Molecular docking-based 3D-QSAR studies of pyrrolo[3,4-c]pyrazole derivatives as Aurora-A inhibitors

G. He[a]; M. H. Qiu[b]; R. Li[a]; X. R. Song[a]; X. Zheng[a]; J. Y. Shi[a]; G. B. Xu[a]; J. Han[a]; L. T. Yu[a]; S. Y. Yang[a]; L. J. Chen[a]; Y. Q. Wei[a]

[a] State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, P.R. China [b] Institute of Botany, The Chinese Academy of Sciences, Kunming, P.R. China

Online publication date: 28 January 2011

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

Molecular docking-based 3D-QSAR studies of pyrrolo[3,4-c]pyrazole derivatives as Aurora-A inhibitors

G. He[a], M.H. Qiu[b], R. Li[a]*, X.R. Song[a], X. Zheng[a], J.Y. Shi[a], G.B. Xu[a], J. Han[a], L.T. Yu[a], S.Y. Yang[a], L.J. Chen[a] and Y.Q. Wei[a]

[a]*State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, P.R. China;* [b]*Institute of Botany, The Chinese Academy of Sciences, Kunming 650204, P.R. China*

In recent years, Aurora kinases have been highlighted as attractive targets for the development of novel anti-cancer agents. To find the correlation between Aurora-A and its inhibitors, structure-based 3D-quantitative structure–activity relationship (QSAR) models were performed on a series of pyrrolo[3,4-c]pyrazole derivatives with comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) methods. Based on the docking results, predictive 3D-QSAR models were established, with cross-validated coefficient values (r_{cv}^2) up to 0.667 for CoMFA and 0.664 for CoMSIA, respectively. Furthermore, the CoMFA and CoMSIA models were mapped back to the binding sites of Aurora-A, which could get a better understanding of vital interactions between the inhibitors and the kinase. Ligands binding in the ATP pocket and the hydrogen bonds with Ala213 and Glu211 were found to be crucial for the potent ligand binding and kinases selectivity. Therefore, these results demonstrated the power of combining the docking/QSAR approach to explore the probable binding conformations of compounds at the active sites of the protein, and further provided useful information for designing new compounds that showed very low binding free energy against the Aurora-A kinase, and which had been shifted for further experimental assay studies.

Keywords: pyrrolo[3,4-c]pyrazole derivatives; 3D-QSAR; molecular docking; CoMFA; CoMSIA

1. Introduction

The Aurora kinase family is a small family of serine/threonine kinase that regulates some important events during mitosis. They play key roles in centrosome maturation and separation, mitotic spindle assembly and chromosome segregation [1]. There are three isoforms of Aurora kinase in mammals, Aurora A, B and C. The Aurora kinase is commonly overexpressed in many tumour cell lines and human primary tumours [2]. Aurora A is localised at the centrosome and spindle poles from late S and early G2 through M phase [3]. In addition, Aurora A has the ability to transport cell lines that are then capable of forming tumours in mice [4]. The role of Aurora A in the cell cycle and tumorigenesis leads to the suggestion that the inhibition of the kinase activity of Aurora A may have remarkable value for the development of small molecular therapeutics for cancer.

In recent years, a number of small molecular inhibitors of Aurora kinase, with wide structural diversity, have been reported. Among these compounds, clinical trials have shown that ZM447439, VX-680, PHA-739358 and other potential inhibitors suppress the growth of tumour cells both *in vitro* and *in vivo*. ZM447439 is a novel quinazoline derivative that inhibits Aurora A and B kinases with IC$_{50}$ values of 110 and 130 nM, respectively [5]. Moreover, this inhibitor is relatively specific to the Aurora kinase family, which shows almost no inhibition of a range of kinases including CDK1 and FLK1 [6]. VX-680 is a 4,6-di-amino pyrimidine that inhibits Aurora A, B and C *in vitro* with inhibition constants of 0.6, 18 and 4.6 nM, respectively, but does not appear to inhibit the majority of other kinases tested with the exception of FLT-3, which is inhibited with a K_i of 30 nM [7]. PHA-739358 is a small molecule pyrrolo[3,4-c]pyrazole derivative with strong activity against Aurora kinases and cross-activities against some receptor tyrosine kinases relevant to cancer. PHA-739358 inhibits all Aurora kinase family members and shows a dominant Aurora kinase inhibition-related cellular phenotype and mechanism of action in cells *in vitro* and *in vivo* [8].

On the basis of our previous research results [9–12], we have explored the binding mode of 3-amino-pyrrolo[3,4-c]pyrazole bicycle inhibitors against Aurora-A, using a molecular docking approach. Based on the docking results, 3D-quantitative structure–activity relationship (QSAR) models are constructed using comparative molecular field analysis (CoMFA) [13] and comparative molecular similarity indices analysis (CoMSIA) [14] methodologies. The main focus of the current study is to demonstrate the common binding mode of pyrrolo[3,4-c]pyrazole

*Corresponding author. Email: li_rui@yahoo.cn

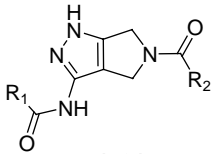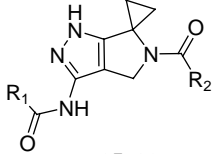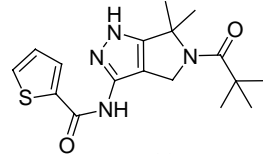Table 1.   Structures, experimental and calculated pIC$_{50}$ values of Aurora-A inhibitors used in this study.
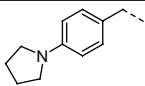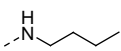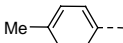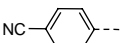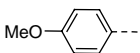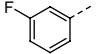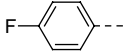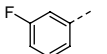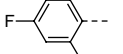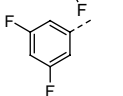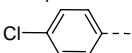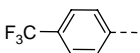
1–24   **25–32**   **33**

Compds	R$_1$	R$_2$	Aurora-A pIC$_{50}$	CoMFA predicted	CoMFA residue	CoMSIA predicted	CoMSIA residue
1	*t*-Bu-phenyl	thiophene-CH$_2$	8.30	8.07	− 0.23	8.12	− 0.18
2	*t*-Bu-phenyl	furan	7.39	7.20	− 0.19	7.25	− 0.14
3[a]	*t*-Bu-phenyl	benzyl	6.89	7.03	0.14	6.69	− 0.20
4[a]	*t*-Bu-phenyl	2,6-diethylanilino-NH	7.00	7.14	0.14	7.03	0.03
5	piperazinyl-phenyl	thiophene-CH$_2$	7.19	7.20	0.01	7.23	0.04
6	piperazinyl-phenyl	furan	6.80	6.98	0.18	6.76	− 0.06
7	piperazinyl-phenyl	benzyl	6.85	6.78	− 0.07	6.92	0.07
8	piperazinyl-phenyl	2,6-diethylanilino-NH	7.57	7.51	− 0.06	7.60	0.03
9[a]	piperazinyl-phenyl	CH(Me)(H)-phenyl	7.62	7.59	− 0.03	7.55	− 0.07
10[a]	piperazinyl-phenyl	CH(HO)(H)-phenyl	8.22	8.15	− 0.07	8.23	0.01
11	piperazinyl-phenyl	CH(F)(H)-phenyl	8.05	8.05	0.00	8.02	− 0.03
12	piperazinyl-phenyl	CH(H)(OMe)-phenyl	6.45	6.48	0.03	6.51	0.06
13	piperazinyl-phenyl	CH(MeO)(H)-phenyl	7.89	7.80	− 0.09	7.83	− 0.06
14	piperazinyl-phenyl	CH(H)(Me)-phenyl	6.35	6.48	0.13	6.41	0.06
15	Me-phenyl	quinoline	6.36	6.32	− 0.04	6.20	− 0.16
16[a]	methylenedioxyphenyl	NH-isopropyl	6.41	6.79	0.38	6.19	− 0.22
17	methylenedioxyphenyl	isopropyl	6.80	6.80	0.00	6.78	− 0.02
18	benzyl	Me	5.63	5.56	− 0.07	5.63	0.00

Table 1 – *Continued*

Compds	R$_1$	R$_2$	Aurora-A pIC$_{50}$	CoMFA predicted	CoMFA residue	CoMSIA predicted	CoMSIA residue
19	pyrrolidinyl-phenyl-CH$_2$	HN-butyl	5.38	5.30	−0.08	5.37	−0.01
20[a]	phenyl	Me	5.51	5.46	−0.05	5.57	0.06
21	Me-phenyl	Me	5.67	5.65	−0.02	5.67	0.00
22	NC-phenyl	Me	5.07	5.06	−0.01	5.08	0.01
23	MeO-phenyl	Me	6.18	6.08	−0.10	6.17	−0.01
24	F-phenyl	Me	5.31	5.24	−0.07	5.35	0.04
25[a]	F-phenyl	tert-butyl	6.47	6.61	0.14	6.18	−0.29
26	F-phenyl	tert-butyl	6.07	6.15	0.08	6.00	−0.07
27[a]	F,F-phenyl	tert-butyl	5.43	5.38	−0.05	5.73	0.30
28	F,F,F-phenyl	tert-butyl	5.50	5.63	0.13	5.63	0.13
29	Cl-phenyl	tert-butyl	6.11	5.95	−0.16	5.84	−0.27
30	F$_3$C-phenyl	tert-butyl	5.70	5.67	−0.03	6.11	0.41
31	thienyl	tert-butyl	7.10	6.68	−0.42	6.86	−0.24
32	cyclopropyl	tert-butyl	6.62	6.67	0.05	6.31	−0.31
33	–	–	6.03	6.26	0.23	5.85	−0.18

[a] Compounds of the test set.

derivatives with Aurora-A, and to validate and predict the accuracy of biological IC$_{50}$ values of small molecular inhibitors against Aurora-A, using molecular docking and CoMFA and/or CoMSIA in combination.

2. Computational details

Table 1 shows a series of Aurora-A inhibitors originally reported by Fancelli et al. during 2004–2007 [15–18], which are divided into a training set and a test set. The training set comprises 25 compounds. The test set consists of eight compounds, randomly selected and structurally diverse molecules possessing activities of a wide range. The *in vitro* IC$_{50}$ values employed in this work were measured under the same experimental conditions, a fundamental requirement for QSAR studies [19,20]. These compounds' inhibitory activities are converted into the corresponding pIC$_{50}$ values ($-\log$ IC$_{50}$) in the CoMFA and CoMSIA.

All of the compounds were built by using the molecular modelling software package Sybyl6.9 [19] and then minimised using Tripos force field [20] and the Gasteiger charge with a distance-dependent dielectric and conjugate gradient method. The optimised structures were used for all subsequent calculations. The X-ray crystal structure of the Aurora-A kinase in complex with inhibitor PHA-680626 was obtained from the Brookhaven Protein Data Bank (entry 2J4Z). The FlexX program [21] interfaced with Sybyl was used to dock the compounds to the binding site of Aurora-A. The active site was defined as including all atoms within a 6.5 Å radius of the co-crystallised ligand. For docking with FlexX, the most critical interactions between ligands and Aurora-A kinase were mapped in X-ray crystal structures, and the default Sybyl/FlexX parameters were used. All the molecules in the training and test sets were docked into the binding site, and among the 30 possible docking conformations generated by the FlexX program, the conformations that were indicated as the most tightly bound by docking scoring were aligned together inside the active site and used directly for CoMFA and CoMSIA.
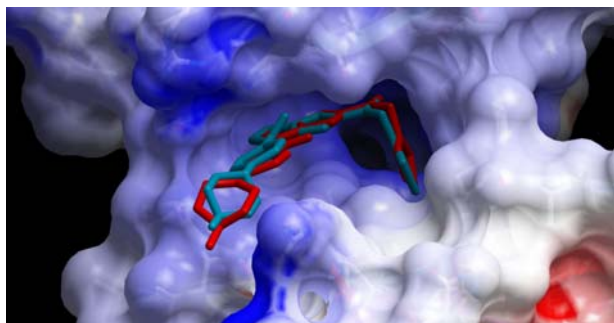
Figure 1. Superimposition of docked (blue) and X-ray crystallographic (red) conformations of compound **5** in active sites of Aurora-A kinase (colour online).

After acquiring the highest percentage of conformation in the active sites of the Aurora-A kinase, five different kinds of partial charges are considered: (1) Gasteiger charges, (2) Gasteiger–Hückel charges, (3) Hückel charges, (4) Pullman charges and (5) MMFF94 charges. All charge calculations in Sybyl are automated with SPL scripts.

To build predictive QSAR models for designing selective inhibitors against Aurora-A, CoMFA and CoMSIA are performed based on the binding confor-mations derived from the molecular docking simulations. Steric and electrostatic interactions are calculated using the Tripos force field with a distance-dependent dielectric constant at all intersections in a regular space (from 1.0 to 3.0 Å) grid, taking an sp^3 carbon atom as the steric probe and $a + 1$ charge as the electrostatic probe. The cut-off

is set to 30 kcal/mol. With standard options for the scaling of variables, the regression analysis is carried out using the full cross-validated partial least-squares (PLS) method of leave-one-out (LOO). The column filtering value is set from 0 to 3.0 kcal/mol to improve the signal-to-noise ratio, by omitting those lattice points whose energy variation was below this threshold. The final model, non-cross-validated conventional analysis, is developed with the optimum number of components to yield a non-cross-validated r^2 value. The CoMSIA method defines five fields: steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor; they are calculated at each lattice interactions of a regularly spaced grid from 1.0 to 3.0 Å. A probe atom with radius 1.0 Å, $+1$ charge, hydro-phobicity $+1.0$ and H-bond donor and acceptor properties of $+1.0$ is used to calculate steric, electrostatic, hydrophobic and H-bond donor and acceptor fields. A distance-dependent Gaussian-type functional form will take into account abrupt changes of potential energy near the molecular surface. The default value of 0.3 is used as the attenuation factor.

3.   Results and discussions

In order to determine the probable binding conformations of these Aurora-A inhibitors, the FlexX package is used to dock all compounds into the active sites of Aurora-A kinase. The docking reliability is validated using the known X-ray structure of Aurora-A kinase in complex with a small molecular ligand PHA-680626 (Figure 1). The ligand is redocked to the binding sites of the protein.

Figure 2.   Three H-bonds (as highlighted by the dashed lines in green colour) between compound **5** and Aurora-A kinase (colour online).

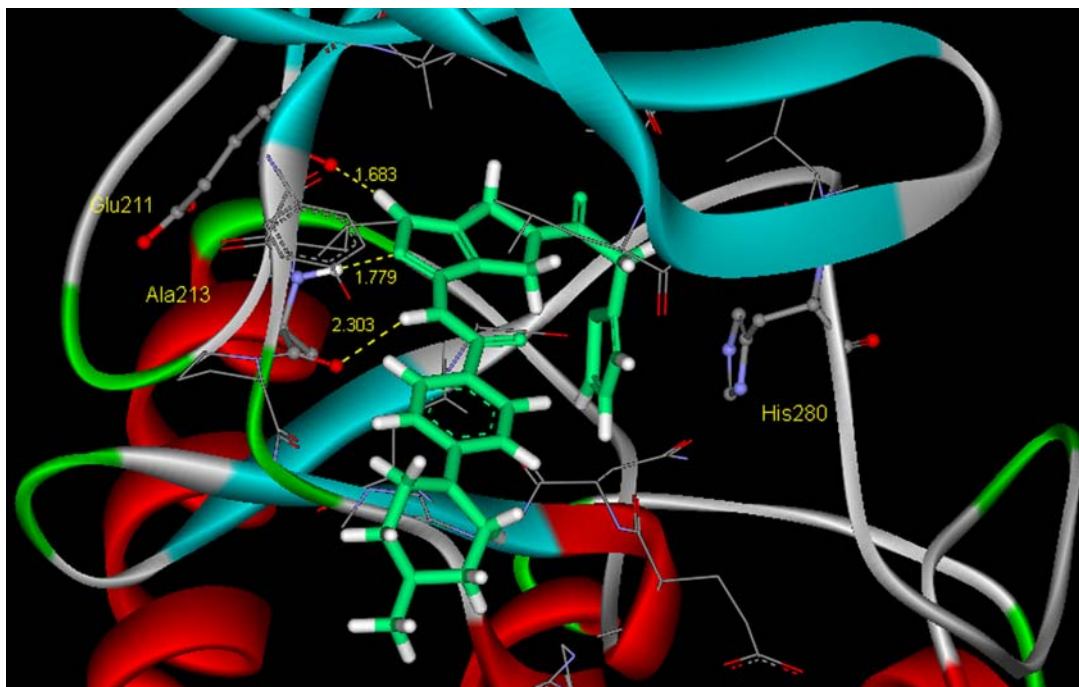All of the inhibitors are bound in the active sites of Aurora-A in a similar conformation of PHA-680626 in the X-ray structure co-crystallised with Aurora-A (Figure 2), and the common chain structures superimposed each other rather well. Figure 3 illustrates the probable binding conformational alignment for the 33 compounds chosen from the docked conformations according to their docking scores, and shows the 3D model of kinase-inhibitor complexes. Based on this set of binding conformations and their alignment, CoMFA and CoMSIA are performed.

The chemical structures of the molecules and their actual pIC_{50} values are shown in Table 1. The predictive power of the 3D-QSAR model which is derived from the training set is assessed by predicting biological activities of the test set molecules. From Table 1, we can see that almost all compounds in test sets yield a good predicted pIC_{50} within 0.5 log unit of the experimental value. The CoMFA and CoMSIA 3D-QSAR methods are based on the assumption that the changes in binding affinities of ligands are related to changes in molecular properties represented by fields. The alignment rule and the bioactive conformation are crucial variables in any 3D-QSAR analysis, as both will influence the outcome of statistical analysis.

Steric and electrostatic CoMFA fields are generated using standard procedures. The different sets of partial charges are used in building the CoMFA models, and all of them exhibit good statistical quality between the predicted and experimentally determined values of pIC_{50}. The statistical details are summarised in Table 2. The q^2 values for Gasteiger–Hückel, Hückel and Pullman are similar (0.618, 0.613 and 0.611); it is better in the case of Gasteiger (0.640) and highest in the case of MMFF94 (0.667) partial charges, respectively. Based on the above observations, the best CoMFA model obtained using MMFF94 partial charge is chosen for further analysis.

In the development of QSAR models with CoMFA, a shift in the q^2 values is observed as the grid spacing is altered [21,22]. To examine this possibility with these data, the different grid boxes with 1.0, 1.5, 2.0, 2.5 and 3.0 Å

Figure 3. Superimposition of 33 Aurora-A inhibitors at the active sites of Aurora-A kinase.

The predicted conformation of PHA-680626 is shown in Figure 1 with the superposition of an X-ray crystallographic one in active sites of Aurora-A. The root mean square deviation between these two conformations is equal to 0.96 Å, suggesting a high docking reliability of FlexX in reproducing the experimentally observed binding mode for Aurora-A. This method and the parameter set can be extended to search the binding conformations of Aurora-A kinase for other inhibitors accordingly.

Table 2. Influence of different partial charges on the CoMFA models.

	CoMFA1	CoMFA2	CoMFA3	CoMFA4	CoMFA5
Partial charge	Gasteiger	Gasteiger–Hückel	Hückel	Pullman	MMFF94
q^{2a}	0.640	0.618	0.613	0.611	0.667
r^{2b}	0.965	0.965	0.958	0.962	0.966
SEE[c]	0.186	0.187	0.205	0.195	0.187
F-value[d]	120.4	118.7	97.9	109.4	126.5
N^{e} contributions	3	3	3	3	3
Steric	0.754	0.730	0.748	0.746	0.646
Electronic	0.246	0.270	0.252	0.254	0.354
Grid spacing: 2.0 Å					

[a] Cross-validated correlation coefficient. [b] Non-cross-validated correlation coefficient. [c] Standard error of estimate. [d] F-test value. [e] Optimum number of components obtained from cross-validated PLS analysis and the same used from final non-cross-validated analysis.

Table 3.   Influence of different grid spacings on the CoMFA models.

	CoMFA1	CoMFA2	CoMFA3	CoMFA4	CoMFA5
Grid spacing (Å)	1.0	1.5	2.0	2.5	3.0
q^{2a}	0.662	0.672	0.667	0.626	0.621
r^{2b}	0.958	0.957	0.966	0.942	0.939
SEE^c	0.205	0.206	0.187	0.241	0.246
F-value^d	98.1	97.2	126.5	69.8	66.5
N^e contributions	4	4	3	3	4
Steric	0.610	0.615	0.646	0.589	0.626
Electronic	0.390	0.385	0.354	0.411	0.372

^{a} Cross-validated correlation coefficient. ^{b} Non-cross-validated correlation coefficient. ^{c} Standard error of estimate. ^{d} F-test value. ^{e} Optimum number of components obtained from cross-validated PLS analysis and the same used from final non-cross-validated analysis.

Table 4.   Influence of different column filtering values on the CoMFA models.

	CoMFA1	CoMFA2	CoMFA3	CoMFA4	CoMFA5	CoMFA6
Column filtering (kcal/mol)	0	1.0	1.5	2.0	2.5	3.0
q^{2a}	0.646	0.644	0.645	0.667	0.650	0.657
r^{2b}	0.961	0.960	0.959	0.966	0.958	0.959
SEE^c	0.197	0.200	0.202	0.187	0.204	0.202
F-value^d	106.5	103.3	101.7	126.5	98.6	101.0
N^e contributions	3	3	3	3	3	3
Steric	0.656	0.717	0.702	0.646	0.695	0.704
Electronic	0.344	0.283	0.298	0.354	0.305	0.296

^{a} Cross-validated correlation coefficient. ^{b} Non-cross-validated correlation coefficient. ^{c} Standard error of estimate. ^{d} F-test value. ^{e} Optimum number of components obtained from cross-validated PLS analysis and the same used from final non-cross-validated analysis.

grid spacing, respectively, are used for the CoMFA calculations. The influence of the different grid spacings on the CoMFA model is very significant (Table 3). From the q^2 value after LOO cross-validation, the standard estimated error (SEE) and the F values, the model with the grid spacing of 2.0 Å is selected as the best model. When the grid spacing is defined as a larger value such as 2.5 or 3.0 Å, some important information about the field properties in some regions may be lost. Also, a lower grid spacing (1.0 or 1.5 Å) may generate more noise in the PLS calculations and require a greater computational effort. The standard error of 2.0 Å grid spacing is the smallest in the five results. The following discussion will only refer to the model CoMFA generated from 2.0 Å grid spacing.

To explore the effect of column filtering on the CoMFA model, different column filtering values are used (with 2.0 Å grid spacing and MMFF94 partial charge) [23]. The column filtering is set to 0, 1.0, 1.5, 2.0, 2.5 and 3.0 kcal/mol (Table 4), respectively, to improve the signal-to-noise ratio by omitting those lattice points on which the energy variation is below this threshold. The results show that when the value of column filtering is bigger than 2.0 kcal/mol, the q^2 value decreases from 0.667 to 0.657 and the standard error is also bigger than that of 2.0 kcal/mol, but they need less computational effort. When the column filtering is set to 0, 1.0 and 1.5 kcal/mol, it needs more time to run. A typical value of 2.0 decreases

the analysis time about an order of magnitude, with typically small effects on the q^2 values obtained.

Figure 4 shows the experimental activities vs. predicted ones in the training and test sets by this CoMFA model. The CoMFA contour plots of steric and electrostatic interactions are shown in Figure 5. As shown in Figure 5(a), there is a major green region and a major yellow region near the thiazole ring of compound **5**, one

Figure 4.   Correlation between predicted activities by CoMFA models and the experimental pIC$_{50}$ values of training and test sets, filled circles represent predictions for the training set, while open circles represent predictions for the test set.
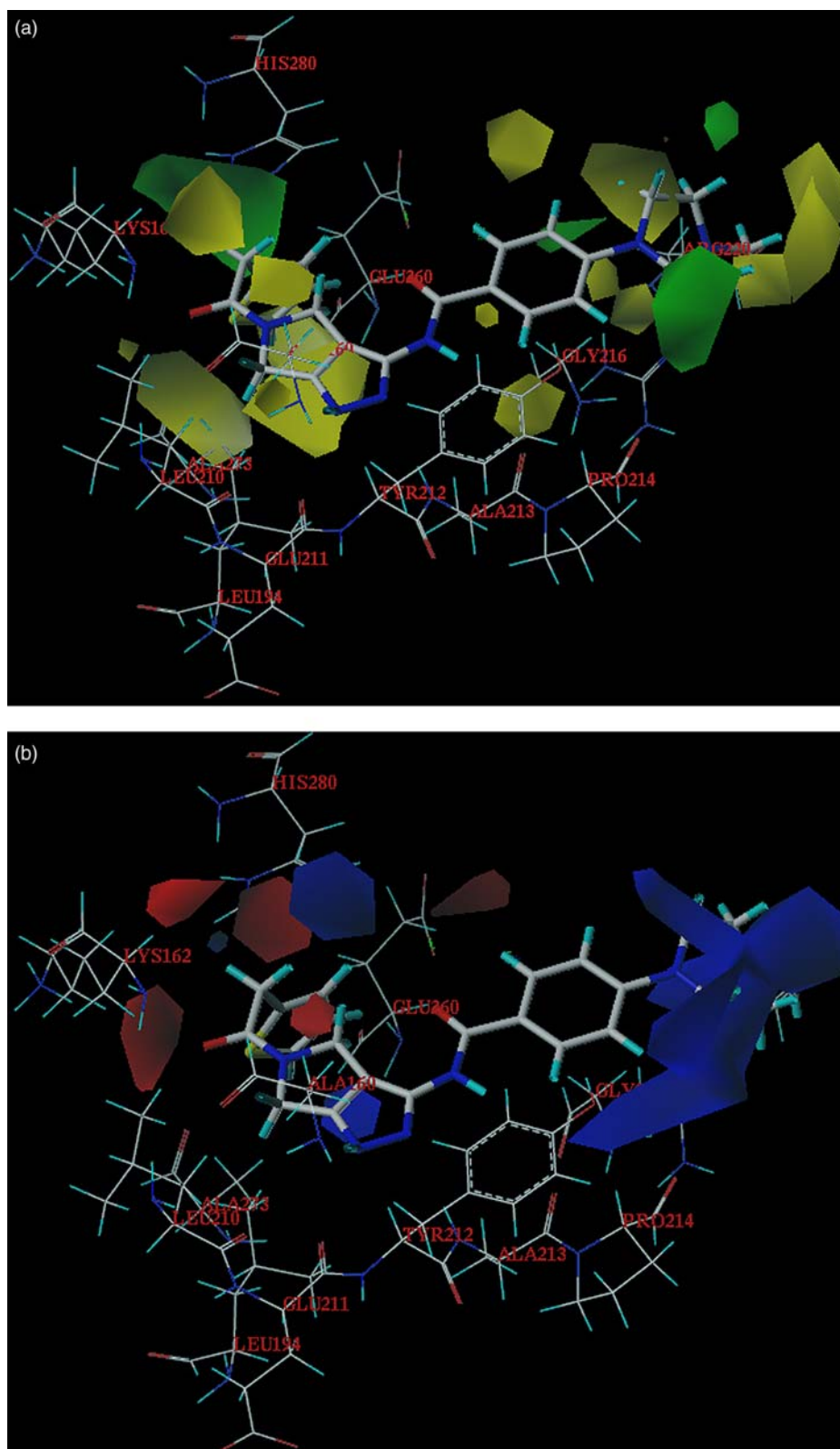
Figure 5.   CoMFA contour maps in combination with inhibitor 5. (a) The steric field distribution. (b) The electrostatic field distribution. Sterically favoured areas are in green; sterically unfavoured areas are in yellow. Positive potential-favoured areas are in blue; positive potential-unfavoured areas are in red (colour online).
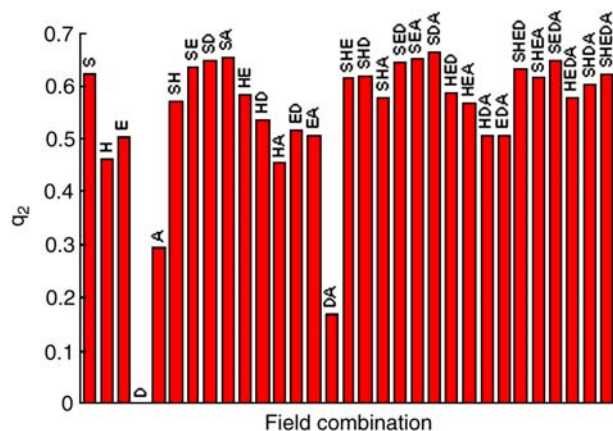
Figure 6.   Results of cross-validated q^2 for different field combinations.

of which is located near His260, indicating that the suitable chirality of R2 substitute makes an obvious influence on the inhibitory activity of Aurora-A. For example, the *S* configuration at the R2 substitution in compound **13** with the *R* configuration gives compound **12** increased inhibitory potency from 354 to 13 nM. In addition, there is a big yellow region near the pyrrole ring of compound **5**. The presence of the yellow region emphasises that a bigger group is unfavourable at this position. The binding mode of compound **5** to the Aurora-A kinase (as shown in Figure 2) indicates that the pyrrolo[3,4-c]pyrazole ring is the main H-bond site between inhibitors and the protein, a bigger group may interfere with the formation of H-bond interactions. The electric contour map (Figure 5(b)) shows two major blue regions near the R1 and R2 substitutes, respectively, that is to say, the positive charges in these two regions are important to the ligand binding, and a charge-withdrawing group linked to this position will enhance the biological activity.

CoMSIA is performed using the following descriptor fields: steric (S), electrostatic (E), hydrophilicity (H), H-bond donor (D) and H-bond acceptor (A). PLS analyses of various CoMSIA models with different combinations

of fields were obtained. The CoMSIA results are summarised in Figure 6, which shows the distribution that resulted from all of the 31 field combinations. The CoMSIA model that includes S, D and A fields performs better than the other field combinations. A cross-validated correlation coefficient q^2 of 0.664 and a non-cross-validated correlation coefficient r^2 of 0.975 with SEE of 0.149 are obtained. These results indicate that a reliable CoMSIA model is successfully constructed. Using the steric (S), hydrogen bond donor (D) and hydrogen bond acceptor (A) field combination, the sensitivity of the CoMSIA models to different grid spacings is also investigated. The difference of the q^2 values for grid spacing from 1.0 to 3.0 is about 0.024 units (Table 5), which is similar to the result of CoMFA models. From the results of the SEE, the model of grid spacing 2.0 Å is optimum, and the following discussion concerned with CoMSIA will only refer to the model generated from 2.0 Å grid spacing.

Figure 7 shows the correlation between the actual values and the predicted values, and the calculated results are listed in Table 1. In summary, the differences between CoMFA and CoMSIA are not striking and both models demonstrated good predictive ability.

The CoMSIA steric and electrostatic field contour plots employing S, D and A fields are shown in Figures 8 and 10. These contour maps are more or less similar to the corresponding CoMFA ones, except that there are two yellow contours near the aromatic ring on the R1 substitution (Figure 8). The CoMSIA H-bond donor and acceptor contours are less than that of CoMFA electric contours, which are difficult to interpret directly in CoMFA contour. The hydrogen bond donor and acceptor contours are shown in Figure 9. In principle, it should highlight the areas near which H-bonding donor on the ligand can form H-bonds with the receptor to influence binding affinity. The presence of the cyan and magenta coloured contour near the R2 substitution indicates that the hydrogen bond donor and acceptor substituted at this position may enhance the Aurora-A inhibitory activity (Figure 9). That is to say, the R2 substitute may form

Table 5.   Influence of different grid spacings on the CoMSIA models.

	CoMSIA1	CoMSIA2	CoMSIA3	CoMSIA4	CoMSIA5
Grid spacing (Å)	1.0	1.5	2.0	2.5	3.0
q^{2a}	0.658	0.664	0.664	0.659	0.640
r^{2b}	0.962	0.965	0.975	0.966	0.949
SEE^c	0.171	0.162	0.149	0.173	0.182
F-value[d]	77.0	100.4	113.2	113.7	124.7
N^e contributions	5	4	4	4	3
Steric	0.540	0.574	0.620	0.642	0.673
H-bond donor	0.243	0.255	0.218	0.203	0.186
H-bond acceptor	0.218	0.171	0.162	0.156	0.142

[a] Cross-validated correlation coefficient. [b] Non-cross-validated correlation coefficient. [c] Standard error of estimate. [d] F-test value. [e] Optimum number of components obtained from cross-validated PLS analysis and the same used from final non-cross-validated analysis.
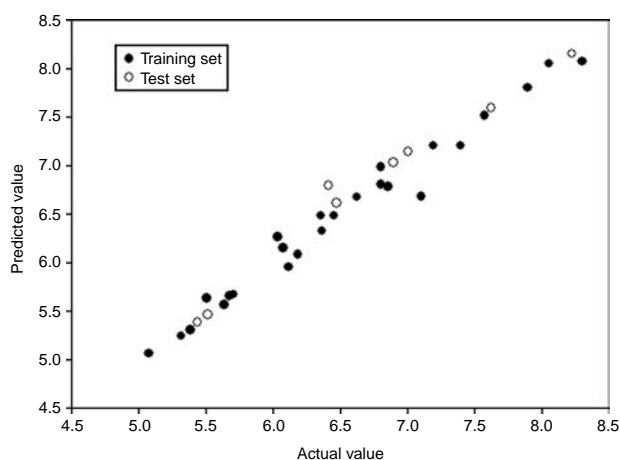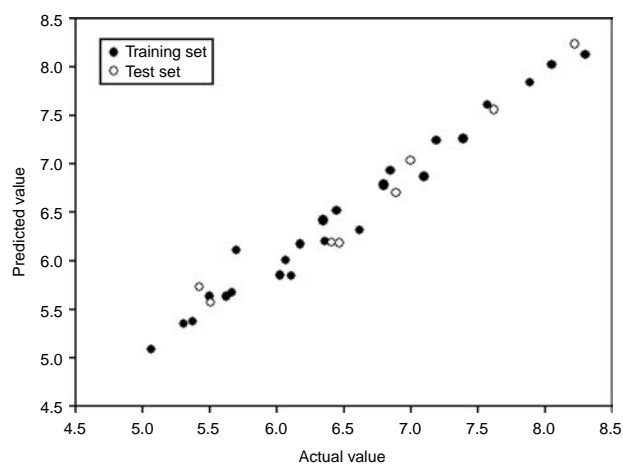
Figure 7. Correlation between predicted activities by CoMSIA models and the experimental pIC_{50} values of training and test sets, filled circles represent predictions for the training set, while open circles represent predictions for the test set.

stable H-bonds with some residues in the protein, which is consistent with the residue of Lys162 located on this region.

In order to take full advantage of the positively charged arginine residues (Arg137 and Arg220) on the epitaxial of the active pocket, and thus further enhance the interaction between the ligand and receptor, we have

the original methyl piperazine group replaced by the following three groups (N1, N2 and N3, Figure 10). Through the MD simulation and molecular mechanical (MM)/Poisson–Boltzmann surface area (PBSA) calculation, the induced fit binding modes and the corresponding binding affinity of compound **1**, N1, N2 and N3 with the receptor protein were obtained (Figure 11(a)–(d)).

For compound **1**, without obvious electrostatic interactions between piperazine and arginine residues, as the result, its binding energy is the smallest (-19.5 kcal/mol). For N1, there is an obvious strong electrostatic interaction between the carboxyl anion groups on the meta position with the positively charged Arg137 residues, which undergo a significant conformational change. As for the Arg220, without the similar electrostatic interaction with the ligand, deflection from the initial position was observed through the MD simulation. As far as N2 was concerted, the carboxyl anion group of naphthalene can interact with the two arginine residues very effectively, so the binding affinity can be stronger than N1. With regard to N3, in addition to the similar electrostatic interactions between the carboxyl anion on the indole group and the two arginine residues, hydrogen bond between the NH group on indole and the histidine residues in the active pockets would also occur. The binding energy can be far greater than the above ligands (-42.6 kcal/mol). Table 6 summarises the results of our
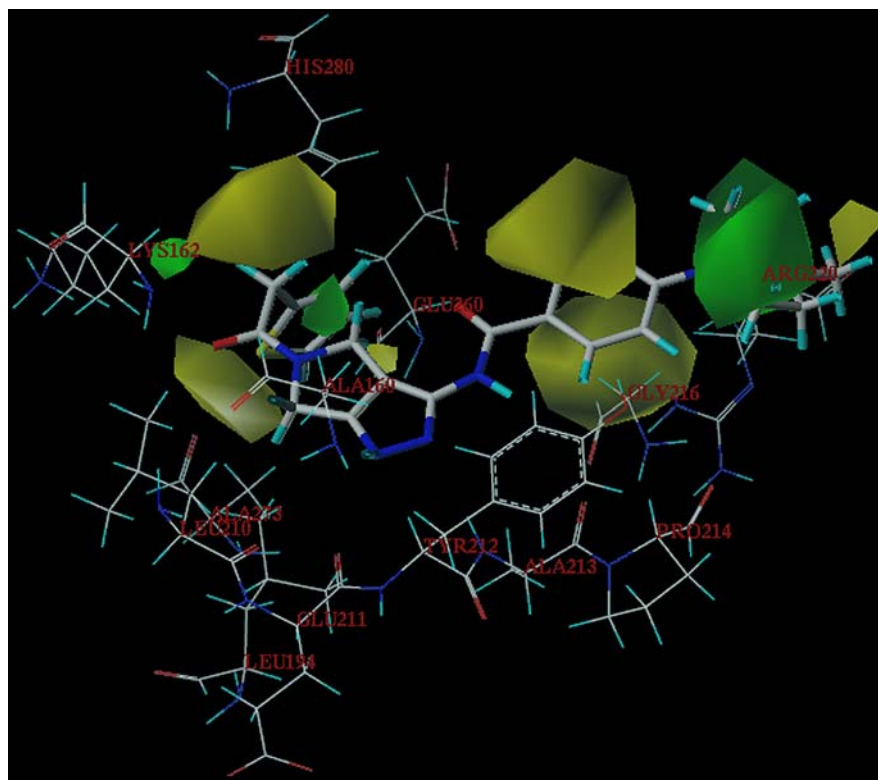

Figure 8. Steric maps from the CoMSIA model. Compound **5** is shown inside the field. Sterically favoured areas are represented by green polyhedra. Sterically disfavoured areas are represented by yellow polyhedra.
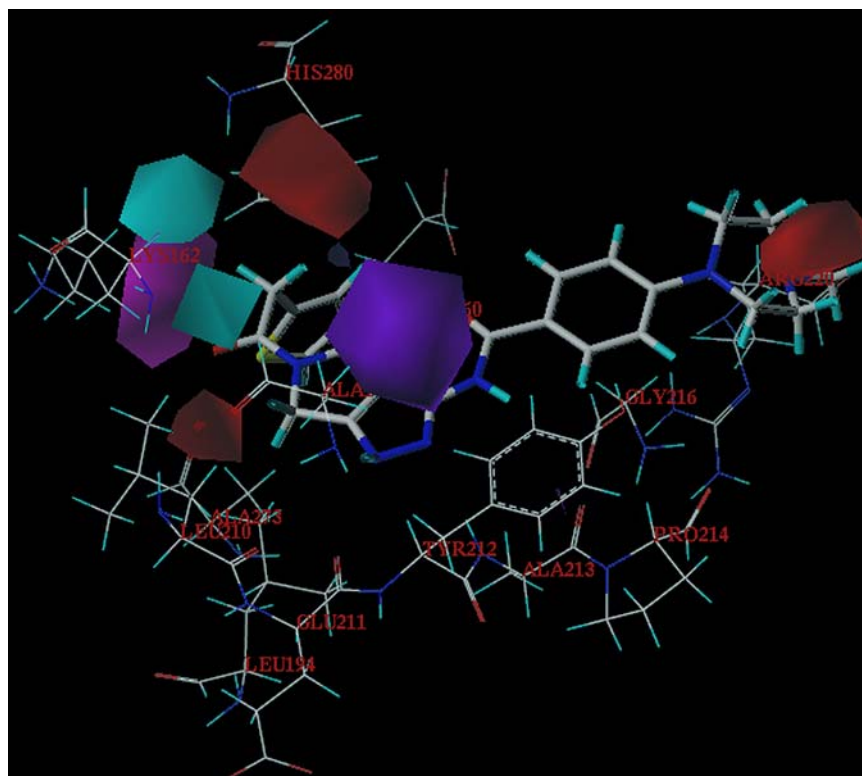
Figure 9.   Hydrogen bond donor and acceptor fields map of CoMSIA model. Compound **5** is shown inside the field. Magenta regions indicate areas where hydrogen bond acceptor groups increase activity, and red regions indicate areas where hydrogen bond acceptor groups decrease activity. Cyan regions indicate areas where hydrogen bond donor groups increase activity.
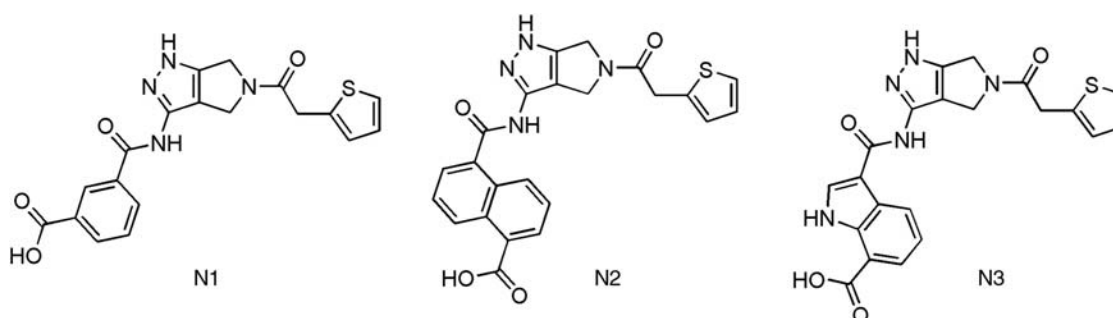
Figure 10.   Chemical structure of N1, N2 and N3.

MM/PBSA calculation, including the energy terms given by the MM/PBSA method for compound **1**, N1, N2 and N3. The low binding free energy in compound N1, N2 and N3 confirms that they are the favourable binding modes.

4.   Conclusions

In summary, using the alignment scheme generated from the docking study, highly predictive CoMFA and CoMSIA models have been developed based on Aurora-A inhibitors. The satisfactory models of CoMFA and CoMSIA are obtained with LOO cross-validated q^2 values of 0.667 and 0.664, respectively, and the non-cross-validated PLS analyses with the optimum components of 3 and 4 having conventional r^2 values of 0.966 and 0.975, respectively. These models match well with the binding conformation of the active site of Aurora-A and both models show similar predictive capabilities. According to this study, we have established the reliable 3D-QSAR models for Aurora-A inhibitors and the built model has

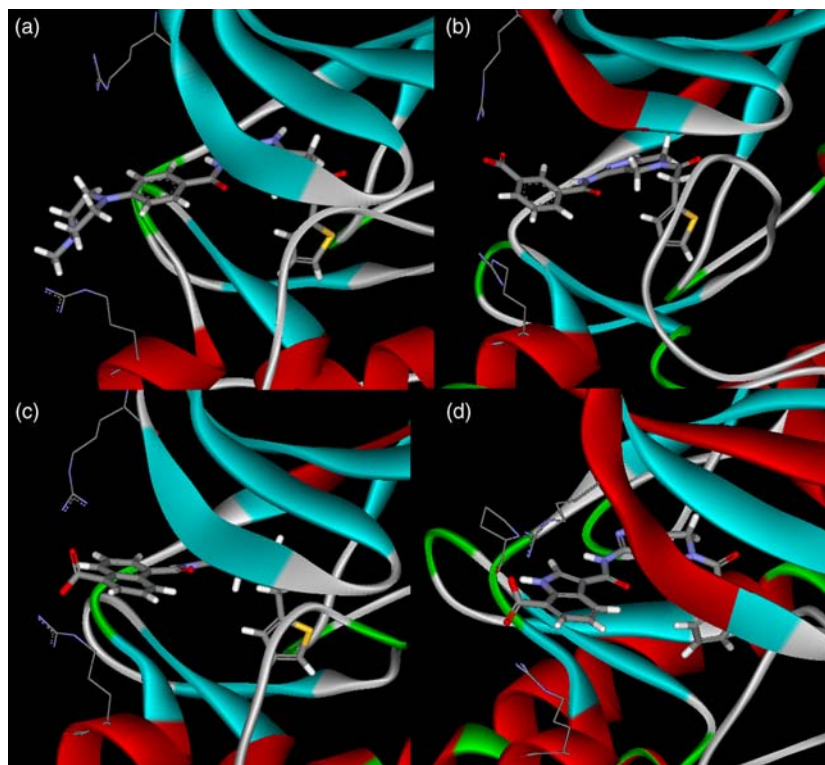Figure 11. Binding mode of compound **1** (a) N1, (b) N2, (c) N3 and (d) inside the active site of Aurora-A kinase (2J4Z.pdb, inactive form).

Table 6. Binding free energy computed by MM/PBSA methods (unit: kcal/mol).

	Compound **1**	N1	N2	N3
ΔE_{vdW}[a]	-53.8 ± 2.7	-47.5 ± 2.4	-57.2 ± 2.7	-56.2 ± 2.8
ΔE_{ele}[b]	-46.6 ± 5.0	-235.2 ± 7.5	-224.9 ± 8.1	-237.5 ± 9.5
ΔE_{PB}[c]	64.1 ± 4.4	235.3 ± 5.2	233.9 ± 5.4	232.5 ± 5.5
ΔE_{SA}[d]	-6.1 ± 0.1	-5.6 ± 0.1	-5.9 ± 0.1	-5.9 ± 0.1
$-T\Delta S$[e]	-23.0 ± 3.7	-21.1 ± 2.1	-21.7 ± 2.4	-24.5 ± 1.8
$\Delta\Delta G$	-19.5	-31.9	-32.6	-42.6

[a] ΔE_{vdW} denotes the sum of MM energies of the molecules from internal and van der Waals energies. [b] ΔE_{ele} denotes the electrostatic energies of the molecules. [c] ΔE_{PB} is the polar solvation energy in continuum solvent, computed using a finite-difference Poisson–Boltzmann (PB) model. [d] ΔE_{SA} is the non-polar solvation energy obtained from the solvent-accessible surface area (SA). [e] $-T\Delta S$ is the solute entropy.

been used for the design and development of novel inhibitors with improved activities. Three new compounds show very low binding free energy against the Aurora-A kinase, and have been shifted for further experimental assay studies.

References

[1] R. Crane, B. Gadea, L. Littlepage, H. Wu, and J.V. Ruderman, *Aurora A, meiosis and mitosis*, Biol. Cell 96 (2004), pp. 215–229.

[2] T.C. Tseng, S.H. Chen, Y.P. Hsu, and T.K. Tang, *Protein kinase profile of sperm and eggs: Cloning and characterization of two novel testis-specific protein kinases (AIE1, AIE2) related to yeast and fly chromosome segregation regulators*, DNA Cell Biol. 17 (1998), pp. 823–833.

[3] M. Kimura, S. Kotani, T. Hattori, N. Sumi, T. Yoshioka, K. Todokoro, and Y. Okano, *Cell cycle-dependent expression and spindle pole localization of a novel human protein kinase, Aik, related to Aurora of* Drosophila *and yeast Ipl1*, J. Biol. Chem. 272 (1997), pp. 13766–13771.

[4] N. Keen and S. Taylor, *Aurora-kinase inhibitors as anticancer agents*, Nat. Rev. Cancer 4 (2004), pp. 927–936.

[5] C. Ditchfield, V.L. Johnson, A. Tighe, R. Ellston, C. Haworth, T. Johnson, A. Mortlock, N. Keen, and S. Taylor, *Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores*, J. Cell Biol. 161 (2003), pp. 267–280.

[6] A. Mortlock, N.J. Keen, F.H. Jung, N.M. Heron, K.M. Foote, R. Wilkinson, and S. Green, *Progress in the development of selective inhibitors of Aurora kinases*, Curr. Topp. Med. Chem. 5 (2005), pp. 199–213.

[7] E.A. Harrington, D. Bebbington, J. Moore, R.K. Rasmussen, A.O. Ajose-Adeogun, T. Nakayama, J.A. Graham, C. Demur, T. Hercend, A. Diu-Hercend, M. Su, J.M.C. Golec, and K.M. Miller, *VX-680, a potent and selective small-molecule inhibitor of the*

Aurora kinases, suppresses tumor growth in vivo, Nat. Med. 10 (2004), pp. 262–267.

[8] A. Gontarewicz, S. Balabanov, G. Keller, R. Colombo, A. Graziano, E. Pesenti, D. Benten, C. Bokemeyer, W. Fiedler, J. Moll, and T.H. Brummendorf, *Simultaneous targeting of Aurora kinases and Bcr-Abl kinase by the small molecule inhibitor PHA-739358 is effective against imatinib-resistant BCR-ABL mutations including T315I*, Blood 111 (2008), pp. 4355–4364.

[9] X.Q. Deng, H.Y. Wang, Y.L. Zhao, M.L. Xiang, P.D. Jiang, Z.X. Cao, Y.Z. Zheng, S.D. Luo, L.T. Yu, Y.Q. Wei, and S.Y. Yang, *Pharmacophore modelling and virtual screening for identification of new Aurora-A kinase inhibitors*, Chem. Biol. Drug Des. 71 (2008), pp. 533–539.

[10] H.Y. Wang, L.L. Li, Z.X. Cao, S.D. Luo, Y.Q. Wei, and S.Y. Yang, *A specific pharmacophore model of Aurora B kinase inhibitors and virtual screening studies based on it*, Chem. Biol. Drug Des. 73 (2009), pp. 115–126.

[11] J. Zou, H.Z. Xie, S.Y. Yang, J.J. Chen, J.X. Ren, and Y.Q. Wei, *Towards more accurate pharmacophore modeling: Multicomplex-based comprehensive pharmacophore map and most-frequent-feature pharmacophore model of CDK2*, J. Mol. Graph. Model. 27 (2008), pp. 430–438.

[12] T.T. Talele and M.L. McLaughlin, *Molecular docking/dynamics studies of Aurora A kinase inhibitors*, J. Mol. Graph. Model. 26 (2008), pp. 1213–1222.

[13] R.D. Cramer, III, D.E. Patterson, and J.D. Bunce, *Comparative molecular-field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins*, J. Am. Chem. Soc. 110 (1988), pp. 5959–5967.

[14] G. Klebe, U. Abraham, and T. Mietzner, *Molecular similarity indexes in a comparative-analysis (COMSIA) of drug molecules to correlate and predict their biological-activity*, J. Med. Chem. 37 (1994), pp. 4130–4146.

[15] D. Fancelli, D. Berta, S. Bindi, A. Cameron, P. Cappella, P. Carpinelli, C. Catana, B. Forte, P. Giordano, M.L. Giorgini, S. Mantegani, A. Marsiglio, M. Meroni, J. Moll, V. Pittala, F. Roletto, D. Severino, C. Soncini, P. Storici, R. Tonani, M. Varasi, A. Vulpetti, and P. Vianello, *Potent and selective Aurora inhibitors identified by the expansion of a novel scaffold for protein kinase inhibition*, J. Med. Chem. 48 (2005), pp. 3080–3084.

[16] P. Pevarello, D. Fancelli, A. Vulpetti, R. Amici, M. Villa, V. Pittala, P. Vianello, A. Cameron, M. Ciomei, C. Mercurio, J.R. Bischoff, F. Roletto, M. Varasi, and M.G. Brasca, *3-Amino-1,4,5,6-tetrahydropyrrolo[3,4-c]pyrazoles: A new class of CDK2 inhibitors*, Bioorg. Med. Chem. Lett. 16 (2006), pp. 1074–1090.

[17] D. Fancelli, J. Moll, M. Varasi, R. Bravo, R. Artico, D. Berta, S. Bindi, A. Cameron, I. Candiani, P. Cappella, P. Carpinelli, W. Croci, B. Forte, M.L. Giorgini, J. Klapwijk, A. Marsiglio, E. Pesenti, M. Rocchetti, F. Roletto, D. Severino, C. Soncini, P. Storici, R. Tonani, P. Zugnoni, and P. Vianello, *1,4,5,6-Tetrahydropyrrolo[3,4-c]pyrazoles: Identification of a potent Aurora kinase inhibitor with a favorable antitumor kinase inhibition profile*, J. Med. Chem. 49 (2006), pp. 7247–7251.

[18] M.G. Brasca, C. Albanese, R. Amici, D. Ballinari, L. Corti, V. Croci, D. Fancelli, F. Fiorentini, M. Nesi, P. Orsini, F. Orzi, W. Pastori, E. Perrone, E. Pesenti, P. Pevarello, F. Riccardi-Sirtori, F. Roletto, P. Roussel, M. Varasi, A. Vulpetti, and C. Mercurio, *6-Substituted pyrrolo[3,4-c]pyrazoles: An improved class of CDK2 inhibitors*, Chem. Med. Chem. 2 (2007), pp. 841–852.

[19] *Sybyl Version 6.9*, Tripos Associates, St Louis, MO, 2001.

[20] M. Clark, R.D. Cramer, and N.V. Opdenbosch, *Validation of the general-purpose Tripos 5.2 force-field*, J. Comput. Chem. 10 (1989), pp. 982–1012.

[21] T.J. Hou, Y.Y. Li, N. Liao, and X.J. Xu, *Three-dimension quantitative structure–activity relationship analysis of some cinnamamides using comparative molecular similarity indices analysis (CoMSIA)*, J. Mol. Model. 6 (2000), pp. 438–455.

[22] Y.S. Guo, J.F. Xiao, Z.R. Guo, F.M. Chu, Y.H. Cheng, and S. Wu, *Exploration of a binding mode of indole amide analogues as potent histone deacetylase inhibitors and 3D-QSAR analyses*, Bioorg. Med. Chem. 13 (2005), pp. 5424–5434.

[23] J. Singh, H. van Vlijmen, W.C. Lee, Y. Liao, K.C. Lin, H. Ateeq, J. Cuervo, C. Zimmerman, C. Hammond, M. Karpusas, R. Palmer, T. Chattopadhyay, and S.P. Adams, *3D QSAR (COMFA) of a series of potent and highly selective VLA-4 antagonists*, J. Comput. Aided Mol. Des. 16 (2002), pp. 201–211.