Quantitative Structure Activity Relationship for The Computational Prediction of Aurora-B Inhibitors

L. Jayashankar[*],B. Syama Sundar

Department of Pharmacy, Acharya Nagarjuna University, Guntur – 522510, Andrapradesh, India

Abstract:
3-Dimensional Quantitative Structure-Activity Relationship (3D-QSAR) study was performed to explore the binding mechanism of some Quinazoline derivatives to Aurora B Kinase. Molecular Field Analysis (MFA) and Receptor surface Analysis (RSA) methods have been carried out to derive best QSAR models. Model developed by MFA and RSA methods has an R^2 (conventional) value of 0.954 and 0.949 respectively. The predictive R2 obtained were 0.9077 and 0.908 for MFA and RSA respectively .These results are suggestive of a statically robust and predictive model. Developed 3D-QSAR models provided crucial information about the field descriptors that could be used for the design of potential inhibitors of Aurora Kinase B.

Keywords: QSAR, Aurora Kinase B, MFA, RSA, Quinazolines

Introduction:

The Aurora family of serine/threonine kinases is essential for mitotic regression. Aurora-A has a crucial role in mitotic spindle formation and centrosome maturation, ensuring faithful segregation of chromosomes into daughter cells. In mammalian cells, abrogation of Aurora kinase-A activity disrupts cell cycle progression. Microinjection of antibodies to Aurora-A or depletion of Aurora-A by RNA interference delays mitotic entry. Very little is known, however, about the ultimate fate of the arrested cells. Aurora-B is a 'chromosomal passenger' protein that is essential for chromosomal congression and cytokinesis. It is associated with centromeres during prometaphase, and with the spindle midzone during anaphase and telophase. Overexpression of kinase-inactive Aurora-B disrupts kinetochore-microtubule interactions, cleavage furrow formation and cytokinesis, leading to polyploidy. This polyploid state may arrest cell-cycle progression through activation of a 'tetraploidy checkpoint'. The function of Aurora-C remains unclear. In normal tissues, the expression of this centrosome-associated kinase is predominantly restricted to germ cells. Expression and activity of the Aurora kinases are tightly regulated during the cell cycle. Activity of all three proteins peaks during the G2 and mitotic phases of the cell cycle, while expression is low or undetectable in resting cells. A variety of Aurora substrates have been identified, the most well characterized being histone H3, a protein involved in chromosome condensation and mitotic entry. Other substrates include CENP-A, myosin II regulatory light chain, protein phosphatase-1, TPX-2, INCENP21, survivin, topoisomerase II alpha, vimentin, MBD-3, MgcRacGAP, desmin, Ajuba8, XIEg5 (in Xenopus), Ndc10p (in budding yeast) and D-TACC (in Drosophila). These proteins all have a role in cell division. Since its discovery in 1997, the mammalian Aurora kinase family has been closely linked to tumorigenesis. Overexpression of Aurora-A transforms mammalian fibroblasts and gives rise to aneuploid cells containing multiple centrosomes and multipolar spindles. The resulting genetic instability is likely to contribute to tumorigenesis. Indeed, amplification of the AURKA locus correlates with chromosomal instability in mammary and gastric tumors. The Aurora kinases are overexpressed in a wide range of human tumors. Elevated expression of Aurora-A has been detected in over 50% of colorectal, ovarian and gastric tumors, and in 94% of invasive duct adenocarcinomas of the breast. In addition, amplification of the AURKA locus (20q13) correlates with poor prognosis for patients with node-negative breast cancer. Aurora-B is highly expressed in multiple human tumor cell lines, and its

levels increase as a function of Duke stage in primary colorectal cancers. Aurora-C, which is normally only found in germ cells, is also overexpressed in a high percentage of primary colorectal cancers and in a variety of tumor cell lines. The advent of targeted therapies for specific cancer phenotypes, such as Gleevec (imatinib) for chronic myelogenous leukemia, has resulted in a surge of optimism across the field of oncology. Gleevec is a small-molecule kinase inhibitor that targets BCR-ABL, c-Kit and platelet-derived growth factor receptor kinases. The clinical success of Gleevec has increased confidence that small-molecule inhibitors of specific kinases may prove to be highly effective anticancer agents. [1-8]

Experimental work:

All molecular modeling calculations and visualizations outlined herein were performed on Silicon Graphics octane workstation running on IRIX 6.5 operating system. The following software packages were used in the present study. Cerius2 Version 4.10

Biological Data and Molecular Structure Generation

The activity data and two-dimensional structures for quinazoline analogs were taken from the literature reported by Carvajal *et al*. Inhibitory constant values (IC50) reported for the compounds were converted to their corresponding pIC50 values, using a simple transformation (-log IC50) where pIC50 represents the value in nanomolar (nM) concentration. All the molecules were initially modeled using 3D Sketcher module of Cerius^2 software. Partial atomic charges were assigned using the Gasteiger method. Initial geometries of the molecules were minimized using the smart minimizer and further geometric optimizations were performed in MOPAC using AM1 method. The dataset compounds were divided into two sets, namely training

set of 27 molecules and test set consisting of 5 molecules. [9-12]

Alignment of 3D QSAR

Alignment was performed using the align module of Cerius^2. Core Substructure Search (CSS) alignment was carried out keeping the align strategy as Consensus.

3D QSAR Studies

Three dimensional quantitative structure activity relationship (3D-QSAR) models were developed using Molecular Field Analysis (MFA) and Receptor Surface Analysis (RSA) methods implemented in Cerius2.

Molecular Field Analysis

Molecular field values were generated on a rectangular grid for all the aligned molecules using CH3 (steric) and H+ (electrostatic) probes. Only 10% from the total variables, with the highest variance were considered as independent variables(Y). The biological activities of all the quinazoline molecules in the training set were used as dependent variables (Table 1). Genetic function algorithm (GFA) combined with partial least square (PLS) approach was used for variable selection and fitting. MFA study was carried out using G/PLS method consisting of 5,000 crossover generations on a population of 100 parent equations. The equation length was set to 10 terms including a constant.

Receptor Surface Analysis

The RSA was used to construct a hypothetical model of the receptor site that embodies essential information about the receptor in terms of hydrophobicity, charge, electrostatics (ELE) potential. The receptor surface was generated, using *van der Waals* field function, with weights proportional to the biological activity. RSA analysis was carried out using G/PLS method consisting of 5,000 crossover generations on a population of 100 parent equations. The equation length was set to 10 terms including a constant.

Table 1: Aurora B inhibitors with observed and calculated biological activities possessing the following scaffold.

Scaffold
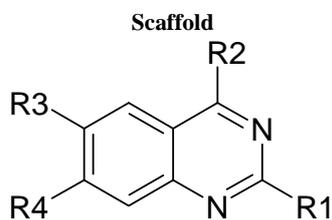
Compds	R1	R2	R3	R4	Activity pIC50	Predicted	
						MFA	RSA
1	H		OCH3		4.44	4.481	4.49
2	H		OCH3		3	3.022	3.10
3	H		OCH3		3.3	3.436	3.38
4	H		OCH3		3.69	3.935	4.02
5	H		OCH3		5.7	5.818	5.62
6	H		OCH3		5.45	4.774	4.64
7	H		OCH3		3.9	3.257	3.26
8	H		OCH3		3	3.363	3.38
9	H		OCH3		3	3.695	3.68
10	-S-Phe		H	H	5	5.583	5.64
11	Ph-O-Me		H	H	5	5.261	5.38
12	H		OCH3	OCH3	5	4.442	4.46
13	H		OCH3	OCH3	5.49	5.508	5.60
14	H		OCH3	OCH3	5.08	4.86	4.76
15	H		OCH3	OCH3	6	5.771	5.82
16	H		OCH3	OCH3	4.39	4.78	4.86
17	H		OCH3	OCH3	5.64	5.174	5.24
18	H		H		3	3.603	3.62
19	H		OCH3		4.98	5.036	5.04
20	H		OCH3		5.72	4.585	4.62
21	H		OCH3		4.68	4.641	4.66
22	H		OCH3		4.75	5.103	5.08
23	H		OCH3		6	5.920	5.94

24	H		OCH3		5.069	5.085	5.06
25	H		OCH3		6	6.361	6.40
26	H		OCH3		6	5.580	5.60
27	H		OCH3		4.6	4.793	4.78
28	H		OCH3		3	2.754	2.76
29	H		OCH3		3.8	4.039	4.04
30	H		OCH3		4.69	5.007	5.02
31	H		OCH3		5.12	4.745	4.76
32	H		OCH3		4.7	4.819	4.82

receptor surface which represents the virtual active site
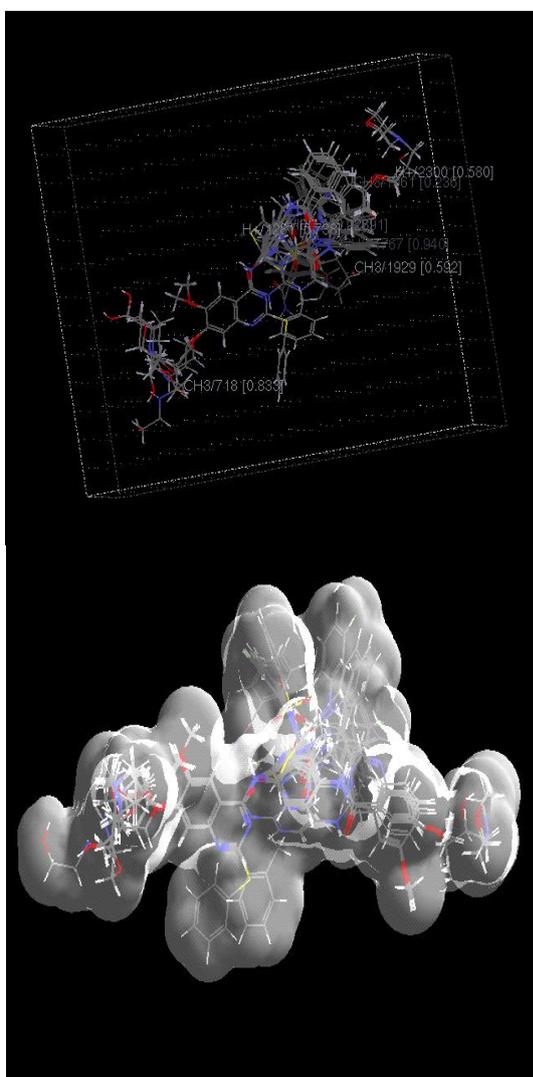
The Stereo view of rectangular molecular field surrounding aligned molecules and
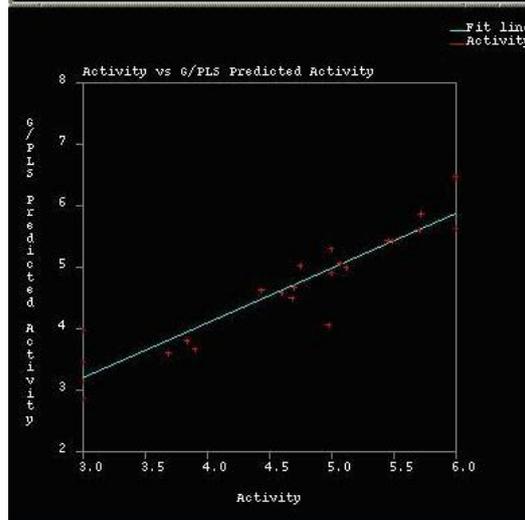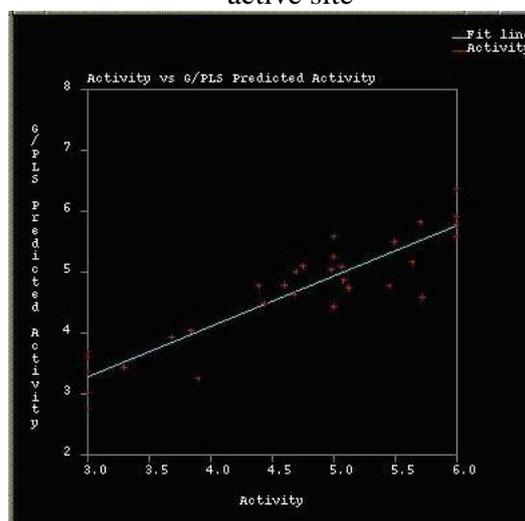
Scatter-plots of actual verses predicted

275

activity for both training and test sets using MFA and RSA

Results and discussion:

Molecular field analysis equation

Activity = 1.50437 -0.014886 *"H+/425" +0.013275 *"H+/236" -0.01349 * "H+/208" -0.019607 * "CH3/352" +0.045393 *"H+/352" (1)

Receptor surface analysis equation

Activity = -0.711179 -105.584 *"ELE/2441" -165.178 *"ELE/1071" -11.2535 *"ELE/2691" -125.784 *"ELE/91" + 00.923963 * "ELE/591" (2)

The equation obtained for MFA and RSA are given in equations 1 and 2 respectively. (CH_3) and electrostatic (H_+) descriptors specify the regions where variations in the structural features (steric or electrostatic) of different compounds in the training set, lead to increased or decreased activities. The number accompanying descriptors represents its position in the 3D dimensional MFA grid. The RSA generated model reveals the importance of the ELE as a descriptor.

Statistical Validation:

Multiple QSAR equations were generated and one with good statistical significance was chosen. The models were assed for the following parameters namely R^2, R^2_{cv}, F values and PRESS value (Table 2).

Table 2: Statistical Parameters for MFA and RSA 2.1 Chemical Data

Statistical Parameters	MFA	RSA
R^2	0.954	0.949
Nobs	26	26
R^2 Pred	0.9077	0.908
LES	0.154	0.152
R	0.9	0.88
XVR^2	0.886	0.852
BSR^2	0.952	0.944

The predictability ability of the model was established using an external test set consisting of 5 compounds, not considered during the model generation process.

The predictive power of the model was calculated using the formula

$$R^2_{Pred} = \frac{SD- PRESS}{SD} \text{------------(1)}$$

Where SD is the sum of squared deviations between the biological activities of each molecule in the test and the mean activity of the training set molecules and PRESS is the sum of squared deviations between the predicted and the actual activities of molecules in the test set.

Conclusion:

3D-QSAR studies were carried out to explore the binding mechanism of Quinazoline derivative to Aurora B kinase. This study shows how chemical features for a set of compounds along with their activities ranging over several orders of magnitudes can be used to generate QSAR equation that can successfully predict the activity. The models were not only predictive within the same series of training compounds but also diversified compounds of test set. The equation identified for the Aurora B Kinase can be used to evaluate how well the newly designed compound shows its biological activity before undertaking any further study including synthesis. This application may help in identifying or designing compounds for further biological evaluation and optimization

References:

[1] Carvajal et al., *Clin Cancer Res.* 2006, 12, 6869 - 6875.
[2] Bedrick B. Gadea, and Joan V. Ruderman.,

PNAS. 2006, 103, 4493 - 4498.

[3] Fiona Girdler, Karen E. Gascoigne, Patrick A. Eyers, Sonya Hartmuth, Claire Crafter, Kevin M. Foote, Nicholas J. Keen, and Stephen S. Taylor., *Journal of Cell Science*. 2006, 119, 3664 - 3675

[4] Barbara Vischioni, Joost J. Oudejans, Wim Vos, Jose A. Rodriguez, Giuseppe Giaccone, Vischioni et al., *Mol Cancer Ther*. 2006, 5, 2905 - 2913

[5] Bedrick B. Gadea and Joan V. Ruderman., *Molecular Biology of the Cell*. 2005, 16, 1305 – 1318.

[6] Silke Hauf, Richard W. Cole, Sabrina LaTerra, Christine Zimmer, et al., *J. Cell Biol*. 2003, 161, 281 - 294

[7] Emilie Montembault, Claude Prigent., *Drugs of the Future*. 2005, 1, 30

[8] Elizabeth. A. Harrington, David Bebbington, JeffMoore et al., *Nat Med*. 2004, 10, 262 - 267

[9] J. A. R. P. Sarma, G. Rambabu, K. Srikanth, D. Raveendra and M. Vithal., *Bioorg. Med. Chem. Lett*. 2002, 12, 2689 – 2693

[10] Jung FH, Pasquet G, Lambert-van der Brempt C et al*., J Med Chem.* 2006, 49, 955 - 70

[11] Andrew Mortlock, Nicholas J. Keen, Frederic H. Jung et al., *Curr Top Med Chem*. 2005, 5, 199 - 213

[12] Nicola M. Heron,a, Malcolm Anderson, David P. Blowers et al., *Bioorg. Med. Chem. Lett.* 2006, 16, 1320 – 1323