

3D-QSAR and molecular docking studies of azaindole derivatives as Aurora B kinase inhibitors

Ping Lan · Wan-Na Chen · Ping-Hua Sun ·
Wei-Min Chen

Received: 17 June 2010 / Accepted: 21 July 2010 / Published online: 11 August 2010
© Springer-Verlag 2010

Abstract The Aurora kinases have been regarded as attractive targets for the development of new anticancer agents. Recently a series of azaindole derivatives with Aurora B inhibitory activities were reported. To explore the relationship between the structures of substituted azaindole derivatives and their inhibition of Aurora B, 3D-QSAR and molecular docking studies were performed on a dataset of 41 compounds. 3D-QSAR, including CoMFA and CoMSIA, were applied to identify the key structures impacting their inhibitory potencies. The CoMSIA model showed better results than CoMFA, with r^2_{cv} value of 0.575 and r^2 value of 0.987. 3D contour maps generated from CoMFA and CoMSIA along with the docking binding structures provided enough information about the structural requirements for better activity. Based on the structure-activity relationship revealed by the present study, we have designed a set of novel Aurora B inhibitors that showed excellent potencies in the developed models. Thus, our results allowed us to design new derivatives with desired activities.

Keywords 3D-QSAR · CoMFA · CoMSIA · Molecular docking · Aurora kinase B

Introduction

The Aurora proteins are a small family of serine/threonine kinases that play critical roles in regulation of the cell cycle,

especially in the later stages from the G2/M checkpoint through the mitotic checkpoint and late mitosis [1–3]. Humans express three Aurora kinase paralogues, Aurora A, B and C; Aurora A and B have distinct roles in mitosis, while the biological role of Aurora C is currently poorly understood [4]. Along with its cellular binding partner TPX2 (target protein of *Xenopus* kinesin-like protein 2), Aurora A plays an essential role in mitotic spindle formation, centrosome maturation, and segregation, while Aurora B is involved in chromosomal condensation, alignment and separation as well as kinetochore-microtubule attachment and cytokinesis with its binding partners INCENP (inner centromere protein) [5–7]. Aurora C is required for spermatogenesis; it may complement the functions of Aurora B, but its function in cell cycle progression is still unknown [8, 9].

Aurora kinases are strongly associated with human cancer, having been reported to be frequently over-expressed in many human cancers, including colorectal, prostate, colon, breast, thyroid cancer and other cancers [10, 11]. Moreover, their elevated expression has been correlated with chromosomal instability and clinically aggressive disease in prostate cancer, and head and neck squamous cell carcinoma [12]. Aurora A over-expression leads to increased degradation of the natural tumor suppressor p53; inhibition of Aurora A causes defects in centrosome separation, with the formation of characteristic monopolar spindles. Over-expression of Aurora B is suggested to induce tumor metastasis; inhibition of Aurora B leads to failure in cytokinesis and abnormal exit from mitosis, ultimately resulting in apoptosis [13–15]. For this reason, Aurora kinases have been indicated as appealing targets for cancer treatment, and a great number of specific inhibitors have been described to date, some of which are under clinical evaluation [16].

P. Lan · W.-N. Chen · P.-H. Sun · W.-M. Chen (✉)
Guangdong Province Key Laboratory of Pharmacodynamic
Constituents of TCM and New Drugs Research, College
of Pharmacy, Jinan University,
Guangzhou 510632, People's Republic of China
e-mail: twmchen@jnu.edu.cn

It is not clear whether inhibition of all three isoforms (Aurora A, B and C) is necessary for optimal anti-tumor efficacy. Previous research has indicated that Aurora B inhibition alone may be sufficient for anti-tumor activity [4, 7]. GSK1070916 (Fig. 1) and a series of its analogues were discovered to function as potent and selective ATP-competitive inhibitors of Aurora B and C [7]. These newly synthesized compounds are reversible ATP-competitive inhibitors and target the Aurora B/INCENP by formation of a rapidly reversible, short-lived enzyme-inhibitor complex [7]. It was presumed by previous research that the azaindole group of these inhibitors interacted with the hinge region of Aurora B/INCENP, and that the pyrazole ring occupied the ATP sugar pocket, while the phenylurea at the C-2 position of pyrazole was essential for the inhibitory selectivity over Aurora A/TPX2; compared to Aurora B/INCENP, the Aurora A/TPX2 has a less accessible hydrophobic pocket at this region [7]. In this paper, 3D-QSAR and molecular docking analysis were performed on a set of GSK1070916 analogues. Partial least square (PLS)-based statistical analysis was carried out on 41 molecules to identify any correlation [17]. The 3D contour maps along with the docking binding structures generated enabled us to explore the structure-activity relationship of these molecules, and guided us to design a series of novel inhibitors with excellent predicted activities.

Materials and methods

Data sets

A set of 41 molecules with reported IC_{50} values for inhibition of Aurora B was available [7]. These values were converted to corresponding pIC_{50} by taking $\text{Log}(1/IC_{50})$. The pIC_{50} values were used as the dependent variables in all the models subsequently developed. The data were divided randomly into a training set of 34 compounds and a test set of 7 compounds. Structures and associated inhibitory activities are shown in Tables 1 and 2.

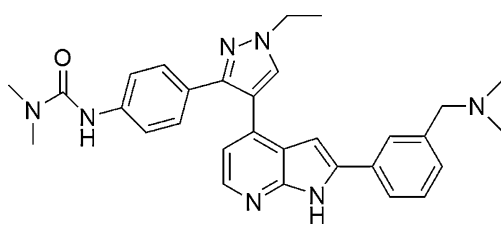


Fig. 1 Chemical structure of GSK1070916

Molecular modeling and alignment

All computational studies were performed using SYBYL 8.1 molecular modeling software from Tripos, Inc. [18]. 3D structures of all compounds were constructed using the Sketch Molecule module. Structural energy minimization was performed using the standard Tripos molecular mechanics force field and Gasteiger-Hückel charge [19, 20]. Pyrrolo[2,3-b]pyridine was selected as the common structure using compound 7—one of the most active compounds in the dataset—as a template. The aligned molecules are shown in Fig. 2.

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 [21]. In the CoMFA method, a sp^3 hybridized carbon atom with a charge of 1e served as the probe atom to calculate steric and electrostatic fields, in which energy values were truncated at 30 kcal mol⁻¹ [22]. As an extension to the CoMFA approach, which has two fields, the CoMSIA method incorporates five different similarity fields, namely steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor. These were applied using a probe atom with radius 1.0 Å, +1.0 charge, and hydrophobic and hydrogen bond properties of +1. The attenuation factor was set to the default value of 0.3 [23].

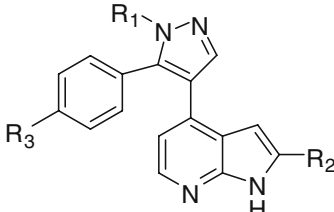
Partial least square analysis

The partial least square (PLS) method was used for all 3D-QSAR analysis. The performance of both the CoMFA and CoMSIA models were evaluated using the leave-one-out (LOO) method, in which one compound is removed from the dataset and its activity is predicted using the model derived from the rest of the dataset [24]. PLS was used in conjunction with the cross-validation option to determine the optimum number of components (ONC), which was then used in deriving the final CoMFA and CoMSIA model without cross-validation. The ONC was the number of components resulting in the highest cross-validated correlation coefficient (r^2_{cv}), which was defined as follows:

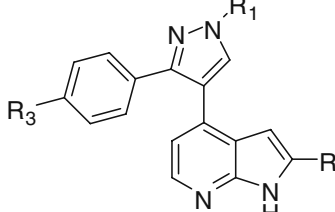
$$r^2_{cv} = 1 - \frac{\sum (Y_{obs} - Y_{pre})^2}{\sum (Y_{obs} - Y_{mean})^2}$$

where Y_{pre} , Y_{obs} and Y_{mean} are predicted, observed, and mean values of the target property (pIC_{50}) [25, 26]. After obtaining the optimum number of components, a PLS analysis was performed with no validation and column filtering 2.0 to generate the highest correlation coefficient (r^2) [27].

Table 1 Structural features and inhibitory activity values of 41 azaindole derivatives



1-3



4-41

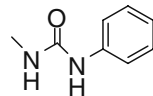
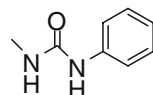
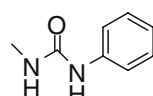
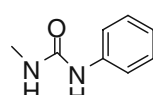
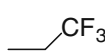
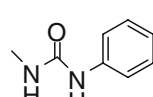
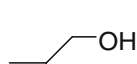
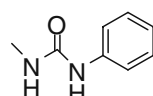
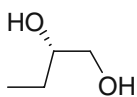
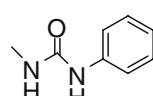
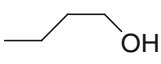
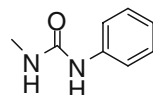
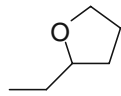
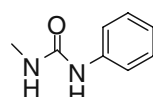
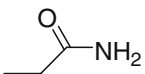
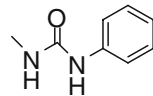
Compd. No.	Substituent			IC ₅₀ (nM)
	R ₁	R ₂	R ₃	
1	H	H		32
2	Et	H		25
3	t-Bu	H		40
4	t-Bu	H		50
5		H		16
6		H		3
7		H		1
8		H		2
9		H		97
10		H		64

Table 1 (continued)

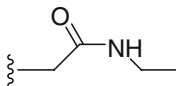
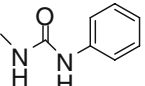
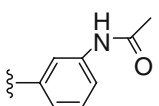
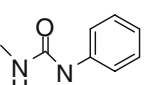
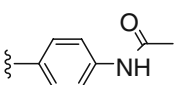
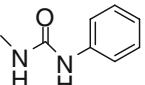
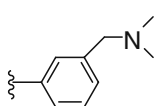
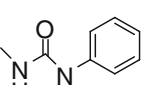
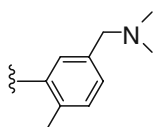
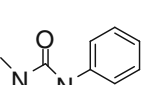
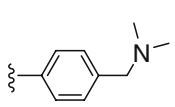
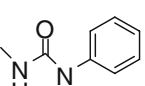
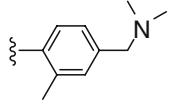
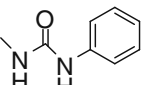
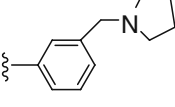
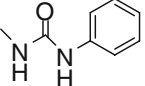
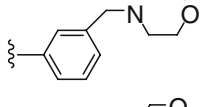
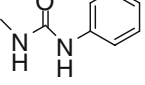
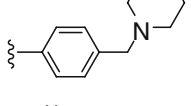
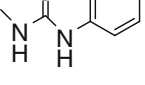
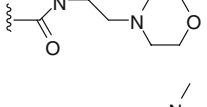
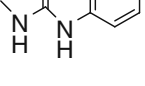
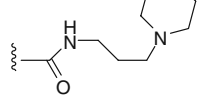
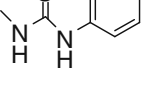
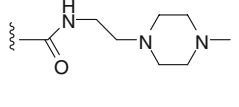
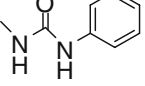

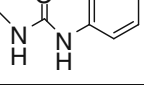
11		H		64
12	Et			44
13	Et			44
14	Et			20
15	Et			60
16	Et			23
17	Et			26
18	Et			11
19	Et			57
20	Et			66
21	Et			5
22	Et			8
23	Et			5
24	Et			58

Table 1 (continued)

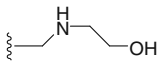
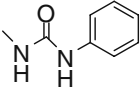
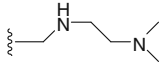
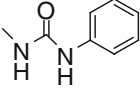
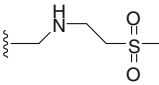
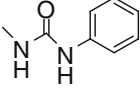
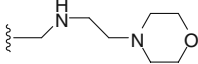
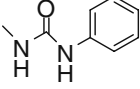
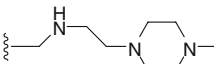
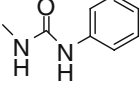
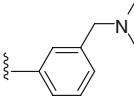
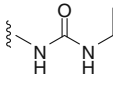
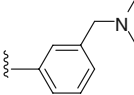
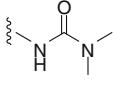
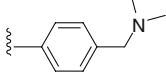
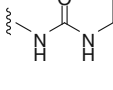
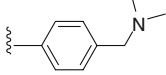
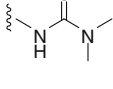
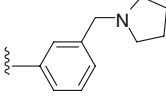
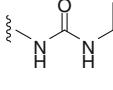
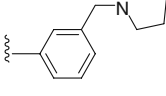
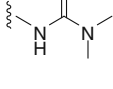
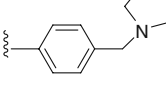
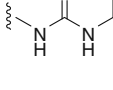
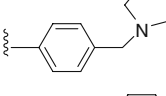
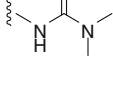
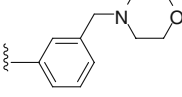
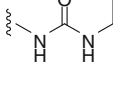
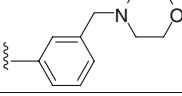
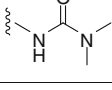
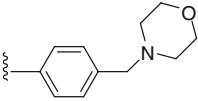
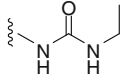
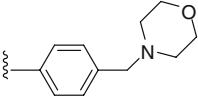
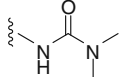
25	Et			46
26	Et			276
27	Et			24
28	Et			22
29	Et			83
30	Et			3
31	Et			5
32	Et			1.7
33	Et			3.6
34	Et			1.4
35	Et			2.1
36	Et			1.5
37	Et			4.5
38	Et			5
39	Et			4.7

Table 1 (continued)

40	Et			4.7
41	Et			8.1

Predictive ability of CoMFA and CoMSIA models

The predictive abilities were determined from a test set of 7 compounds that were not included in the training set. These molecules were aligned to the template and their pIC_{50} values were predicted. The predictive correlation coefficient (r^2_{pred}), based on the molecules of test set, was defined as follows: $r^2_{\text{pred}} = (\text{SD}-\text{PRESS})/\text{SD}$.

Where SD is the sum of the squared deviations between the inhibitory activities of the test set and mean activities of the training molecules, and PRESS is the sum of squared deviations between predicted and actual activity values for each molecule in the test set [28, 29].

Molecular docking

The crystal structure of Aurora B was retrieved from the RCSB Protein Data Bank (PDB entry code: 2VGO). The Aurora B structure was utilized in subsequent docking experiments without energy minimization. All ligands and water molecules were removed and polar hydrogen atoms were added. The Surflex-Dock using an empirical scoring function and a patented search engine to dock ligands into a protein's binding site was applied to study molecular docking [18]. To visualize the binding mode between the inhibitor and ATP pocket, the MOLCAD (molecular computer aided design) program was employed. MOLCAD calculates and displays the surfaces of channels and cavities, as well as the separating the surface between protein subunits [18]. MOLCAD provides several molecular surfaces to investigate the size and volume of proteins. These surfaces can also be used as screens for the color coded display of many molecular properties; in the present work, the electrostatic potential (EP) and lipophilicity (LP) were generated. Other parameters were established by default software settings.

Results and discussion

CoMFA and CoMSIA results

CoMFA and CoMSIA models were derived from a training set of 34 molecules with pIC_{50} values ranging

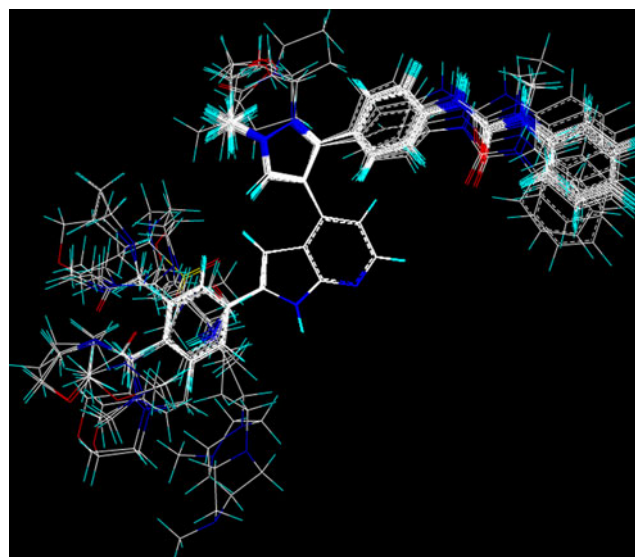
from 6.56 to 9.00. The statistical parameters associated in CoMFA and CoMSIA models are listed in Table 3. The CoMFA model using both steric and electrostatic fields gave a cross-validated correlation coefficient (r^2_{cv}) of 0.549 (> 0.5) with an optimized component of 6. A high non-cross-validated correlation coefficient (r^2) of 0.983 with a low standard error estimate (SEE) of 0.093, F value of 257.043 and predictive correlation coefficient (r^2_{pred}) of 0.759 was obtained. The contributions of steric and electrostatic fields were 0.596 and 0.404, respectively. In comparison to CoMFA, CoMSIA methodology has the advantage of exploring more fields. The best CoMSIA model included steric, electrostatic hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields and gave a r^2_{cv} of 0.575 (> 0.5) with an optimized component of 6. A high non-cross-validated r^2 of 0.987 with a low SEE of 0.081, F value of 342.803 and r^2_{pred} of 0.791 was obtained. Contributions of steric, electrostatic hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields were 0.159, 0.208, 0.218, 0.100 and 0.315, respectively. The correlations between the predicted and experimental pIC_{50} values of all compounds in the CoMFA and CoMSIA models are shown in Fig. 3a,b. The PLS statistics of both CoMFA and CoMSIA models indicated that CoMSIA produced better results than CoMFA.

CoMFA contour map analysis

To visualize the information content of the derived 3D-QSAR model, CoMFA contour maps were generated to rationalize the regions in 3D space around the molecules where changes in the steric and electrostatic fields were predicted to increase or decrease activity. CoMFA steric and electrostatic contour maps are shown in Fig. 4 using compound 34 as a reference structure. In Fig. 4a, the green contours (80% contribution) indicate a steric contribution to potency, while the yellow contours (20% contribution) indicate regions of steric hindrance to activity. In Fig. 4b, the electrostatic field is indicated by blue and red contours, indicating regions where electron-donating group and electron-withdrawing group would be favorable, respectively.

Table 2 The experimental pIC₅₀ values, predicted pIC₅₀ values (*Pred.*) and their residuals (*Res.*) of the training and test set molecules

Compound no.	Experimental	CoMFA		CoMSIA	
		Pred.	Res.	Pred.	Res.
1	7.495	7.470	0.025	7.610	-0.115
2	7.602	7.750	-0.148	7.531	0.071
3	7.398	7.350	0.048	7.442	-0.044
4	7.301	7.411	-0.110	7.371	-0.070
5 ^a	7.796	7.851	-0.055	7.616	0.180
6	8.523	8.420	0.103	8.342	0.181
7	9.000	8.921	0.079	9.123	-0.123
8	8.699	8.752	-0.053	8.672	0.027
9	7.013	6.882	0.131	6.917	0.096
10 ^a	7.194	7.475	-0.281	7.446	-0.252
11	7.194	7.255	-0.061	7.209	-0.015
12	7.357	7.447	-0.091	7.427	-0.071
13	7.357	7.359	-0.003	7.353	0.003
14	7.699	7.667	0.032	7.657	0.042
15	7.222	7.246	-0.024	7.406	-0.184
16	7.638	7.676	-0.038	7.655	-0.017
17	7.585	7.515	0.070	7.544	0.041
18	7.959	7.886	0.073	7.930	0.029
19	7.244	7.209	0.035	7.275	-0.031
20	7.181	7.107	0.074	7.171	0.010
21 ^a	8.301	7.951	0.350	7.849	0.452
22	8.097	8.031	0.066	8.122	-0.025
23	8.301	8.209	0.092	8.324	-0.023
24 ^a	7.237	7.537	-0.300	7.331	-0.094
25	7.337	7.147	0.190	7.162	0.175
26	6.559	6.733	-0.174	6.609	-0.050
27	7.620	7.666	-0.046	7.644	-0.024
28	7.658	7.555	0.103	7.629	0.029
29	7.081	7.198	-0.117	7.075	0.006
30	8.523	8.576	-0.053	8.503	0.020
31 ^a	8.301	8.509	-0.208	8.394	-0.093
32	8.770	8.819	-0.049	8.769	0.001
33 ^a	8.444	8.730	-0.286	8.521	-0.077
34	8.854	8.880	-0.026	8.861	-0.007
35	8.678	8.709	-0.031	8.698	-0.020
36	8.824	8.736	0.088	8.754	0.070
37 ^a	8.347	8.633	-0.286	8.498	-0.151
38	8.301	8.323	-0.022	8.361	-0.060
39	8.328	8.376	-0.048	8.266	0.062
40	8.328	8.403	-0.075	8.341	-0.013
41	8.092	8.129	-0.037	8.063	0.028

^a Test set molecules**Fig. 2** Alignment of the compounds used in the training set

As shown in Fig. 4a, the R₁ position was oriented toward a large green contour indicating that a bulky group would be favored. This may explain why compounds **6–8**, **21–23** and **30–41** which have relatively bulkier groups (e.g., ethyl, hydroxyethyl, dihydroxyethyl, hydroxypropyl) at R₁ showed significantly increased activities compared to compound **1**, which has a hydrogen atom at R₁. The presence of the green contour around the R₂ position suggested that a bulky group at this region would

Table 3 Partial least squares (PLS) result summary of CoMFA and CoMSIA models. *ONC* Optimal number of components, *SEE* Standard error of estimate

PLS statistics	CoMFA	CoMSIA
r^2_{cv} ^a	0.549	0.575
r^2 ^b	0.983	0.987
ONC	6	6
SEE	0.093	0.081
F value	257.043	342.803
r^2_{pred} ^c	0.759	0.791
Field contribution		
Steric	0.596	0.159
Electrostatic	0.404	0.208
Hydrophobic	-	0.218
H-bond donor	-	0.100
H-bond acceptor	-	0.315

^a Leave one out (LOO) cross-validated correlation coefficient^b Non-cross-validated coefficient^c Predictive correlation coefficient for test set compounds

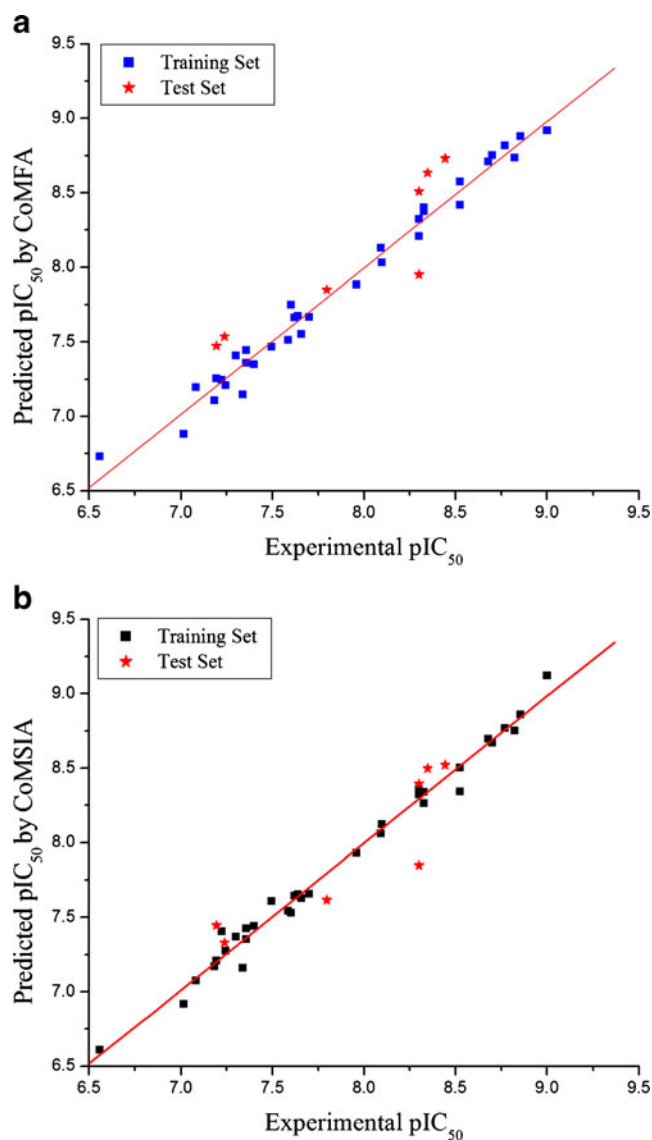


Fig. 3 Graph of actual versus predicted pIC_{50} of the training set and the test set using CoMFA (a) and CoMSIA (b)

be favorable. Comparing compounds **21–23** (pIC_{50} = 8.301, 8.097, 8.301) and **30–41** (pIC_{50} = 8.525, 8.424, 8.624, 8.563, 8.800, 8.682, 8.329, 8.287, 8.312, 8.394, 8.379, 8.183), which possess bulky groups [e.g., *N*-(2-morpholinoethyl)formamide, *N*-(2-(4-methylpiperazin-1-yl)ethyl)formamide, pyrrolylmethyl phenyl] at R_2 with compounds **3**, **4**, **9–11**, and **24–26** (pIC_{50} = 7.833, 7.365, 6.817, 7.347, 7.197, 7.255, 7.139, 6.779), which have relatively minor substituents (e.g., H, dimethylamino-methyl), it was found that their activity discrepancies can be explained easily by this green contour. A huge green contour near the middle of the R_3 side chain (carbamido group) demonstrated that a relatively bulkier substituent at the carbamylamino position would increase activity, while

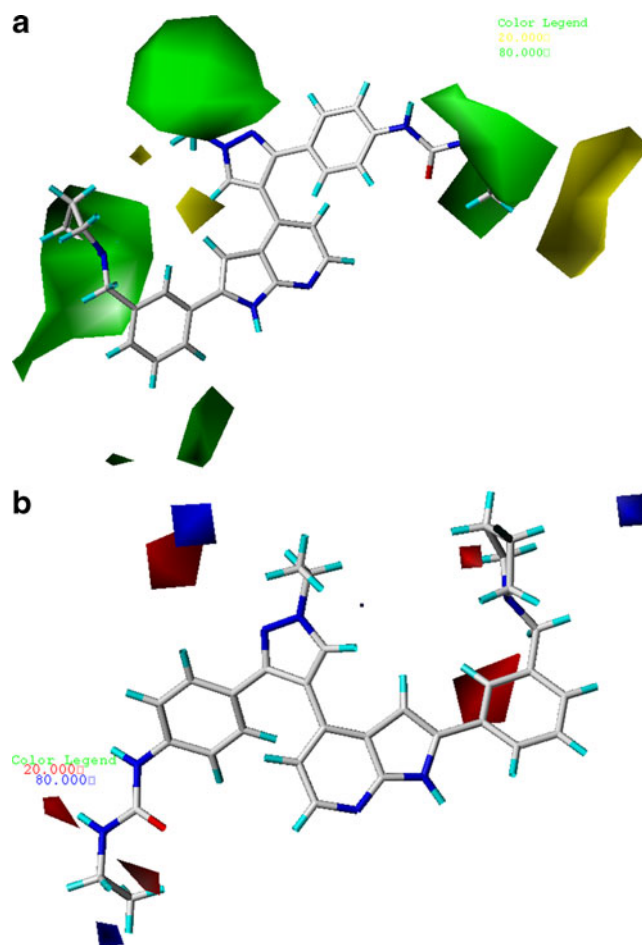


Fig. 4 CoMFA Std* coeff contour maps illustrating steric, electrostatic features in combination with compound **34**. **a** Steric fields: green contours regions where bulky groups increase activity, yellow contours regions where bulky groups decrease activity. **b** Electrostatic fields: blue contours regions where electron-donating groups increase activity, red contours regions where electron-withdrawing groups increase activity

a yellow contour at the terminal of the R_3 side chain indicated that bulky groups would be unfavorable. In general, compounds **30**, **32**, **34**, **36**, **38**, **40** (R_3 = ethyl-carbamylamino) showed better activities than the corresponding compounds **31**, **33**, **35**, **37**, **39**, **41** (R_3 = dimethylcarbamylamino). Meanwhile, compounds **14**, **16**, **18–20** with a bulky phenyl group at the terminus of the R_3 side chain (phenylurea) showed significantly decreased potency compared to the analogous compounds **30–41**. Furthermore, the most inactive compounds **9–11**, **24–26** and **29** all possessed a phenylurea at R_3 ; the reason for their poor potency will be discussed further below.

The CoMFA electrostatic field contour map is shown in Fig. 4b. A red contour and a blue contour near the R_1 position revealed that the electrostatic field was not important for the R_1 position. Most of the active derivatives

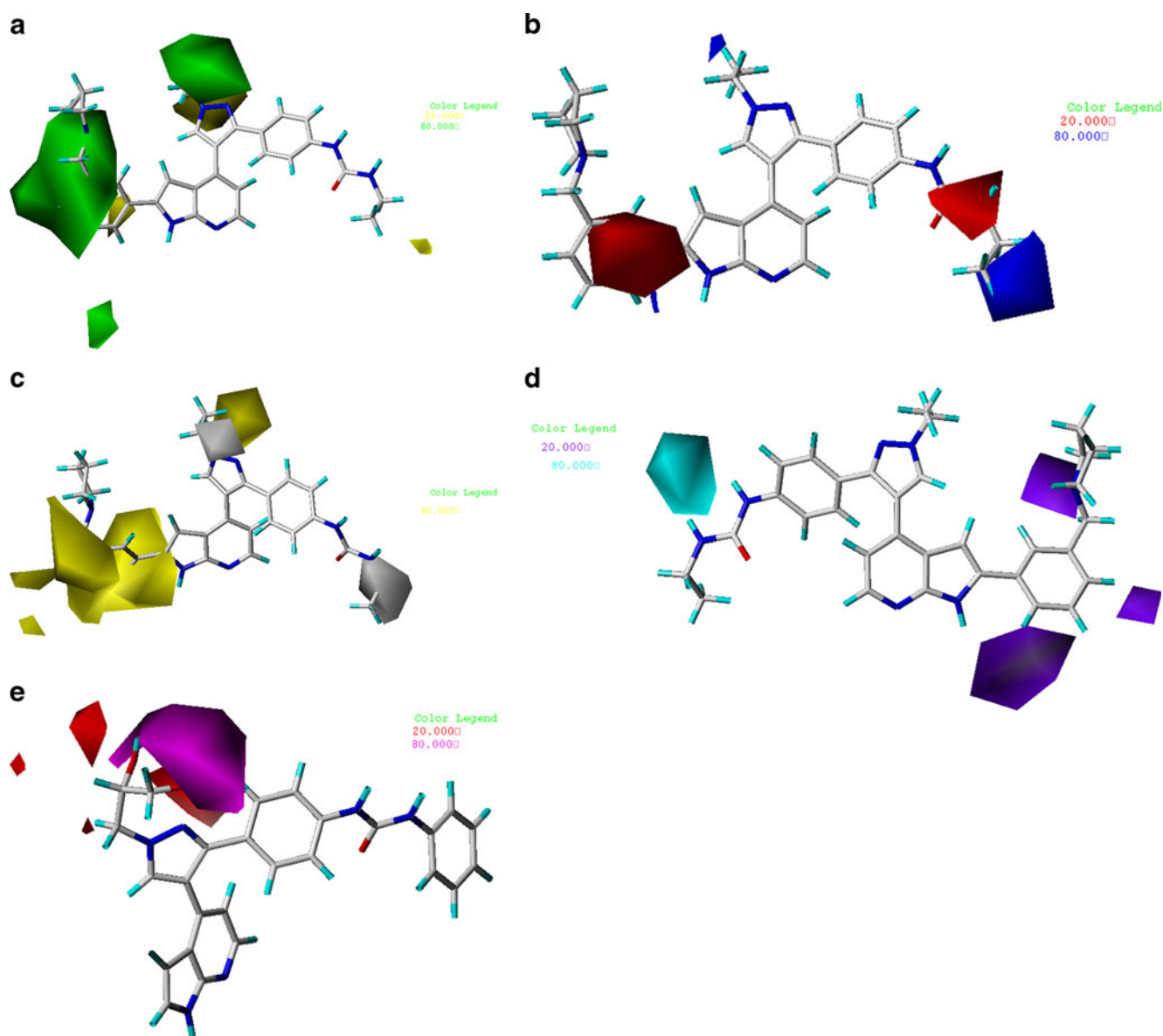


Fig. 5 Std* coeff CoMSIA contour maps illustrating steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor features in combination with compounds **34** and **7**. **a** Steric contour map. *Green contours* Sterically favored regions, *yellow contours* sterically disfavored regions. **b** Electrostatic contour map. *Blue contours* Regions where electron-donating groups are favored, *red contours* regions where electron-withdrawing groups are favored. **c** Hydrophobic

contour map. *White contours* Regions where hydrophilic substituents are favored, *yellow contours* regions where hydrophobic substituents are favored. **d** Hydrogen bond donor contour map. *Cyan and purple contours* indicate favorable and unfavorable hydrogen bond donor groups, respectively. **e** Hydrogen bond acceptor contour map. *Magenta contours* Hydrogen bond acceptor groups increase activity, *red contours* disfavored region

had an electron-donating group (e.g., ethyl, *t*-Bu), but some compounds with an electron-withdrawing group (e.g., hydroxyl, trifluoromethyl) also showed excellent activity. The red contour near the R₂ position indicated that an electron-withdrawing group would be favored. Comparing derivatives **21–23** (pIC₅₀= 8.301, 8.097, 8.301), which had an electron-withdrawing amide group at R₂, with **24–29** (pIC₅₀= 7.255, 7.139, 6.779, 7.712, 7.615, 7.239) with electron-donating alkylamine groups, revealed that their

activity discrepancies can be explained by this red contour. The blue contour around the terminal of R₃ position suggested that an electron-donating substituent at the carbamylamino position would increase activity. This may explain why compounds **30–41**, with an ethyl or methyl group at the carbamylamino position, showed excellent activities, while replacing the ethyl or methyl group with an electron-withdrawing phenyl group (**14**, **16**, **18–20**) showed decreased potencies.

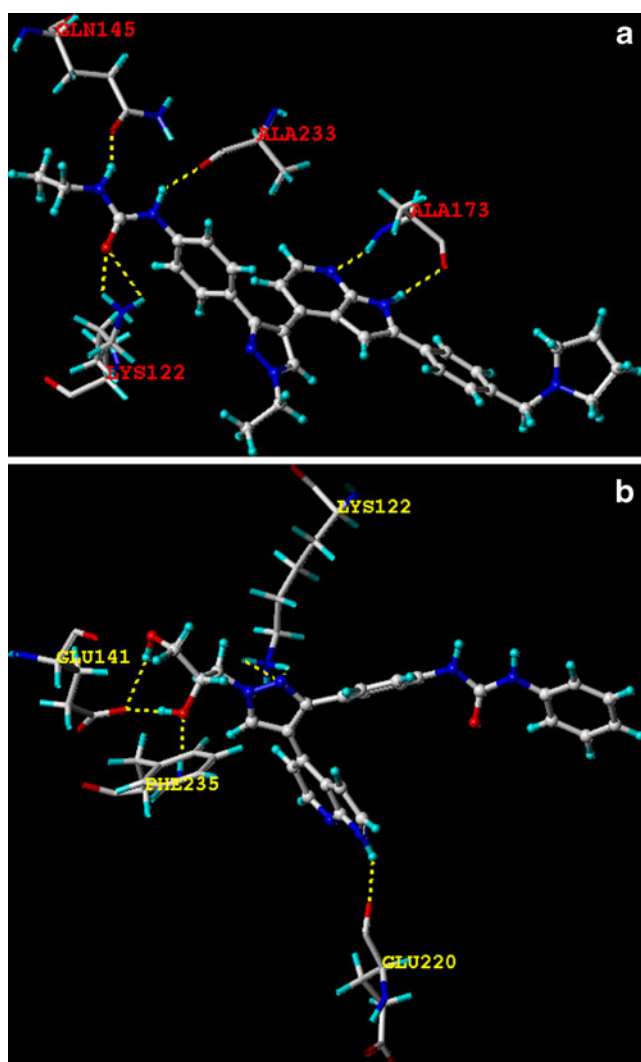


Fig. 6 Binding mode between compound **36** (a) and **7** (b) and the ATP pocket of Aurora B (PDB code 2VGO). Key residues and hydrogen bonds are labeled

CoMSIA contour map analysis

CoMSIA steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor field contour maps are shown in Fig. 5. The CoMSIA steric and electrostatic field contour maps closely resembled the corresponding CoMFA contour maps.

The hydrophobic field contour map shown in Fig. 5c using compound **34** was used as a reference structure; white (20% contribution) and yellow (80% contribution) contours highlighted areas where hydrophilic and hydrophobic properties were favored. A white contour and a yellow contour near the R_1 position revealed that a hydrophobic field for the R_1 position was not important. Four sections of yellow contour around the R_2 position strongly indicated

that a hydrophobic group at this position would be essential. Most of the derivatives possessed a hydrophobic phenyl group at R_2 ; compounds **25–29** with hydrophilic amine substituent at R_2 showed decreased activities. The white contour near the R_3 position revealed the extreme importance of the carbamylamino group in R_3 ; when removed it resulted in decreased activity.

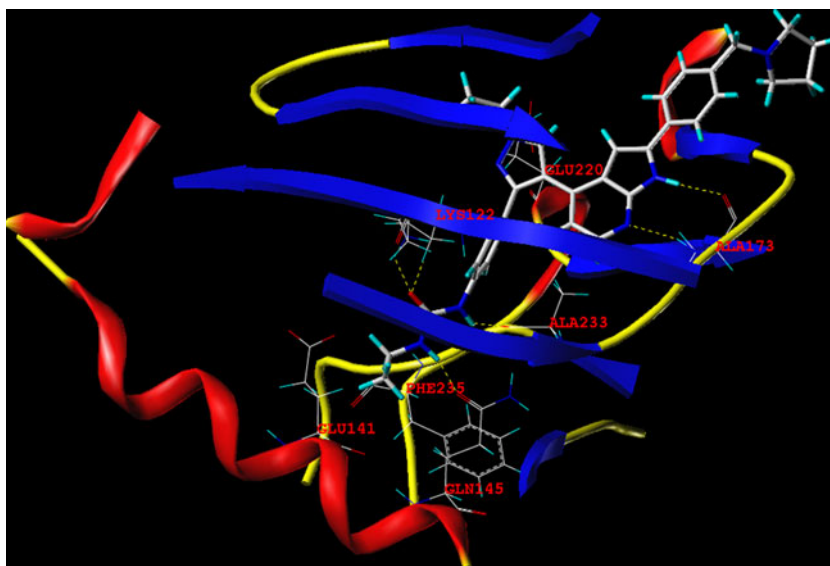
The hydrogen bond donor field contour map shown in Fig. 5d using compound **34** was used as a reference structure. The cyan (80% contribution) and purple (20% contribution) contours indicate favorable and unfavorable hydrogen bond donor groups. The cyan contour near the carbamylamino group of the R_3 position demonstrated the extreme importance of the two $-NH$ groups. In fact, they acted as hydrogen bond donors and formed H-bonds with the residues of Aurora B. Compounds **30, 32, 34, 36, 38,** and **40** (R_3 = ethylcarbamylamino) with two $-NH$ groups that can act as hydrogen bond donors showed relatively increased activities compared to **31, 33, 35, 37, 39,** and **41** (R_3 = dimethylcarbamylamino) with only one $-NH$ group. A purple contour around the N atom of the pyrrolyl group suggested that a hydrogen bond donor at this position would be unfavorable. Derivatives **30–41**, with a tertiary amine at this position, were the most active compounds.

The hydrogen bond acceptor field contour map shown in Fig. 5e using compound **7** was used as a reference structure. The magenta (80% contribution) and red (20% contribution) contours identified favorable and unfavorable positions for hydrogen bond acceptors. A magenta and two red contours near the two hydroxyl groups in the R_1 position of compound **7** revealed that these may act as hydrogen bond acceptor and donor. The huge magenta contour around the N-2 atom of azaindole indicated that the N-2 atom may act as a hydrogen bond acceptor. In fact, the 2'-OH acted as both hydrogen bond donor and acceptor, and formed an H-bond with residue PHE235 and GLU141; the 3'-OH acted as a hydrogen bond donor and formed H-bonds with GLU141. The N-2 atom was a hydrogen bond acceptor and formed an H-bond with LYS122. The CoMSIA hydrogen bond acceptor field contour map was in agreement with the observations subsequently obtained from docking analysis.

Docking analysis

To elucidate the mechanisms of interaction between these inhibitors and the receptor, compounds **7** and **36** were selected for more detailed analysis. These two compounds represented different binding modes of these derivatives. Fig. 6 shows the hydrogen bond binding of compounds **7** and **36** with the ATP pocket of Aurora B. As shown in Fig. 6a, the N atom on pyridyl and the N-H on pyrrolyl groups acted as hydrogen bond acceptor and donor,

Fig. 7 MOLCAD Robbin surfaces structure of selected compound **36** in complex with the ATP-binding site. Alpha helices are shown as *helices* or *cylinders*, while beta sheets are shown as *arrows* and the loop regions as *tubes*. Key residues and hydrogen bonds are labeled



respectively, by forming H-bonds with the $-NH$ and carbonyl groups of ALA173. The two $-NH$ of ethylcarbamylamino formed H-bonds with the carbonyl groups of ALA223 and GLN145, while the carbonyl group of ethylcarbamylamino formed two H-bonds with the $-NH_2$ of LYS122. In Fig. 6b, the situation is different: only the two hydroxyl groups at the R_1 position, the N-2 atom of azaindole and the N-H of pyrrolyl group formed H-bonds with PHE235, GLU141, LYS122 and GLU220, respectively. It can be inferred that the bulky phenyl group at the terminus of the R_3 position can affect the binding between the carbamylamino group and the residues of the ATP pocket, resulting in different binding modes. This may also explain why compounds with a phenylurea group at R_3 position showed relatively decreased potencies compared to those with an ethylcarbamylamino or dimethylcarbamylamino group.

To visualize the secondary structure elements of the binding structure, the MOLCAD Robbin surfaces program was applied. Figure 7 shows the secondary structures of compound **36** in complex with the ATP pocket of Aurora B.

The MOLCAD surface of ATP-binding site was developed and displayed with electrostatic potential to examine the CoMFA electrostatic contour map. Figure 8a shows docking of compound **36** into the ATP-binding site; the red color shows the electron-withdrawing zone and purple color shows the electron-donating zone. The observations taken from Fig. 8a satisfactorily matched those of the CoMFA electrostatic contour map. In detail, the R_2 regions were in the red zone, which suggested that an electron-donating substituent would be favorable; the R_3 position was in a blue zone, which indicated that electron-withdrawing groups may benefit potency.

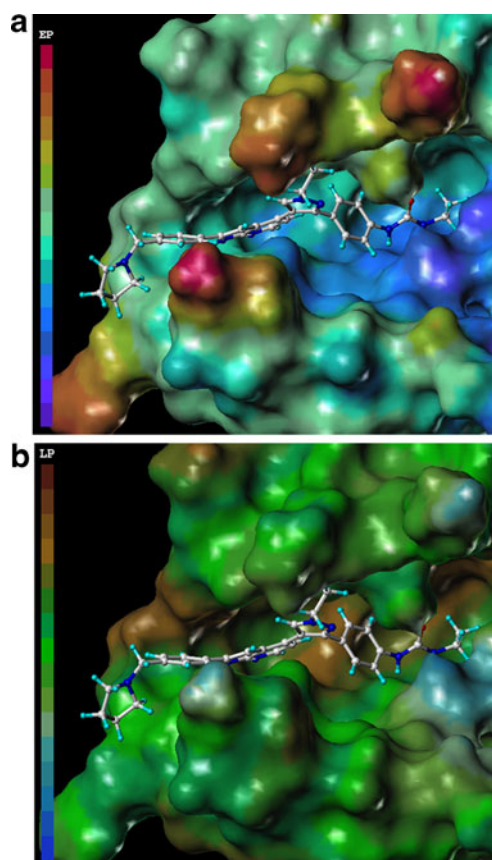
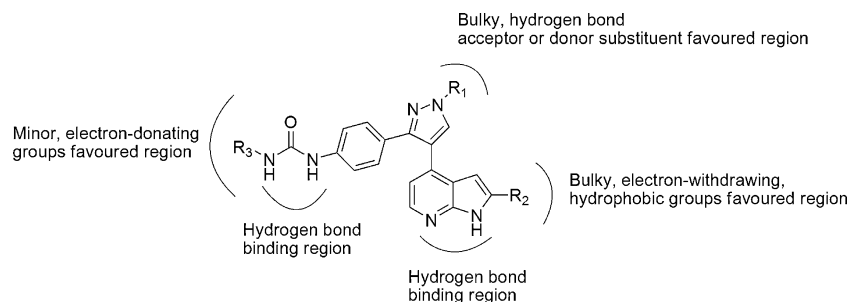


Fig. 8 The MOLCAD electrostatic (a) and lipophilic potential (b) surfaces of the ATP-binding site of Aurora (PDB code 2VGO) within compound **36**. The color ramp for electrostatic potential (EP) ranges from *red* (most positive) to *purple* (most negative). The color ramp for lipophilicity (LP) ranges from *brown* (highest lipophilic area of the molecule) to *blue* (highest hydrophilic area)

Fig. 9 Structure-activity relationship revealed by 3D-QSAR and docking



The MOLCAD surface of the ATP-binding site was also developed and displayed with LP to examine the CoMSIA hydrophobic contour map. The color ramp for LP ranges from brown (highest lipophilic area of the molecule) to blue (highest hydrophilic area). The R_2 side chain was in the green and brown area, which suggested that hydrophobic groups would increase potency; the phenyl group of the C-2 position of azaindole was in a hydrophobically favored brown area, while the urea group of R_3 position was found to be in a hydrophilically favored white area, which demonstrated the extreme importance of the phenylurea group at the C-2 position of azaindole. The observations gained from Fig. 8b were in agreement with those of the CoMSIA hydrophobic contour map, and the predictions of previous research.

Design of novel inhibitors

3D-QSAR and docking studies provided enough information about the structural requirements for better activity. In

detail, bulky and hydrogen bond donor and acceptor groups at the R_1 position increase activity; bulky, electron-withdrawing and hydrophobic substituent are favored at R_2 ; and minor, electron-donating substituents at R_3 benefit potency. Pyrrolo[2,3-b]pyridine and carbamylamino groups were essential for binding to the ATP pocket. The structure-activity relationship revealed by this study is summarized in Fig. 9. Based on this structure-activity relationship, we designed a series of new inhibitors. These designed molecules were aligned in the database, and their pIC_{50} values were predicted by the previously established CoMFA and CoMSIA models. The structures of newly designed derivatives and their predicted pIC_{50} values are shown in Table 4 and Fig. 10. As shown in Table 4, most of the designed derivatives showed better potencies than compounds 7, 34 and 36, which were the most active derivatives in the database. Compounds D1, D3, D4 and D5 were almost 10-fold more active than compound 7 in the CoMSIA model. These results validate the structure-activity relationship obtained from 3D-QSAR and docking studies.

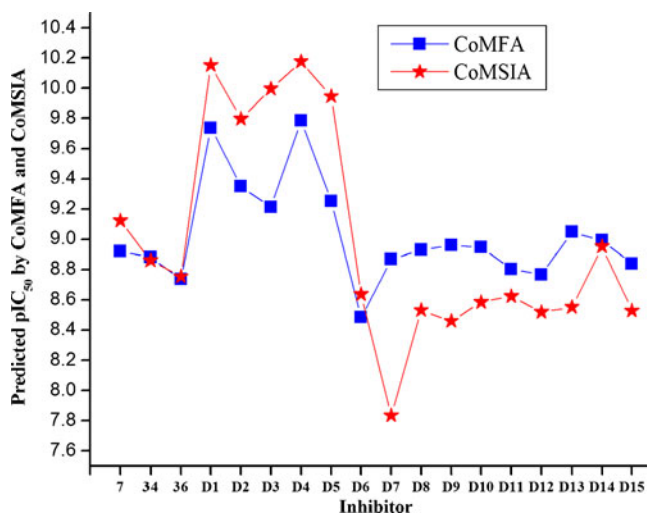


Fig. 10 Predicted pIC_{50} values of designed inhibitors using CoMFA and CoMSIA


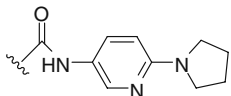

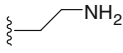
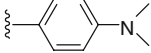
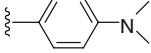
Conclusions

The present study successfully applied 3D-QSAR and molecular docking analysis to characterize a set of recently synthesized Aurora B inhibitors. The CoMFA and CoMSIA models generated both exhibited reliable correlative and predictive abilities. Furthermore, the CoMFA and CoMSIA contour maps, along with the docking binding structures, provided enough information to understand the structure-activity relationship and to identify structural features influencing inhibitory activity. Based on the structure-activity relationship revealed by this study, we successfully designed a set of new inhibitors with excellent predicted activities in CoMFA and CoMSIA models. These results will serve as a useful guideline for designing novel GSK 1070916 analogues with desired activities.

Table 4 Structure and predicted pIC₅₀ values of newly designed inhibitors

No.	Substituent			Predicted pIC ₅₀	
	R ₁	R ₂	R ₃	CoMFA	CoMSIA
D1			Et	9.735	10.153
D2			<i>i</i> -Pr	9.351	9.798
D3			<i>t</i> -Bu	9.212	9.997
D4			Et	9.785	10.178
D5			<i>i</i> -Pr	9.251	9.946
D6			Et	8.486	8.638
D7	Et		Et	8.867	7.834
D8	Et		Et	8.929	8.531
D9	Et		Et	8.959	8.460
D10	Et		Et	8.947	8.585
D11	Et		Et	8.804	8.625
D12	<i>i</i> -Pr		<i>i</i> -Pr	8.768	8.520

Table 4 (continued)

D13				9.048	8.553
D14			Et	8.991	8.953
D15	Et		<i>i</i> -Pr	8.839	8.529

Acknowledgments We gratefully acknowledge support for this research from the Natural Science Foundation of Guangdong Province (No. 9151063201000053) and the Fundamental Research Funds for the Central Universities (No. 21610405), China.

References

- Moriarty KJ, Koblisch HK, Garrabrant T, Maisuria J, Khalil E, Ali F, Petrounia IP, Crysler CS, Maroney AC, Johnson DL, RAG Jr (2006) The synthesis and SAR of 2-amino-pyrrolo[2, 3-d] pyrimidines: a new class of Aurora-A kinase inhibitors. *Bioorg Med Chem Lett* 16:5778–5783
- Bavetsias V, Sun C, Bouloc N, Reynisson J, Workman P, Linardopoulos S, McDonald E (2007) Hit generation and exploration: imidazo[4, 5-b]pyridine derivatives as inhibitors of Aurora kinases. *Bioorg Med Chem Lett* 17:6567–6571
- Zhong M, Bui M, Shen W, Baskaran S, Allen DA, Elling RA, Flanagan WM, Fung AD, Hanan EJ, Harris SO, Heumann SA, Hoch U, Ivy SN, Jacobs JW, Lam S, Lee H, McDowell RS, Oslob JD, Purkey HE, Romanowsky MJ, Silverman JA, Tangonan BT, taverna P, Yang W, Yoburn JC, Yu CH, Zimmerman KM, O'Brein T, Lew W (2009) 2-Aminobenzimidazoles as potent Aurora kinase inhibitors. *Bioorg Med Chem Lett* 19:5158–5161
- Mortlock AA, Foote KM, Heron NM, Jung FH, Pasquet G, Lohmann J-JM, Warin N, Renaud F, Savi CD, Roberts NJ, Johnson T, Dousson CB, Hill GB, Perkins D, Hatter G, Wilkinson RW, Wedge SR, Heaton SP, Odedra R, Keen NJ, Crafter C, Brown E, Thompson K, Brightwell S, Khatri L, Brady MC, Kearney S, McKillop D, Rhead S, Parry T, Green S (2007) Discovery, synthesis, and in vivo activity of a new class of pyrazoloquinazolines as selective inhibitors of Aurora B kinase. *J Med Chem* 50:2213–2224
- Kishore AH, Vadamurthy BM, Mantelingu K, Agrawal S, Reddy BAA, Roy S, Rangappa KS, Kundu TK (2008) Specific small-molecule activator of Aurora kinase A induces autophosphorylation in a cell-free system. *J Med Chem* 51:792–797
- Medina JR, Grant SW, Axten JM, Miller WH, Donatelli CA, Hardwicke MA, Oleykowsky CA, Liao Q, Plant R, Xiang H (2010) Discovery of a new series of Aurora inhibitors through truncation of GSK1070916. *Bioorg Med Chem* 20:2552–2555
- Adams ND, Adams JL, Burgess JL, Chaudhari AM, Copeland RA, Donatelli CA, Drewry DH, Fisher KE, Hamajima T, Hardwicke MA, Huffman WF, Koretke-Brown KK, Lai ZV, McDonald OB, Nakamura H, Newlander KA, Oleykowski CA, Parrish CA, Patrick DR, Plant R, Sarpong MA, Sasaki K, Schmidt SJ, Silva DJ, Sutton D, Tang J, Thompson CS, Tummino PJ, Wang JC, Xiang H, Yang J, Dhanak D (2010) Discovery of GSK1070916, a potent and selective inhibitor of Aurora B/C kinase. *J Med Chem* 53:3973–4001
- Coumar MS, Wu JS, Leou JS, Tan UK, Chang CY, Chang TY, Lin WH, Hsu JTA, Chao YS, Wu SY, Hsieh HP (2008) Aurora kinase A inhibitors: identification, SAR exploration and molecular modeling of 6, 7-dihydro-4H-pyrazolo-[1, 5-a]pyrrolo[3, 4-d]pyrimidine-5, 8-dione scaffold. *Bioorg Med Chem Lett* 18:1623–1627
- Bebbington D, Binch H, Charrier J-D, Everitt S, Fraysse D, Golec J, Kay D, Knechtel R, Mak C, Mazzei F, Miller A, Mortimore M, O'Donnell M, Patel S, Pierard F, Pinder J, Pollard J, Ramaya S, Robinson D, Rutherford A, Studley J, Westcott J (2009) The discovery of the potent aurora inhibitor MK-0457 (VX-680). *Bioorg Med Chem Lett* 19:3586–3592
- Foote KM, Mortlock AA, Heron NM, Jung FH, Hill GB, Pasquet G, Brady MC, Green S, Heaton SP, Kearney S, Keen NJ, Odedra R, Wedge SR, Wilkinson RW (2008) Synthesis and SAR of 1-acetanilide-4-aminopyrazole-substituted quinazolines: selective inhibitors of Aurora B kinase with potent anti-tumor activity. *Bioorg Med Chem* 18:1901–1909
- Li M, Jung A, Ganswindt U, Marini P, Friedl A, Daniel PT, Lauber K, Jendrosseck V, Belka C (2010) Aurora kinase inhibitor ZM447439 induces apoptosis via mitochondrial pathways. *Biochem Pharm* 79:122–129
- Rawson TE, Ruth M, Blackwood E, Burdick D, Corson L, Dotson J, Drummond J, Fields C, Georges GJ, Goller B, Halladay J, Hunsaker T, Kleinheinz T, Krell HW, Li J, Liang J, Limberg A, McNutt A, Moffat J, Phillips G, Ran Y, Safina B, Ultsh M, Walker L, Wiesmann C, Zhang B, Zhou A, Zhu BY, Ruger P, Cochran AG (2008) A pentacyclic Aurora kinase inhibitor (AKI-001) with high in vivo potency and oral bioavailability. *J Med Chem* 51:4465–4475
- Myrianthopoulos V, Magiatis P, Ferandin Y, Skaltsounis AL, Meijer L, Mikros E (2007) An integrated computational approach to the phenomenon of potent and selective inhibition of Aurora kinases B and C by a series of 7-substituted indirubins. *J Med Chem* 50:4027–4037
- Aihara A, Tanaka S, Yasen M, Matsumura S, Mitsunori Y, Murakata A, Noguchi N, Kudo A, Nakamura N, Ito K, Arai S (2010) The selective Aurora B kinase inhibitor AZD1152 as a novel treatment for hepatocellular carcinoma. *J Hepatol* 52:63–71
- Aliagas-Martin I, Burdick D, Corson L, Dotson J, Drummond J, Fields C, Huang OW, Hunsaker T, Kleinheinz T, Krueger E, Liang J, Moffat J, Phillips G, Pulk R, Rawson TE, Ultsh M, Walker L, Wiesmann C, Zhang B, Zhu BY, Cochran AG (2009) A class of 2, 4-bisanilinopyrimidine Aurora A inhibitors with unusually high selectivity against Aurora B. *J Med Chem* 52:3300–3307
- Howard S, Berdini V, Boulstridge JA, Carr MG, Cross DM, Curry J, Devine LA, Early TR, Fazal L, Gill AL, Heathcote M, Maman S, Matthews JE, McMenamin RL, Navarro EF, O'Brien MA, O'Reilly M, Rees DC, Reule M, Tisi D, Williams G, Vinkovic M,

- Wyatt PG (2009) Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent Aurora kinase activity. *J Med Chem* 52:379–388
17. Sivan SK, Manga V (2010) Molecular docking and 3D-QSAR studies on triazolinone and pyridazinone, non-nucleoside inhibitor of HIV-1 reverse transcriptase. *J Mol Model* 16:1169–1178
 18. Sybyl 8.1, Tripos Inc, St. Louis, MO, <http://tripos.com/>
 19. Song QL, Sun PH, Chen WM (2010) Exploring 3D-QSAR for ketolide derivatives as antibacterial agents using CoMFA and CoMSIA. *Lett Drug Des Discovery* 7:149–159
 20. Zhang N, Zhong R (2010) Docking and 3D-QSAR studies of 7-hydroxycoumarin derivatives as CK2 inhibitors. *Eur J Med Chem* 45:292–297
 21. Ibrahim DA, El-Metwally AM (2010) Design, synthesis, and biological evaluation of novel pyrimidine derivatives as CDK2 inhibitors. *Eur J Med Chem* 45:1158–1166
 22. Hu R, Barbault F, Delamar M, Zhang R (2009) Receptor- and ligand-based 3D-QSAR study for a series of non-nucleoside HIV-1 reverse transcriptase inhibitors. *Bioorg Med Chem* 17:2400–2409
 23. Murumkar PR, Gupta SD, Zambre VP, Giridhar R, Yadav MR (2009) Development of predictive 3D-QSAR CoMFA and CoMSIA models for β -aminohydroxamic acid-derived tumor necrosis factor- α converting enzyme inhibitors. *Chem Biol Drug Des* 73:97–107
 24. Lu XY, Chen YD, Sun NY, Jiang YJ, You QD (2010) Molecular-docking-guided 3D-QSAR studies of substituted isoquinoline-1, 3-(2H, 4H)-diones as cyclin-dependent kinase (CDK4) inhibitors. *J Mol Model* 16:163–173
 25. Sun J, Cai S, Yan N, Mei H (2010) Docking and 3D-QSAR studies of influenza neuraminidase inhibitors using three-dimensional holographic vector of atomic interaction field analysis. *Eur J Med Chem* 45:1008–1014
 26. Pan J, Liu GY, Cheng J, Chen XJ, Ju XL (2010) CoMFA and molecular docking studies of benzoxazoles and benzothiazoles as CYP450 1A1 inhibitors. *Eur J Med Chem* 45:967–972
 27. Pan X, Tan N, Zeng G, Huang H, Yan H (2010) 3D QSAR studies on ketoamides of human cathepsin K inhibitors based on two different alignment methods. *Eur J Med Chem* 45:667–681
 28. Cichero E, Cesarini S, Mosti L, Fossa P (2010) CoMFA and CoMSIA analyses on 4-oxo-1, 4-dihydroquinoline and 4-oxo-1, 4-dihydro-1, 5-, -1, 6- and -1, 8-naphthyridine derivatives as selective CB2 receptor agonists. *J Mol Model* 16:677–691
 29. Li Y, Wang YH, Yang L, Zhang SW, Liu CH, Yang SL (2005) Comparison of steroid substrates and inhibitors of P-glycoprotein by 3D-QSAR analysis. *J Mol Struct* 733:111–118