan contour* indicates regions where hydrogen bond donor groups increase activity. *Purple* is disfavored. Compound **34** is shown inside the field

1,3-(2H,4H)-dione suggests that any bulky substitution at this position is likely to decrease activity. That is why addition of a triple bond between the phenyl and the core, compound **17**, shows much reduced activity. Also, the yellow contours near the piperidine ring indicate that any bulky substitutions at this position are disfavored in terms

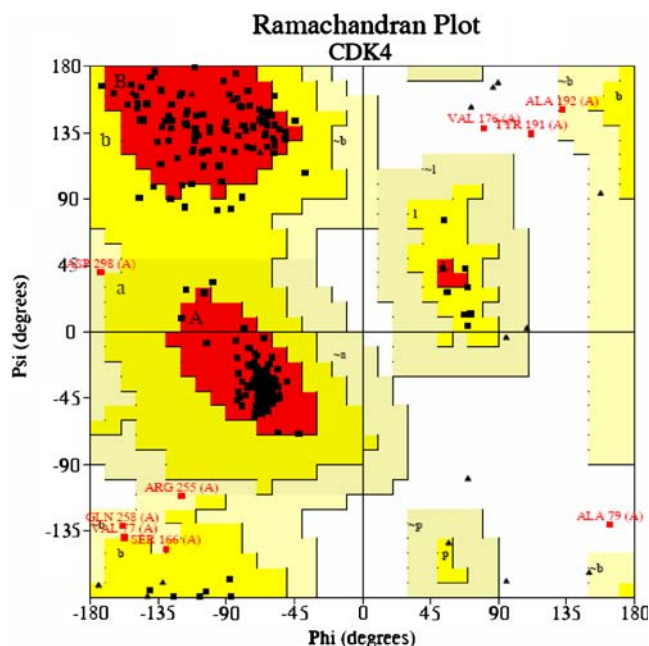


Fig. 4 Ramachandran (Phi–Psi) plot of the modeled CDK4 obtained using PROCHECK. Residues in generously allowed regions and disallowed regions are labeled

of inhibitory activities. This can explain the fact that the activity of compound **22** is higher than that of compounds **39–42**.

The CoMSIA electrostatic contour is shown in Fig. 3b. The two blue contours near the 6- and 7-positions of the scaffold of the isoquinoline-1,3-(2H,4H)-diones indicate that positive charge in these regions is favored for inhibitory activity. This can be seen with compounds **6**, **8**, **10**, **21**, **23** and **37**, which bear electron-withdrawing substituents such as nitro, *N,N*-dimethylcarboxamido, acet-

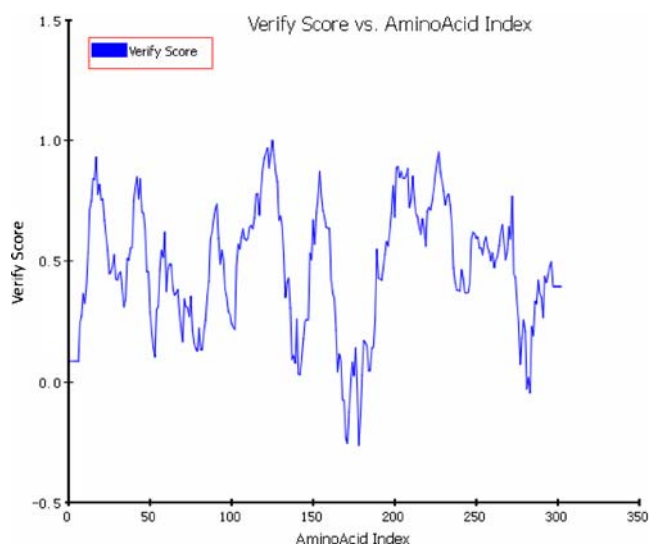


Fig. 5 Evaluation of the final CDK4 structure by the program PROFILE-3D

Table 3 Summary of the consensus scores of ten docked poses

ID	ChemScore	LigScore1	LigScore2	PLP1	PLP2	Jain	PMF	Ludi1	Ludi2	Ludi3	Consensus
34(1)	33.53	2.27	1.99	93	91.65	5.84	102.14	823	623	603	6
34(2)	33.22	4.53	4.65	90.07	88.6	5.06	84.45	736	538	622	6
34(3)	32.23	4.31	4.43	86.34	88.37	5.72	101.26	735	569	667	6
34(4)	32.21	5.48	6.1	84.46	82.19	4.19	94.17	697	524	629	3
34(5)	31.09	5.13	5.88	85.87	82.54	4.83	91.74	721	532	635	3
34(6)	31.06	4.96	5.41	88.06	84.24	4.59	85.45	718	528	628	2
34(7)	30.82	4.43	4.75	86.44	82.17	4.75	91.15	688	528	635	2
34(8)	30.80	1.82	0.09	90.66	96.37	6.95	82.16	823	623	603	6
34(9)	30.71	2.56	4.09	100.31	95.33	5.07	103.66	617	530	527	4
34(10)	30.15	4.66	4.68	89.57	86.68	4.34	93.56	739	542	643	7

amido and cyano at the 6- and 7-positions, resulting in dramatic decreases in CDK4 activity. It also explains the good inhibitory activities of compounds **20** and **36**, both of which have a positive charge (OCH₃) at these two regions.

The CoMSIA hydrogen bond donor contour plot is displayed in Fig. 3c. The cyan contour near the 3-position of the phenyl ring indicates that a hydrogen bond donor group at this position enhances activity. This accounts for the better activity of the most potent compound **34** and compound **35** as compared to the other molecules such as compounds **12**, **29**, and **36**. The magenta contours indicate that hydrogen bond acceptor groups in these regions are required for high activity.

Homology model of CDK4

A homology model of CDK4 was built to analyze the biologically active conformation of the most potent compound **34** at the binding site. In the modeling, ten homology models were generated in Insight II. The best model, as determined by the lowest value of the Modeler objective function, was minimized with CVFF force fields. The resulting model had 86.9% of the residues in the most favored region of the Ramachandran plot (Fig. 4), and was used for the docking study. Checking with profile-3D showed that the overall self-compatibility score for this protein was 129.81, which was much higher than the minimum score of 61.843 and close to the expected score of 137.429 (Fig. 5)—a compatibility score above zero indicates the presence of residues in the favorable regions; residues below zero lie far away from the active site of CDK4 and have no impact on this study [28].

Docking analysis

Although docking programs perform well in predicting and generating the correct biologically active conformation of a

ligand, current scoring functions are less successful at correct identification. Considering this objective, a consensus scoring strategy was adopted. The most potent two-molecular, **34**, was docked into the active site of the homology model of CDK4 using the GOLD 3.1 docking program with the ChemScore function. The resulting ten docked poses were transferred to a consensus scoring approach with multiple scoring functions. The scaled scores were then summed to give the final consensus scores shown in Table 3. The highest consensus score of 7 was selected as the bioactive conformation. The final binding pose of compound **34** at the ATP active site of homology model of CDK4 is shown in Fig. 6.

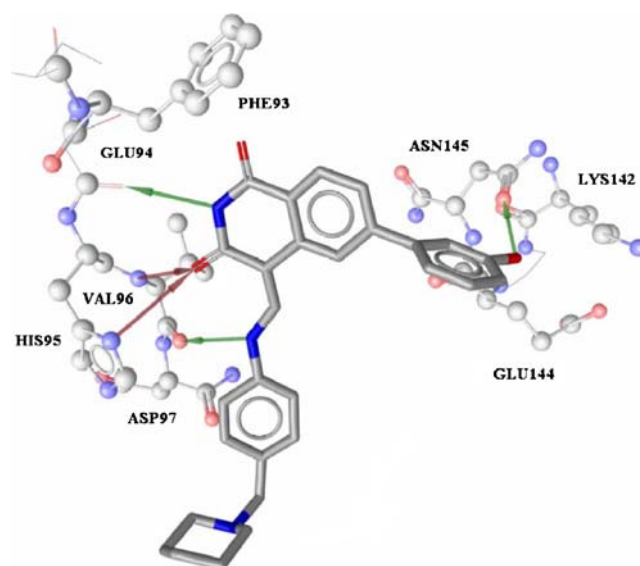


Fig. 6 The docked conformation of compound **34** in the active site of homology model of CDK4 with highest consensus score function. *Green arrows* Hydrogen bonds (limited within 3.5 Å). There are 13 differences in amino acid sequence between the homology model of CDK4 and the CDK4 mimic model

The modeling binding conformation represents the interaction model of the most potent compound **34** with CDK4. The basic scaffold of the isoquinoline-1,3-(2*H*,4*H*)-dione of compound **34** makes several strong hydrogen bonds with Glu94, His95 and Val96 amino acid residues at the ATP active site of CDK4. It can be seen clearly from Fig. 6 that the NH of the isoquinoline-1,3-(2*H*,4*H*)-dione core forms an H-bond with the backbone carbonyl of Glu94. Additionally, a carbonyl oxygen of isoquinoline-1,3-(2*H*,4*H*)-dione forms two interactions with the backbone NHs of His95 and Val96, while the amino NH of the enamine headpiece forms an H-bond with the backbone carbonyl oxygen of Glu94. Moreover, the 6-phenyl substituent of compound **34** is surrounded by residues Lys142, Glu144 and Asn145, mainly through hydrogen bond interactions between the hydroxyl group at the *m*-position of the 6-phenyl and Asn145. The modeling hydrogen bond interactions (limited to within 3.5 Å) between compound **34** and CDK4 are identical to those described in a previous study [16]. This docking guided conformation of compound **34** was then further used for 3D-QSAR analysis.

Conclusions

To explore the structure–activity relationships of substituted isoquinoline-1,3-(2*H*,4*H*)-diones for their CDK4 inhibitory activity and to build a statistically significant model with good predictive power, a 3D-QSAR study was performed. The high degree of sequence identity between the CDK2 and CDK4 kinase domains enabled us to construct a reasonable good homology model of CDK4. The docked conformation of compound **34** with highest consensus score was used as the template to build and align the rest of the compounds to it. The CoMSIA-SED model demonstrated better predictive ability than the COMFA with $r_{cv}^2 = 0.707$ and $r^2 = 0.988$. Contour plots show that optimum bulky substituents with hydrogen bond donors at the *m*- and *p*-positions of the phenyl ring at the C-6 of the isoquinoline-1,3-(2*H*,4*H*)-dione are important for high activity. This suggests that the groups in these positions can form hydrogen bond interactions with Lys142 and Asn145 of the homology model of CDK4. Furthermore, substitution of an electropositive group at the 6- and 7-positions of the scaffold of the isoquinoline-1,3-(2*H*,4*H*)-diones is essential for activity. In addition, bulky substitution at the piperidine ring is unfavorable for inhibitory activity. The results provide useful information on the structural requirements of isoquinoline-1,3-(2*H*,4*H*)-diones derivatives for CDK4 inhibitory activity, which could be utilized in the future design of more potent CDK4 inhibitors.

References

- Harper JW, Adams PD (2001) Cyclin-dependent kinases. *Chem Rev* 101:2511–2526
- Massague J (2004) G1 cell-cycle control and cancer. *Nature* 432:298–306
- Knockaert M, Greengard P, Meijer L (2002) Pharmacological inhibitors of cyclin-dependent kinases. *Trends Pharmacol Sci* 23:417–425
- Yu Q, Sicinska E, Geng Y et al (2006) Requirement for CDK4 kinase function in breast cancer. *Cancer Cell* 9:23–32
- Meijer L, Borgne A, Mulner O et al (1997) Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. *Eur J Biochem* 243:527–536
- Senderowicz AM, Headlee D, Stinson SF et al (1998) Phase I trial of continuous infusion flavopiridol, a novel cyclin-dependent kinase inhibitor, in patients with refractory neoplasms. *J Clin Oncol* 16:2986–2999
- Fuse E, Kuwabara T, Sparreboom A et al (2005) Review of UCN-01 development: a lesson in the importance of clinical pharmacology. *J Clin Pharmacol* 45:394–403
- Fry DW, Harvey PJ, Keller PR et al (2004) Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Mol Cancer Ther* 3:1427–1438
- Misra RN, Xiao HY, Kim KS et al (2004) N-(cycloalkylamino) acyl-2-aminothiazole inhibitors of cyclin-dependent kinase 2. N-[5-[[[5-(1,1-dimethylethyl)-2-oxazolyl]methyl]thio]-2-thiazolyl]-4-piperidinecarboxamide (BMS-387032), a highly efficacious and selective antitumor agent. *J Med Chem* 47:1719–1728
- Wang S, Meades C, Wood G et al (2004) 2-Anilino-4-(thiazol-5-yl)pyrimidine CDK inhibitors: synthesis, SAR analysis, X-ray crystallography, and biological activity. *J Med Chem* 47:1662–1675
- Singh SK, Dessalew N, Bharatam PV (2006) 3D-QSAR CoMFA study on indenopyrazole derivatives as cyclin dependent kinase 4 (CDK4) and cyclin dependent kinase 2 (CDK2) inhibitors. *Eur J Med Chem* 41:1310–1319
- Dessalew N, Singh SK (2008) 3D-QSAR CoMFA and CoMSIA study on benzodipyrzoles as cyclin dependent kinase 2 inhibitors. *Med Chem* 4:313–321
- Dessalew N, Bharatam PV (2007) 3D-QSAR and molecular docking study on bisarylmaleimide series as glycogen synthase kinase 3, cyclin dependent kinase 2 and cyclin dependent kinase 4 inhibitors: an insight into the criteria for selectivity. *Eur J Med Chem* 42:1014–1027
- Mascarenhas NM, Ghoshal N (2008) Combined ligand and structure based approaches for narrowing on the essential physicochemical characteristics for CDK4 inhibition. *J Chem Inf Model* 48:1325–1336
- Caballero J, Fernández M, González NF (2008) Structural requirements of pyrido2,3-d]pyrimidin-7-one as CDK4/D inhibitors: 2D autocorrelation, CoMFA and CoMSIA analyses. *Bioorg Med Chem* 16:6103–6115
- Tsou HR, Otteng M, Tran T (2008) 4-(Phenylaminomethylene) isoquinoline-1,3(2*H*,4*H*)-diones as potent and selective inhibitors of the cyclin-dependent kinase 4 (CDK4). *J Med Chem* 51:3507–3525
- Sybyl ver.6.9, TRIPOS, St Louis, <http://trijos.com>
- Jones G, Willett P, Glen RC et al (1997) Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 267:727–748
- Wold S, Albano C, Dunn WJ III et al (1984) In: Kowalski BR (ed) *Chemometrics-mathematics and statistics in chemistry*. Riedel, Dordrecht, pp 17–95

20. Stahle L, Wold S (1987) Partial least squares analysis with cross-validation for the two-class problem: a. Monte Carlo study. *J Chemom* 1:185–196
21. Cramer RD, Bunce JD, Patterson DE (1988) Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. *Quant Struct Act Relat* 7:18–25
22. Podlogar BL, Ferguson DM (2000) QSAR and CoMFA: a perspective on the practical application to drug discovery. *Drug Des Discov* 17:4–12
23. Honma T, Hayashi K, Aoyama T (2001) Structure-based generation of a new class of potent Cdk4 inhibitors: new de novo design strategy and library design. *J Med Chem* 44:4615–4627
24. Protein Data Bank <http://www.rcsb.org/pdb/explore.do?structureId=1PKD>
25. Accelrys Inc. <http://www.accelrys.com>.
26. Kontoyianni M, McClellan LM, Sokol GS (2004) Evaluation of docking performance: comparative data on docking algorithms. *J Med Chem* 47:558–565
27. Mascarenhas NM, Ghoshal N (2008) An efficient tool for identifying inhibitors based on 3D-QSAR and docking using feature-shape pharmacophore of biologically active conformation—a case study with CDK2/CyclinA. *Eur J Med Chem* 43:2807–2818
28. Schulz GE, Schirmer RH (1974) Topological comparison of adenylyl kinase with other proteins. *Nature* 250:142–144