Mol Divers (2010) 14:27–38
DOI 10.1007/s11030-009-9137-9

FULL LENGTH PAPER

3D-QSAR and molecular docking studies of 4-anilinoquinazoline derivatives: a rational approach to anticancer drug design

Sisir Nandi · Manish C. Bagchi

Received: 2 January 2009 / Accepted: 27 February 2009 / Published online: 28 March 2009

Abstract The present article is an attempt to formulate the three-dimensional quantitative structure–activity relationship (3D-QSAR) modeling of 4-anilinoquinazoline derivatives having promising anticancer activities inhibiting epidermal growth factor (EGFR) kinase. Molecular field analysis was applied for the generation of steric and electrostatic descriptors based on aligned structures. Partial least-squares (PLS) method was applied for QSAR model development considering training and test set approaches. The PLS models showed some interesting results in terms of internal and external predictability against EGFR kinase inhibition for such type of anilinoquinazoline derivatives. Steric and electrostatic field effects are discussed in the light of contribution plot generated. Finally, molecular docking analysis was carried out to better understand of the interactions between EGFR target and inhibitors in this series. Hydrophobic and hydrogen-bond interactions lead to identification of active binding sites of EGFR protein in the docked complex.

Keywords Murine tumors · Anilinoquinazoline derivatives · Molecular field analysis · Steric and electrostatic descriptors · Partial least squares · Binding affinity

Introduction

A considerable amount of experimental studies have been carried out with 4-anilinoquinazoline derivatives which are potent and highly selective inhibitors of epidermal growth factor (EGFR) phosphorylation at the ATP binding site. These compounds cause inhibition of EGFR produced by abnormal signal transduction via hyperactivation of tyrosine protein kinases due to overexpression or mutation, thus leading to anticancer activities against human lung cancer, breast cancer, squamous head, and neck carcinomas [1]. A number of 5-substituted 4-anilinoquinazoline derivatives were synthesized by Ballard et al. [2] and these compounds were evaluated in erbB2 and EGFR kinase assays measuring inhibition of phosphorylation at the ATP binding site. Rewcastle and co-workers [3] prepared two series of 4-(phenylmethyl) amino and 4-(3-bromophenyl) amino quinazoline compounds and evaluated their inhibitory activities against EGFR tyrosine kinases that ultimately led to structure–activity relationships of these compounds. Structure–activity relationships of a series of quinazoline derivatives studied by Gibson et al. [4] identified 4-(4-iso quinolylamino) quinazoline and 4-(trans-2-phenyl cyclopropylamino) quinazoline as potent EGFR inhibitors against a tumor xenograft model (A431 vulval carcinoma in nude mice). In order to study the structure–activity relationships, Hennequin and co-workers [5] synthesized a number of 4-anilinoquinazoline compounds, and it was shown that anilinoquinazolines possessing C-6 aminomethyl side-chains act as potent and selective inhibitors of EGFR kinase. Structure–activity relationships for 4-anilinoquinazolines and modeling of the binding of these compounds to EGFR have also been studied by Denny [6]. Bridges et al. [7] synthesized numerous 4-anilinoquinazoline derivatives acting as EGFR-mediated potential tyrosine kinase inhibitors, and the anticancer activities of these compounds against human A431 carcinoma cell vesicles have been reported. However, hardly any quantitative structure–activity relationships (QSARs) based on structural parameters of the 4-anilinoquinazoline derivatives have been presented. The present article is an attempt to develop QSAR models based on

S. Nandi · M. C. Bagchi (✉)
Structural Biology and Bioinformatics Division, Indian Institute of Chemical Biology, 4, Raja S.C. Mullick Road, Jadavpur, Calcutta 700032, India
e-mail: mcbagchi@iicb.res.in

three-dimensional quantitative structure–activity relationship (3D-QSAR) methods for 4-anilinoquinazoline compounds.

For the development of 3D-QSARs, molecular field analysis (MFA) [8] has been applied to evaluate specific contributions of steric and electrostatic field effects necessary for the activity variation of 4-anilinoquinazolines. These steric and electrostatic field descriptors are useful for the better understanding of molecular modeling studies of these series of compounds in terms of ligand–receptor interactions. In the present study, an attempt has been made to formulate 3D-QSAR models using partial least-squares (PLS) [9,10] methodology. The concept of training and test sets has been introduced for the prediction of EGFR kinase inhibitory activity of structurally diverse sets of compounds. It is natural to associate molecular docking studies for clarification of the binding mode of these series of compounds. Interesting features have been obtained from the ligand–receptor docked conformation in terms of binding affinity and interaction between the ligand and active binding sites. It is expected that such 3D-QSARs and ligand–receptor-based molecular modeling studies of 4-anilinoquinazolines will provide better tools for rational design of promising EGFR inhibitors having greater therapeutic safety and efficacy [11–16].

Materials and methods

Biological activity data

A number of 4-anilinoquinazoline derivatives having anticancer activities by EGFR kinase inhibition [7] were considered in the present study. The experimental biological activities, in the form of IC_{50} (nM), were converted into pIC_{50} ($-\log IC_{50}$, μM), where IC_{50} represents the concentration of these compounds that produce 50% kinase inhibition. Our aim is to utilize these activity data for the development of a valid 3D-QSAR model based on steric and electrostatic fields that gives a deep insight into structure–property–activity correlations. Table 1 shows the structure of 64 such compounds along with their biological activity values. Out of these 64 compounds, small aliphatic substituents such as $-$OMe, $-NH_2$, $-NO_2$, and $-OC_2H_5$, are attached at 6- and 7-positions of compounds 1–57, whereas compounds 58–64 contain methoxy group at 6- and 7-positions. Moreover, $-NH_2$, $-$Me, and $-$OMe groups are attached to 2-, N^4-, and 5-positions of compounds 58–64.

From Table 1, it is observed that mostly small structural substituents are introduced at 3′-position of the aniline group, whereas 6- and 7-positions are modified by bulky aliphatic substituents. Small lipophilic electron-withdrawing substituents such as Cl, Br, and CF_3 at the 3′-position of the aniline group are favorable. Electron-donating groups such as 6,7-$(OCH_3)_2$, 6,7-$(OC_2H_5)_2$, and 6,7-$(OC_3H_8)_2$ should be substituted at the 6- and 7-positions of the 4-anilinoquinazoline ring to generate more potent compounds. Compounds bearing 6-NH_2 group or 6,7-$(NH_2)_2$ in case of 3-bromoanilino quinazolines are highly active in this series.

Geometry optimization

Three-dimensional quantitative structure–activity relationship studies of 4-anilinoquinazoline derivatives were carried out by using Molecular Design Suite software version 3.5 [17]. Three-dimensional structures of all compounds have been constructed using MDS 3.5 and their geometries were subsequently optimized to make the conformations having least potential energy. Energy minimizations were performed using Merck molecular force field (MMFF) and MMFF charge [18] followed by considering distance-dependent dielectric constant of 1.0 and convergence criterion of 0.01 kcal/mol. The total energy of a conformation can be calculated using MMFF by the relation

$$E_{\text{total}} = E_{\text{B}} + E_{\text{A}} + E_{\text{BA}} + E_{\text{OOP}} + E_{\text{T}} + E_{\text{vdw}} + E_{\text{elec}},$$

where

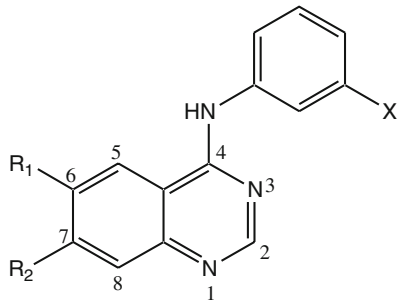
E_{B} = energy of bond stretching
E_{A} = energy of angle bending
E_{BA} = energy of bond stretching and angle bending
E_{OOP} = out-of-plane bending energy
E_{T} = torsion energy term
E_{vdw} = van der Waals energy
E_{elec} = electrostatic energy

Alignment of molecules

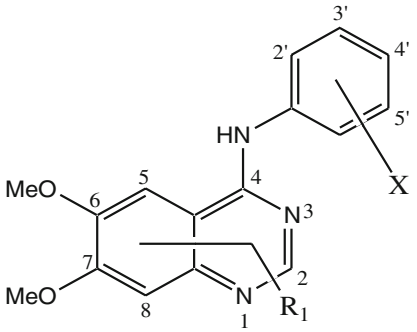
Molecular alignment is a crucial step in 3D-QSAR study to obtain meaningful results. This method is based on moving of molecules in 3D space, which is related to the conformational flexibility of molecules. The goal is to obtain optimal alignment between the molecular structures necessary for ligand–receptor interactions [19]. All molecules in the data set were aligned by template-based method where a template is built by considering common substructures in the series. The structure of 4-anilinoquinazoline template is shown in Fig. 1.

A highly bioactive energetically stable conformation in this class of compounds is chosen as a reference molecule on which other molecules in the data set are aligned, considering template as a basis for the alignment. For this purpose, the crystal conformation of erlotinib [20] was selected as a reference for generating the alignment in the present investigation. The aligned view of 4-anilinoquinazolines is presented in Fig. 2.

Table 1 Structures of 4-anilinoquinazoline derivatives with activities



1-57



58-64

Comp No.	Substituents			IC_{50} (nM)	pIC_{50} (μM)	PLP score (kcal/mol)
	R_1	R_2	X			
1	H	H	H	344	0.463	−49.300
2*	H	H	F	56	1.251	−50.858
3*	H	H	Cl	23	1.638	−49.058
4	H	H	Br	27	1.568	−50.516
5*	H	H	I	80	1.096	−53.814
6	H	H	CF_3	577	0.238	−51.265
7*	OMe	H	H	55	1.259	−54.089
8	OMe	H	Br	30	1.522	−52.137
9*	NH_2	H	H	770	0.113	−50.449
10*	NH_2	H	CF_3	574	0.241	−51.998
11*	NH_2	H	Br	0.78	3.107	−56.570
12	NO_2	H	H	5000	−0.698	−55.528
13*	NO_2	H	Br	900	0.045	−53.415
14	H	MeO	H	120	0.920	−56.860
15*	H	MeO	Br	10	2.000	−54.465
16	H	NH_2	H	100	1.000	−53.032
17*	H	NH_2	F	2	2.698	−54.298
18	H	NH_2	Cl	0.25	3.602	−53.636
19*	H	NH_2	Br	0.1	4.000	−57.083
20	H	NH_2	I	0.35	3.455	−51.938
21	H	NH_2	CF_3	3.3	2.481	−58.723
22*	H	NO_2	H	12000	−1.079	−54.651
23*	H	NO_2	F	6100	−0.785	−57.359
24*	H	NO_2	Cl	810	0.091	−55.319
25	H	NO_2	Br	1000	0.000	−53.921
26*	H	NO_2	I	540	0.267	−54.426
27	OMe	OMe	H	29	1.537	−61.904
28*	OMe	OMe	F	3.8	2.420	−57.418
29	OMe	OMe	Cl	0.31	3.508	−59.088
30	OMe	OMe	Br	0.025	4.602	−65.423
31*	OMe	OMe	I	0.89	3.050	−60.996
32	OMe	OMe	CF_3	0.24	3.619	−60.691
33*	NHMe	H	Br	4	2.397	−56.509
34	NMe_2	H	Br	84	1.075	−53.287
35	$NHCO_2Me$	H	Br	12	1.920	−63.560

Table 1 continued

Comp No.	Substituents			IC_{50} (nM)	pIC_{50} (μM)	PLP score (kcal/mol)
	R_1	R_2	X			
36*	H	OH	Br	4.7	2.327	−58.106
37	H	NHAc	Br	40	1.397	−62.214
38*	H	NHMe	Br	7	2.154	−54.697
39	H	NHEt	Br	12	1.920	−56.608
40	H	NMe_2	Br	11	1.958	−52.324
41	NH_2	NH_2	Br	0.12	3.920	−55.921
42	NH_2	NHMe	Br	0.69	3.161	−48.377
43*	NH_2	NMe_2	Br	159	0.798	−58.141
44	NH_2	OMe	Br	3.8	2.420	−57.088
45	NH_2	Cl	Br	6.5	2.187	−53.772
46	NO_2	NH_2	Br	53	1.275	−64.780
47	NO_2	NHMe	Br	68	1.167	−55.634
48	NO_2	NMe_2	Br	2000	−0.301	−54.627
49	NO_2	NHAc	Br	28	1.552	−63.707
50	NO_2	OMe	Br	15	1.823	−58.164
51	NO_2	Cl	Br	25	1.602	−49.889
52	OCH_2O		Br	15	1.823	−60.877
53	OH	OH	Br	0.17	3.769	−55.698
54	OEt	OEt	Br	0.006	5.221	−70.960
55	OPr	OPr	Br	0.17	3.769	−67.286
56	OBu	OBu	Br	105	0.978	−70.503
57	5,6-diOMe		Br	1367	−0.135	−51.838
58	2-NH_2		3′-Br	463	0.360	−52.243
59	N^4-Me		3′-Br	152	0.818	−56.877
60	5-OMe		3′-Br	0.67	3.173	−62.789
61	H		2′-Br	128	0.892	−57.918
62	H		4′-Br	0.96	3.017	−55.623
63	H		3′,4′-diBr	0.072	4.142	−61.725
64	H		3′,5′-diBr	113	0.946	−57.199

* Compounds belonging to test set

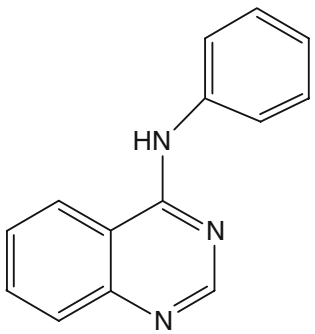


Fig. 1 4-Anilinoquinazoline (template)

Computation of steric and electrostatic fields

The aligned biologically active conformations of 4-anilinoquinazolines are used for the calculation of molecular fields. Molecular fields are the steric and electrostatic interaction energies which are used to formulate a relationship between steric and electrostatic properties together with the biological activities of compounds. Each conformation is taken in turn, and the molecular fields around it are calculated. This is done by generating three-dimensional rectangular grids around the molecule and calculating the interaction energy between the molecule and probe group placed at each grid point. Steric and electrostatic fields are computed at

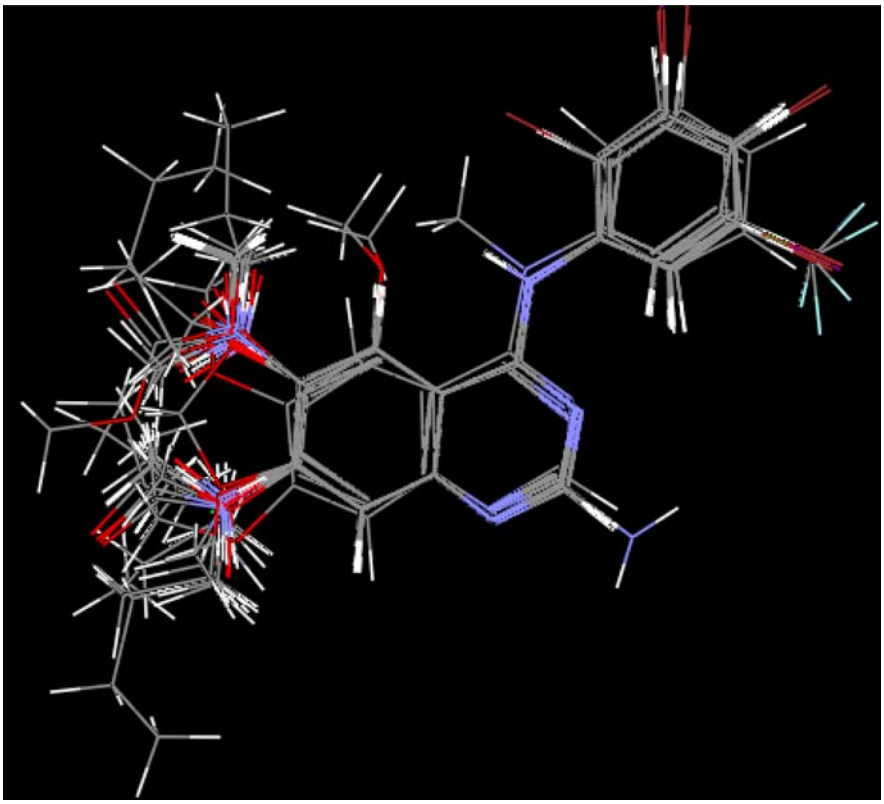


Fig. 2 3D view of aligned molecules

each grid point considering MMFF charges [18]. Methyl probe of charge +1 with 10.0 kcal/mole electrostatic and 30.0 kcal/mole steric cutoff were used for fields generation. A value of 1.0 is assigned to the distance-dependent dielectric constant. Steric and electrostatic field descriptors were calculated using Lennard–Jones and Coulomb potentials [19].

Several 3D-QSAR techniques such as comparative molecular field analysis (COMFA), comparative molecular similarity analysis (COMSIA), and *k*-nearest neighbor (*k*NN) [8,19] are being used in modern QSAR research. In the present study, molecular field analysis coupled with partial least squares (PLS) was applied to obtain a 3D-QSAR model based on steric and electrostatic descriptors. Since multicollinearity among the computed descriptors may detrimentally affect the regression analysis, PLS is frequently used as the regression method in 3D-QSAR. The calculated steric and electrostatic field descriptors were used as independent variables and pIC_{50} values were used as dependent variables in partial least-squares regression analysis [9,10] to derive the 3D-QSAR models using MDS software. The internal predictability of the models was evaluated in terms of cross-validated q^2 by the following equation:

$$q^2 = 1 - \frac{\sum_{i=1}^{N} \left(y_i - y_{\text{pred},i}\right)^2}{\sum_{i=1}^{N} \left(y_i - \hat{y}\right)^2}, \tag{1}$$

where $\sum_{i=1}^{N} (y_i - y_{\text{pred},i})^2$ is the predicted sum of squared deviations between the observed activities (y_i) and predicted activities ($y_{\text{pred},i}$) of the i-th molecule in the training set, whereas $\hat{y}$ is the mean of observed activities of all molecules in the training set. To test the utility of the model as a predictive tool, it was validated over the external test set of compounds for the prediction of activities. The external predictability of the developed model is denoted by predictive r^2 (Pred_r^2), which is given by Eq. 2.

$$\text{Pred\_}r^2 = 1 - \frac{\text{PRESS}}{\text{SSD}}, \tag{2}$$

where PRESS is the predicted sum of squared deviations between the observed and predicted activities of compounds in the test set, and SSD indicates the sum of squared deviations between the observed activities of the test molecules and the mean of observed activities of the training molecules.

Molecular docking study

Piecewise linear pairwise potential (PLP)-based molecular docking of 4-anilinoquinazoline derivatives has been performed using the docking module of Molecular Design software [21,22], which involves the use of the PLP function summed over energy interactions between all pairs of protein and ligand atoms. Molecular docking energy evaluations are usually carried out with the help of a scoring function.

There are several scoring functions such as dock score, PLP score, potential of mean force (PMF) score, steric and electrostatic score, etc. For our interest, the energy of interactions between ligand and protein was calculated in terms of PLP score, which depends upon the following different atom type parameters: hydrogen-bond donors, hydrogen-bond acceptors, both donor and acceptors, and nonpolar atoms such as carbon.

The PLP function is incorporated by the MDS software in the GRIP docking method that calculates the ligand–receptor binding affinity in terms of the PLP score. The PLP score is designed to enable flexible docking of ligands to perform a full conformational and positional search within a rigid binding site. All the optimized ligands were docked into active binding sites of EGFR target protein that can be obtained in a co-crystallized state with erlotinib (protein data bank, PDB entry 1M17) [20], which was considered as the reference to define the active binding sites in the present investigation. Water molecules and HET ATOM-like bound ligand data were removed from the PDB file of EGFR protein during docking study. A rotation angle of 30° was set so that ligand would be rotated inside the receptor cavity to generate different ligand poses inside the receptor cavity. After completion of the docking process, the minimum interaction energy between each ligand and EGFR protein for the best ligand pose inside the receptor cavity was obtained as the PLP score, which is presented in Table 1. For our interest, docking of compounds 54 (highly active), 56 (moderately active), and 25 (low activity) with minimum PLP score are discussed to explore the interaction patterns in the "Results and discussion" section.

Results and discussion

3D-QSAR modeling and its validation

In the present study, PLS coupled with stepwise variable selection method was used to develop 3D-QSAR models of 4-anilinoquinazoline derivatives based on steric and electrostatic fields. The total data set was divided into training and test sets using the sphere exclusion algorithm [23] for diversity of the sampling procedure. Compounds marked with an asterisk (*) in Table 1 were selected as test-set molecules. The quality of the model was assessed by cross-validated q^2 in the training set and external validation was performed by calculating predictive r^2 (Pred_r^2) from the test-set compounds. The 3D-QSAR model for EGFR inhibition developed in the training set may be written as

$$\begin{aligned} \mathrm{pIC_50} = {} & -22.929 + (0.072)\,\mathrm{E\_403} - (1.721)\,\mathrm{E\_1039} \\ & - (0.057)\,\mathrm{E\_389} + (3.479)\,\mathrm{S\_215} \\ & + (0.102)\mathrm{S\_414} - (3.722)\mathrm{S\_924} \\ & + (0.819)\mathrm{S\_671} \end{aligned} \tag{3}$$

Table 2 3D-QSAR-derived predicted activities of test-set compounds

Test molecule	Observed activity	Predicted activity
2	1.251	1.151
3	1.638	1.127
5	1.096	1.113
7	1.259	1.162
9	0.113	0.927
10	0.241	0.226
11	3.107	1.592
13	0.045	−0.078
15	2.000	2.566
17	2.698	1.777
19	4.000	2.762
22	−1.079	0.49
23	−0.785	0.524
24	0.091	0.501
26	0.267	0.486
28	2.420	2.943
31	3.050	2.732
33	2.397	2.496
36	2.327	0.845
38	2.154	2.653
43	0.798	1.594

$n = 43$, df = 40, $r^2 = 0.714$, $q^2 = 0.626$, F-test = 49.807, no. of optimum components = 2.

Here, n represents number of observations, df is the degrees of freedom, r is the square root of the multiple R-squared for regression, q^2 is the cross-validated r^2, and F is the F-statistic for the regression model.

E_403, E_1039, E_389, S_215, S_414, S_924, and S_671 are the steric and electrostatic field energy of interactions between probe (CH_3) and compounds at their corresponding spatial grid points of 403, 1039, 389, 215, 414, 924, and 671.

The above model is validated by predicting the biological activities of the test molecules, as indicated in Table 2.

The plot of observed versus predicted activities for the test compounds is represented in Fig. 3. From Table 2 it is evident that the predicted activities of all the compounds in the test set are in good agreement with their corresponding experimental activities and optimal fit is obtained.

The external predictability of the above 3D-QSAR model using the test set was determined by Pred_r^2, which is 0.685. So the above results indicate that 3D-QSAR model for EGFR generates 62.6% and 68.5% internal and external model prediction, respectively.

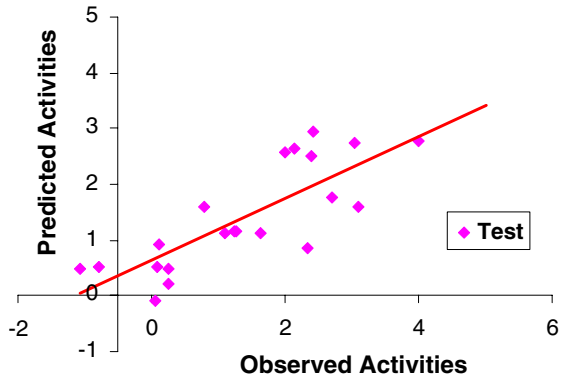


Fig. 3 Observed versus predicted activities according to the model shown in Eq. 3

Steric and electrostatic contribution plot

The plot of contributions of steric and electrostatic field interactions (Fig. 4) indicates relative regions of the local fields (steric and electrostatic) around the aligned molecules [24]. Green and blue balls represent steric and electrostatic field effects, respectively.

In the QSAR model, steric descriptors with positive coefficients represent regions of high steric tolerance; bulky substituent is favorable in this region. Steric descriptors with negative coefficients indicate regions where bulky substituent is disfavored. Electrostatic field descriptors with positive coefficients represent regions where electropositive (electron-withdrawing) groups are favorable, whereas negative coefficient indicates that electronegative (electron-rich or electron-donating) groups are favorable in this region [25].

From 3D-QSAR model Eq. 3 and Fig. 4 it is observed that electrostatic field with negative coefficient (E_1039) is far from the anilino moiety, indicating that electronegative groups are unfavorable on this site and presence of electronegative groups decrease the activity of 4-anilinoquinazoline compounds. Electrostatic descriptor with negative coefficient (E_389) around 7-position of the quinazoline ring corroborates that electronegative (electron-donating) group is preferred at 7-position of quinazolines. These results are in close agreement with the experimental observations that compounds 27, 28, 29, 30, 31, and 32 have methoxy substituent and compound 54 contains ethoxy group at 6- and 7-positions. These compounds produce greater activity due to electronegative substituents on the 6- and 7-positions of the quinazoline ring [7]. Presence of electrostatic field with positive coefficient (E_403) suggests that electropositive (electron-withdrawing) substituent may be favorable on the 5-position of template. Presence of steric descriptors with positive coefficients simultaneously at 6- and 7-positions of the quinazoline ring, such as S_414 and S_215, suggests the favorability of bulky groups in these regions for producing potent EGFR inhibitors. This is also well supported by the docking study, which confirms that the bulky substituents attached to 6- and 7-positions of the quinazoline are situated in hydrophobic pocket formed by GLY 695, LEU 694, GLY 772, and PHE 771. Bulkiness of the substituents at 6- and 7-position makes the compounds more active (compounds 28, 29, 30, 31, 32, 54, and 55) because of increased steric field effect near this region [7]. A bulky aromatic anilino substituent is essential at 4-position of the quinazoline ring for producing kinase inhibition, as indicated by the presence of steric field with positive coefficient (S_671) in

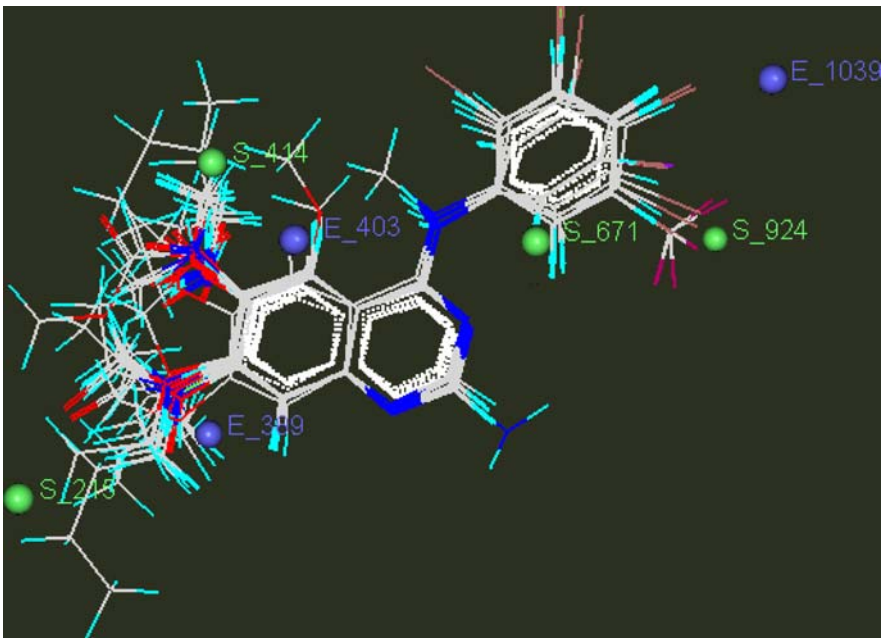


Fig. 4 Contribution plot of steric and electrostatic field interactions

this region. It is inferred from the docking results that the 4-anilino moiety is located in a deep hydrophobic pocket formed by LEU 820, VAL 702, THR 830, LYS 721, ASP 831, and ALA 719. Steric field with negative coefficient (S_924) near 3′-position of the anilino group indicates that sterically unfavorable bulky group in this position would have a detrimental effect on the EGFR kinase inhibitory activity of these quinazoline compounds. Thus, the contribution plot arising out of 3D-QSAR studies provide some useful insights for better understanding of the structural features of these compounds responsible for producing significant EGFR kinase inhibitory activity, which conforms with the docking results.

4-Anilinoquinazoline–EGFR docking

Molecular docking helps the study of ligand–receptor interaction to identify active binding sites of receptor proteins, which helps to obtain the best geometry of ligand–receptor complex so that the energy of interaction between ligand and receptor is minimum. The minimum energy of interaction is represented by different scoring functions. This utility allows screening of a set of compounds for lead optimization. The affinity of ligands to a receptor can be predicted with reasonable accuracy, yielding a relative rank ordering of the docked compounds with respect to affinity. Prediction of affinity is referred to as scoring.

The present investigation deals with docking of highly active, moderately active, and low-activity 4-anilinoquinazolines with the EGFR protein and makes a comparison between EGFR-4–anilinoquinazoline docking and crystallographic EGFR–erlotinib study. Erlotinib has been chosen for its potential drug-like activity in anilinoquinazoline congeneric series. The detailed results of interactions are given in Table 3.

Comparison of docked complexes provides an insight into the activity patterns of various 4-anilinoquinazoline compounds in terms of hydrogen-bond and hydrophobic interactions. Figures 5, 6, and 7 represent the interaction patterns of a highly active (compound 54), a moderately active (compound 56), and a low-activity (compound 25) compound, respectively, with EGFR for a clear understanding of prediction of binding sites.

Binding-site analysis and its graphical interpretation

From the above study, it is clear that the quinazoline ring is surrounded by hydrophobic residues, as indicated in Table 3. The anilino group substituted at the 4-position of the quinazoline ring is bounded by a hydrophobic pocket consisting of residues such as LEU 820, VAL 702, THR 830, LYS 721, ASP 831, and ALA 719. These observations may be compared with binding sites and interaction patterns of erlotinib in case of anilino moiety and quinazoline ring. The docked models reveal that $N-1$ of the quinazoline forms a hydrogen bond with hydrogen atom of amino backbone of MET 769. The hydrogen-bonding distances for compounds 54, 56, and 25 are 2.6 Å, 1.9 Å, and 3.0 Å, respectively. The quinazoline ring plays a crucial role for producing biological activity by interacting with MET 769 [26], an important active residue for binding affinity of the inhibitor, which correlates with the results obtained from crystallographic study of erlotinib–EGFR [20]. The minimum PLP score of −70.960 kcal/mol for compound 54 indicates high binding affinity of the ligand toward EGFR. For compound 54, the methylene carbons of ethoxy group at 6-position of the quinazoline produce strong hydrophobic interactions with GLY 695 and LEU 694, and the methylene carbons of ethoxy group at 7-position are interacting with GLY 772 and PHE 771. GLY 695, LEU 694, GLY 772, and PHE 771 are nonpolar hydrophobic residues having higher hydropathy indices [27]. The moderately active compound 56 produces good PLP score of −70.503 kcal/mol, but it is less active than compound 54 since the butoxy group at 6-position of the quinazoline is associated with hydrophobic interactions with GLY 772, CYS 773, and ASP 776, where CYS 773 and ASP 776 are hydrophilic residues having lower hydropathy indices, which may decrease the activity of the compound. The butoxy group at 7-position has hydrophobic interactions with nonpolar residues such as PHE 771, LEU 694, and GLY 772, and polar residue such as PRO 770. PRO 770 possesses hydrophilic characteristic with lower hydropathy index, causing decrease in activity of the compound. Compound 25 shows poor affinity toward EGFR, as denoted by PLP score of −53.921 and there are no hydrophobic interactions at 6- and 7-positions due to presence of $-NO_2$ as deactivating group.

Concluding remarks

The contribution plot of steric and electrostatic field interactions generated by 3D-QSAR shows that electronegative groups at aniline moiety are unfavorable. This finding is in close agreement with the structures of these compounds, where presence of electropositive groups is found in the anilino moiety. It also suggests that bulky electronegative (electron-donating) groups are favorable at 7-position of the template. This finding supports the experimental observations, where presence of bulky electronegative groups at 7-position signifies increase in activities of compounds. From the molecular docking studies, it is evident that hydrophobic groups substituted at 6- and 7-positions of the quinazoline ring possessing strong hydrophobic interactions with nonpolar active residues are likely to enhance EGFR kinase inhibition. On the other hand, presence of hydrophilic residues or polar hydrophobic residues with lower hydropathy indices in this region of interactions may decrease the activity of the 4-anilinoquinazoline compounds.

Table 3 Hydrogen-bonding and hydrophobic interactions of 4-anilinoquinazolines with EGFR kinase

Compounds		Hydrogen-bonding distance between $N-1$ of quinazoline and H atom of amino backbone of MET 769	Hydrophobic interactions (within 5 Å)			
			Anilino moiety	Quinazoline ring	Substituent attached to 6-position of quinazoline	Substituent attached to 7-position of quinazoline
Erlotinib (lead)		1.6 Å	LEU 820, VAL 702, THR 830, LYS 721, ASP 831, ALA 719, THR 766	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	GLY 695, LEU 694, VAL 702	GLY 772, PHE 771, LEU 694
Highly active	Comp 54	2.6 Å	LEU 820, VAL 702, THR 830, LYS 721, ASP 831, ALA 719	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	GLY 695, LEU 694	GLY 772, PHE 771
	Comp 30	2.7 Å	LEU 820, VAL 702, THR 830, LYS 721, ASP 831, ALA 719	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	LEU 694	GLY 772, PHE 771
Moderately active	Comp 56	1.9 Å	LEU 820, VAL 702, THR 830, LYS 721, ASP 831, ALA 719	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	GLY 772, CYS 773, ASP 776	GLY 772, PHE 771, LEU 694, PRO 770
	Comp 47	2.1 Å	LEU 820, VAL 702, THR 830, LYS 721	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	No hydrophobic interactions due to presence of $-NO_2$ as deactivating group	LEU 694, PRO 770
	Comp 43	2.8 Å	LEU 820, VAL 702, LYS 721	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	No hydrophobic interactions due to presence of $-NO_2$ as deactivating group	LEU 694, GLY 772, PRO 770
Low activity	Comp 25	3.0 Å	LEU 820, VAL 702, THR 830, LYS 721	MET 769, LEU 694, LEU 820, LEU 768, GLY 772, VAL 702, ALA 719, PRO 770	No hydrophobic interactions due to presence of $-NO_2$ as deactivating group	No hydrophobic interactions due to presence of $-NO_2$ as deactivating group

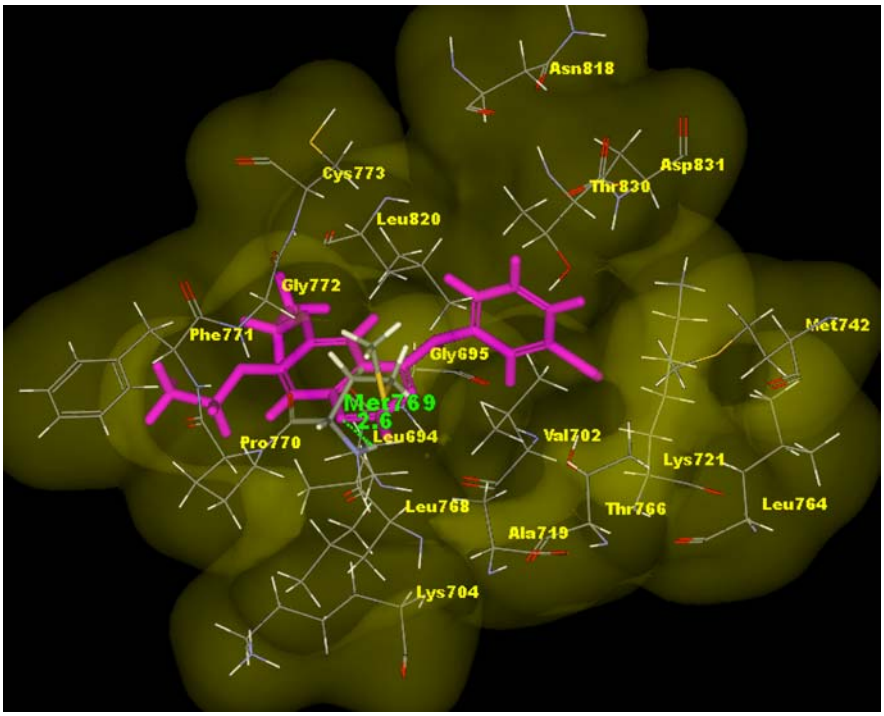


Fig. 5 Highly active ligand 54 bound with the active binding sites of EGFR, represented by molecular surface; the bound ligand is represented as stick model (pink color). The residues within 5 Å of the inhibitor are displayed. Dotted line represents H-bonding between the $N-1$ of quinazoline ring and hydrogen atom of amino backbone of MET 769 (stick model)

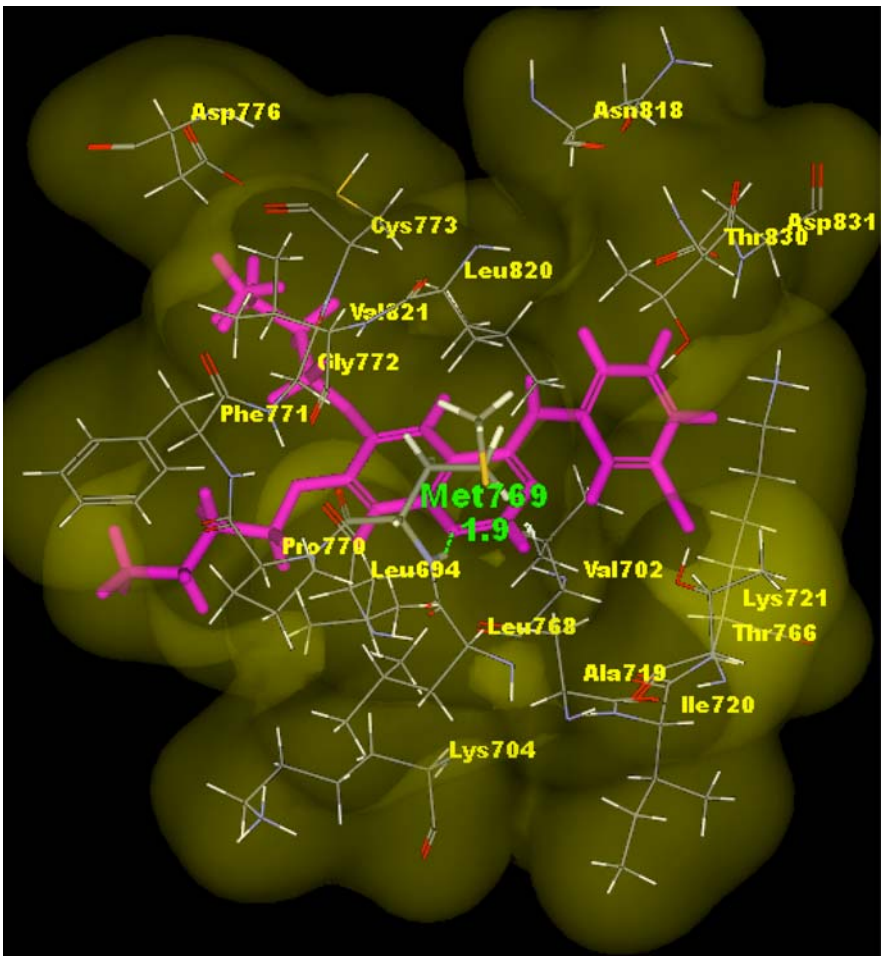


Fig. 6 Moderately active ligand 56 bound with the active binding sites of EGFR, represented by molecular surface; the bound ligand is represented as stick model (pink color). The residues within 5 Å of the inhibitor are displayed. Dotted line represents H-bonding between the $N-1$ of quinazoline ring and hydrogen atom of amino backbone of MET 769 (stick model)

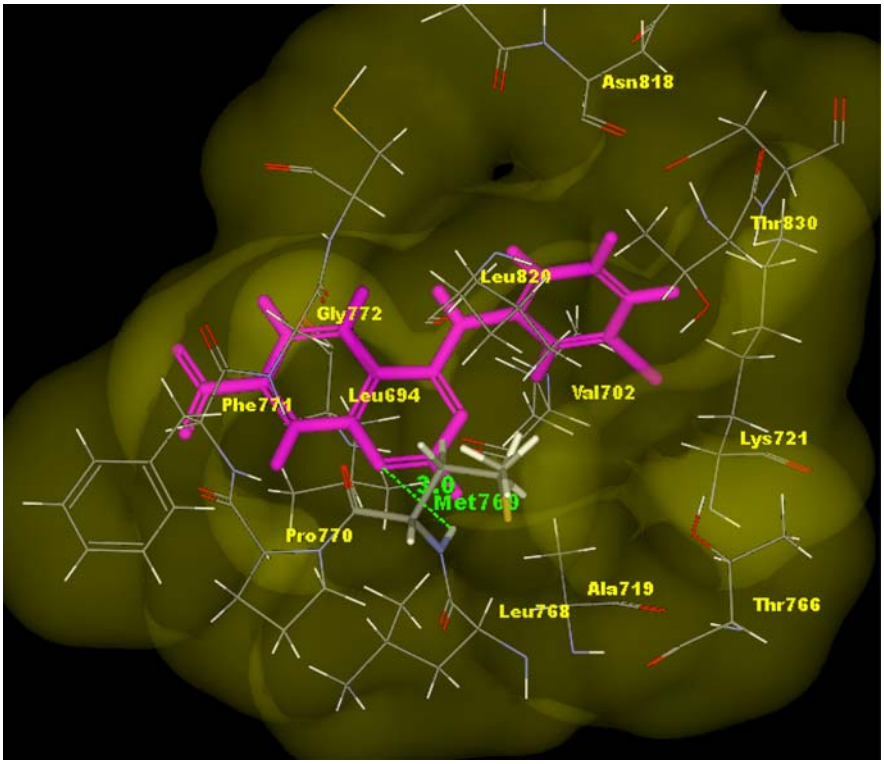


Fig. 7 Low-activity ligand 25 bound with the active binding sites of EGFR, represented by molecular surface; the bound ligand is represented as stick model (pink color). The residues within 5 Å of the inhibitor are displayed. Dotted line represents H-bonding between the $N-1$ of quinazoline ring and hydrogen atom of amino backbone of MET 769 (stick model)

Acknowledgements Sisir Nandi thanks the Council of Scientific and Industrial Research, New Delhi 110001, India for the grant of a Senior Research Fellowship to him.

References

1. Herbst RS, Fukuoka M, Baselga J (2004) Gefitinib-a novel targeted approach to treating cancer. Nat Rev Cancer 4:956–965. doi:10.1038/nrc1506
2. Ballard P, Bradbury RH, Hennequin LFA, Hickinson DM, Johnson PD, Kettle JG, Klinowska T, Morgentin R, Ogilvie DJ, Olivier A (2005) 5-Substituted 4-anilinoquinazolines as potent, selective and orally active inhibitors of erbB2 receptor tyrosine kinase. Bioorg Med Chem Lett 15:4226–4229. doi:10.1016/j.bmcl.2005.06.068
3. Rewcastle GW, Denny WA, Bridges AJ, Zhou H, Cody DR, McMichael A, Fry DW (1995) Tyrosine kinase inhibitors. 5. Synthesis and structure–activity relationships for 4-[(phenylmethyl) amino]- and 4-(phenylamino)quinazolines as potent adeonosine-5′-triphosphate binding site inhibitors of the tyrosine kinase domain of the epidermal growth factor receptor. J Med Chem 38:3482–3487. doi:10.1021/jm00018a008
4. Gibson KH, Grundy W, Godfrey AA, Woodburn JR, Ashton SE, Curry BJ, Scarlett L, Barker AJ, Brown DS (1997) Epidermal growth factor receptor tyrosine kinase: structure–activity relationships and antitumour activity of novel quinazolines. Bioorg Med Chem Lett 7:2723–2728. doi:10.1016/S0960-894X(97)10059-2
5. Hennequin LFA, Ballard P, Boyle FT, Delouvrie B, Ellston RPA, Halsall CT, Harris CS, Hudson K, Kendrew J, Pease JE, Ross HS, Smith P, Vincent JL (2006) Novel 4-anilinoquinazolines with C-6 carbon-linked side chains: synthesis and structure–activity relationship of a series of potent, orally active, EGF receptor tyrosine kinase inhibitors. Bioorg Med Chem Lett 16:2672–2676. doi:10.1016/j.bmcl.2006.02.025
6. Denny WA (2001) The 4-anilinoquinazoline class of inhibitors of the *erb*B family of receptor tyrosine kinases. Il Farmaco 56:51–56. doi:10.1016/S0014-827X(01)01026-6
7. Bridges AJ, Zhou H, Cody DR, Rewcastle GW, McMichael A, Showalter HDH, Fry DW, Kraker AJ, Denny WA (1996) Tyrosine kinase inhibitors. 8. An unusually steep structure–activity relationship for analogues of 4-(3-Bromoanilino)-6,7-dimethoxyquinazoline (PD 153035), a potent inhibitor of the epidermal growth factor receptor. J Med Chem 39:267–276. doi:10.1021/jm9503613
8. Ajmani S, Jadhav K, Kulkarni SA (2006) Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. J Chem Inf Model 46:24–31. doi:10.1021/ci0501286
9. Wold H (1975) Soft modelling by latent variables: the non-linear iterative partial least squares approach. In: Gani J (ed) Perspectives in probability and statistics, papers in honor of Bartlett MS. Academic, London
10. Hoskuldsson A (1995) A combined theory for PCA and PLS. J Chemometr 9:91–123. doi:10.1002/cem.1180090203
11. Bagchi MC, Maiti BC, Mills D, Basak SC (2004) Usefulness of graphical invariants in quantitative structure–activity correlations of tuberculostatic drugs of the isonicotinic acid hydrazide type. J Mol Model 10:102–111. doi:10.1007/s00894-003-0173-6
12. Bagchi MC, Maiti BC (2003) On application of atom pairs on drug design. J Mol Struct THEOCHEM 623:31–37. doi:10.1016/S0166-1280(02)00659-0
13. Bagchi MC, Maiti BC, Bose S (2004) QSAR of antituberculosis drugs of INH type using graphical invariants. J Mol Struct THEOCHEM 679:179–186. doi:10.1016/j.theochem.2004.04.013
14. Nandi S, Bagchi MC (2007) QSAR analysis of BABQ compounds via calculated molecular descriptors. Med Chem Res 15:393–406. doi:10.1007/s00044-006-0010-4
15. Nandi S, Vracko M, Bagchi MC (2007) Anticancer activity of selective phenolic compounds: QSAR studies using ridge regression and neural networks. Chem Biol Drug Des 70:424–436. doi:10.1111/j.1747-0285.2007.00575.x

16. Bagchi MC, Mills D, Basak SC (2007) Quantitative structure–activity relationship (QSAR) studies of quinolone antibacterials against *M. fortuitum* and *M. smegmatis* using theoretical molecular descriptors. J Mol Model 13:111–120. doi:10.1007/s00894-006-0133-z
17. Molecular Design Suite 3.5, VLife Technologies, Pune, India. www.vlifesciences.com
18. Halgren TA (1996) Merck molecular force field. III. Molecular geometries and vibrational frequencies. J Comput Chem 17:553–586. doi:10.1002/(SICI)1096-987X(199604)17:5/6<553::AID-JCC3>3.0.CO;2-T
19. Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA) 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110:5959–5967. doi:10.1021/ja00226a005
20. Stamos J, Sliwkowski MX, Eigenbrot C (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J Biol Chem 277:46265–46272. doi:10.1074/jbc.M207135200
21. Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DB, Fogel LJ, Freer ST (1995) Molecular recognition of the inhibitor AC-1343 by HIV-l protease: conformationally flexible docking by evolutionary programming. Chem Biol 2:317–324. doi:10.1016/1074-5521(95)90050-0
22. Verkhivker GM, Bouzida D, Gehlhaar DK, Rejto PA, Arthurs S, Colson AB, Freer ST, Larson V, Luty BA, Marrone T, Rose PW (2000) Deciphering common failures in molecular docking of ligand–protein complexes. J Comput Aided Mol Des 14:731–751. doi:10.1023/A:1008158231558
23. Hudson BD, Hyde RM, Rahr E, Wood J (1996) Parameter based methods for compound selection from chemical databases. Quant Struct–Act Rel 15:285–289. doi:10.1002/qsar.19960150402
24. Kubinyi H (1993) (ed) 3D QSAR in drug design. Theory, methods and applications, ESCOM. Science, Leiden, pp 486–502
25. Nandi S, Bagchi MC (2009) QSAR of aminopyrido[2,3-d]pyrimidin-7-yl derivatives: anticancer drug design by computed descriptors. J Enzyme Inhib Med Chem. doi:10.1080/14756360802519327
26. Palmer BD, Trumpp-Kallmeyer S, Fry DW, Nelson JM, Showalter HDH, Denny WA (1997) Tyrosine kinase inhibitors. 11. Soluble analogues of pyrrolo- and pyrazoloquinazolines as epidermal growth factor receptor inhibitors: synthesis, biological evaluation, and modeling of the mode of binding. J Med Chem 40:1519–1529. doi:10.1021/jm960789h
27. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132. doi:10.1016/0022-2836(82)90515-0