

3D-QSAR CoMFA analysis of C₅ substituted pyrrolotriazines as HER2 (ErbB2) inhibitors

Mahendra Awale, C. Gopi Mohan *

*Centre for Pharmacoinformatics, National Institute of Pharmaceutical Education and Research (NIPER),
Sector 67, S.A.S. Nagar, 160 062 Punjab, India*

Received 19 August 2007; received in revised form 17 October 2007; accepted 21 October 2007

Available online 4 December 2007

Abstract

Human cancers are characterized by an up-regulation of some of the RTKs (EGFR and HER2) and have been clinically validated as targets for cancer therapy. C₄ and C₅ substituted pyrrolotriazines showed dual inhibition of HER2 and EGFR protein tyrosine kinases. To explore the relationship between the structures of the aforementioned classes of molecules and their HER2 inhibition, 3D-QSAR CoMFA analysis have been performed. The developed CoMFA model showed statistically significant results with good predictive ability.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Receptor tyrosine kinase; EGFR; Human epidermal growth factor receptor; CoMFA

1. Introduction

Despite improvements in survival rates, cancer remains the second leading cause of death world wide. The epidermal growth factor receptor (EGFR, ErbB1 or HER1) and the human epidermal growth factor receptor2 (HER2, ErbB2) are members of the ErbB family of receptor tyrosine kinases (RTKs). HER2, like other EGFR members, is a transmembranous glycoprotein with intrinsic tyrosine kinase activity encoded by the HER2 protooncogene located on the long arm of chromosome 17. In normal cells, activation of this RTK family triggers a rich network of signaling pathways, which control normal cell growth, differentiation, motility, and adhesion in several cell lineages [1]. Many human cancers are characterized by an up-regulation of some of these RTKs and have been clinically validated as targets for cancer therapy. These receptors (except that of HER3) share the same molecular structure with an extracellular cysteine rich ligand binding domain, a single alpha-helix transmembrane domain, and an intracellular domain with tyrosine kinase (TK) activity at the carboxy-terminal tail [2]. The TK domains of HER2 and

HER4 showed an 80% homology with that of EGFR members. Ligand binding induces EGFR homodimerization, as well as heterodimerization with other types of HER proteins. EGFR/EGFR homodimers are unstable, whereas EGFR/HER2 heterodimers are stable, and recycle more rapidly to the cell surface [3]. The dimerization activates the RTKs, which phosphorylate special tyrosine residues on proteins. These phosphorylated tyrosine residues initiate downstream multiple signaling pathways associated with cell growth (or differentiation) mainly including the Ras/MAP kinase pathway and PKB/Akt pathway. HER2 over expression is identified on many tumor cells and the relationship between HER2 status and clinicopathological characteristics in breast cancer has been well investigated [4]. Statistically, over expression of HER2 occurs in a number of human cancers, including 25–30% of breast cancer, 5–28% of pulmonary adenocarcinoma and 17% of colorectal aden carcinoma [5]. Its up-regulated expression is also related to rapid disease progression, chemoresistance, accelerated relapse as well as poor prognosis and mortality. Consequently, they have become targets of intense drug discovery efforts to identify novel anti-cancer agents. The frequent co-expression of HER2 and EGFR in variety of tumor types and their capacity to form heterodimers with other members of the ErbB family provided a strong rationale for simultaneously targeting both of these receptors. There are currently several small molecules of dual EGFR and HER2

* Corresponding author. Tel.: +91 172 2214682x2019; fax: +91 172 2214692.

E-mail addresses: cmohan@niper.ac.in, cgopimohan@yahoo.com
(C.G. Mohan).

kinase inhibitors in clinical development which include: lapatinib (GW572016), AEE-788, 4 and BMS-599626. The latter utilizes the bicyclic pyrrolotriazine ring system as a scaffold for the construction of an ATP mimic [6]. Its lipophilic C₄ substituent provides potent and selective kinase inhibition while its C₆ solubilizing side chain imparts good potency and pharmacokinetics.

Review of literature shown *in silico* studies to understand the mode of action and the relationship between physico-chemical properties and the inhibitory activities of different kind of EGFR/HER2 inhibitors. Shi et al. performed QSAR analysis of tyrosine kinase inhibitors on 4-(3-bromoanilino)-6,7-dimethoxyquinazoline series using modified ant colony optimization and multiple linear regression method [7]. A receptor-guided alignment-based comparative 3D-QSAR studies was carried out by Kamath and Buolamwini on benzylidene malonitrile tyrphostins as EGFR and HER2 kinase inhibitors [8].

The structure–activity relationship of 5-methyl pyrrolotriazines with different C₄ substituents showed analogs with small C₄ anilines or bicyclic heterocycles and are selective EGFR inhibitors [9]. Appending a lipophilic benzyl group to the C₄ aniline increased HER2 kinase inhibition without reducing EGFR inhibition. To further explore the relationship between the structure of the a fore mentioned classes of molecules and their HER2 inhibitor activity, three dimensional quantitative structure–activity relationship (3D-QSAR) studies using comparative molecular field analysis (CoMFA) were performed.

2. Methodology

Comparative molecular field analysis (CoMFA): 3D-QSAR method CoMFA was introduced by Cramer et al. [10], in which an assumption is made that the interaction between an inhibitor and its molecular target is primarily non-covalent in nature and shape dependant. Therefore, QSAR may be derived by sampling the steric and electrostatic fields surrounding a set of ligands and correlating the difference in these fields to biological activity. CoMFA calculate steric field using Lennard–Jones potential and electrostatic field using Coulomb potential.

3. Computational details

3.1. Dataset for analysis

In vitro biological data of C₅ substituted pyrrolotriazine dual inhibitors of HER2 and EGFR protein tyrosine kinases (in human tumor xenograft models) reported by Mastalerz et al. was used [9] to construct CoMFA model and for analysis of physico-chemical features. The structure and experimental value of activity for the 32 molecules used in this study are shown in Table 1. The molecules under study were built using SYBYL7.1 [11] molecular modeling package installed on a Silicon Graphics Fuel Work station running IRIX 6.5. Since crystal structure of tyrosine kinase domain of HER2 is not

available, and we presumed that its homology model may not be good enough for docking analysis, in order to obtain accurate bioactive conformation of the various ligands under study, in the present CoMFA analysis. Therefore, bioactive conformation of the molecules are computed by selecting, the basic skeleton and conformation of most active molecule **11** (IC₅₀ value 0.018 μ M), to which Gastaiger–Huckel charges were applied and systematic search were carried out to obtain low energy conformer in gaseous phase, which is subsequently energy minimized by Powell method using Tripos force field with 0.05 kcal/mol energy gradient convergence criterion. The rest of the molecules were built by changing required substitution on template molecule **11** and energy minimized as stated previously. These molecules were then used to construct 3D-QSAR model. The IC₅₀ value in micro molar (μ M) range were converted into molar (M) range and then its logarithmic scale (pIC₅₀, M) were then used for subsequent QSAR analysis (Eq. (1)) as the response variable

$$\text{pIC}_{50} = -\log \text{IC}_{50} \quad (1)$$

3.2. Molecular alignment

Molecular alignment is the most sensitive parameter in 3D-QSAR analysis. This renders the spatial alignment of molecules under study as one of the most sensitive and determining factors in obtaining robust and meaningful models. In the present study geometry optimized molecules were aligned on the template molecule **11** (Supplementary Fig. S1) by common substructure alignment using ALIGN DATABASE command in SYBYL. The common substructure used for alignment and superimposed structure after alignment is presented in Fig. 1.

3.3. CoMFA interaction energies

CoMFA steric and electrostatic potential fields were calculated at each lattice intersection on a regularly spaced grid of 2.0 Å units in *x*, *y*, and *z* directions on the aligned dataset. The pattern of 3D cubic lattice generated automatically by SYBYL/CoMFA routine, extended at least 4.0 Å beyond the volumes of all investigated molecules along these axes. The van der Waals potential and Coulombic terms, which represent

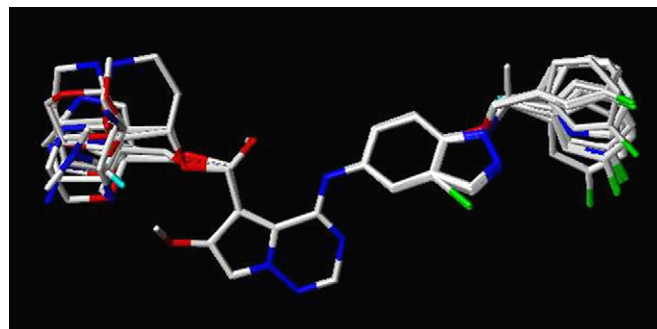


Fig. 1. Alignment of molecules.

steric and electrostatic fields, respectively, were calculated using the standard Tripos force field method. A distance dependent dielectric constant of 1.0 was used and sp^3 hybridized carbon atom with +1 charge served as probe atom

to calculate the steric and electrostatic fields. The steric and electrostatic contributions were truncated to +30.0 kcal/mol and the later were ignored at the lattice intersections with maximal steric interactions.

Table 1

Structures, IC_{50} , pIC_{50} values (actual and predicted) and residuals of C_4 and C_5 substituted pyrrolotriazines (molecules 1–32)

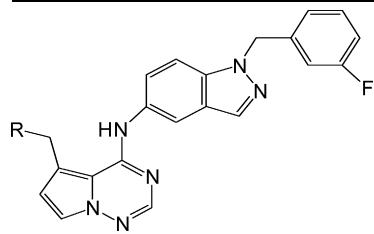
ID	Ar	$^aIC_{50}$ (μ M)	pIC_{50}		Residual
			Actual	Predicted	

5-Methyl pyrrolotriazines with different C_4 substitutions (molecules 1–6)

1		1	6.00	5.96	0.04
2		1	6.00	5.99	0.01
3		0.43	6.37	6.42	−0.04
4		0.11	6.96	6.93	0.03
5 ^b		0.31	6.51	6.26 (6.31) ^c	0.25 (0.20) ^c
6		0.25	6.60	6.63	−0.03

Table 1 (Continued)

ID	R	^a IC ₅₀ (μM)	pIC ₅₀		Residual
			Actual	Predicted	



Structures and activities of C₅ substituted pyrrolotriazines (molecules **7–24**)

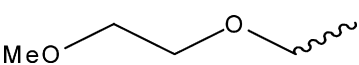
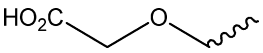
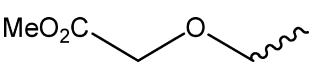
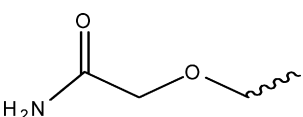
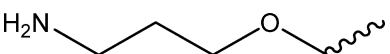
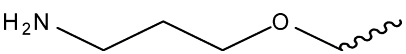
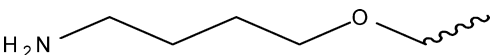
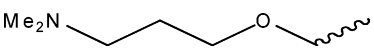
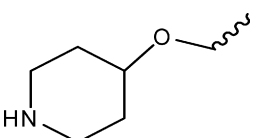
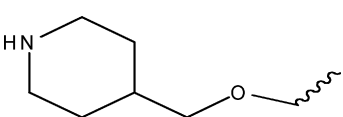
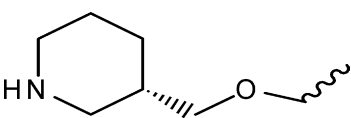
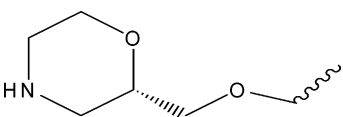
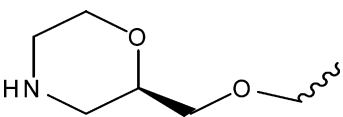
7^b	H	0.088	7.05	7.36 (7.27) ^c	−0.31 (−0.22) ^c
8^b	OH	0.11	6.96	7.25 (7.14) ^c	−0.29 (−0.18) ^c
9^b	MeO	0.42	6.38	7.28 (7.15) ^c	0.9 (−0.77) ^c
10		0.42	6.38	6.38	0.00
11		0.018	7.74	7.70	0.04
12^b		0.28	6.55	7.63 (6.79) ^c	1.08 (0.24) ^c
13		0.11	6.96	6.96	0.00
14		0.09	7.04	7.06	−0.02
15		0.06	7.22	7.26	−0.04
16		0.13	6.89	6.96	−0.07
17		0.56	6.25	6.21	0.04
18		0.07	7.15	7.06	0.09
19		0.12	6.92	6.93	−0.01
20		0.17	6.77	6.74	0.02
21		0.055	7.24	7.15	0.09
22^b		0.060	7.22	6.52 (6.91) ^c	0.70 (−0.31) ^c

Table 1 (Continued)

ID	R	^a IC ₅₀ (μM)	pIC ₅₀		Residual
			Actual	Predicted	
23		0.59	6.23	6.24	−0.01
24 ^b		0.14	6.85	6.52 (6.50) ^c	0.33 (0.35) ^c

ID	Ar	^a IC ₅₀ (μM)	pIC ₅₀		Residual
			Actual	Predicted	
Structures and activities with modification at C ₄ of molecule 21 (molecules 25–32)					
25		0.56	6.25	6.22	0.03
26 ^b		0.38	6.42	6.01 (5.93) ^c	0.41 (0.49) ^c
27		0.41	6.39	6.41	−0.02
28		0.13	6.89	6.97	−0.08
29		0.85	6.06	6.07	−0.01
30		0.041	7.38	7.38	0.00
31		0.51	6.29	6.24	0.05
32		0.25	6.60	6.62	−0.02

^a HER2 sequence expressed in sf9 insect cells, then purified and IC₅₀ value determined.^b Test set molecules (molecules **9** and **12** behave as outlier in case of model A).^c Values in brackets indicates predicted pIC₅₀ values and its residual for test set using model C.

3.4. Partial least square (PLS) analysis

The relationship between the structural parameters (CoMFA interaction energies) and the biological activities has been quantified by the PLS algorithm [12]. The cross-validation analysis was performed using leave-one-out (LOO) method, where one molecule is removed from the dataset and its activity is predicted using the model derived from the rest of the dataset. The cross-validated r_{cv}^2 that resulted in optimum number of components and lowest standard error of prediction was selected. To speed up the analysis with reduced noise, a minimum column filtering value of 2.0 kcal/mol was used for the cross-validation. Final analysis (non-cross-validation) was performed to calculate non-cross-validated (r^2) using the optimum number of components obtained from the LOO cross-validation analysis. To further assess the statistical confidence and robustness of derived model, a 100-cycle bootstrap analysis was performed.

3.5. Predictive correlation coefficient (r_{pred}^2)

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

$$r_{pred}^2 = \frac{\text{S.D.} - \text{PRESS}}{\text{S.D.}} \quad (2)$$

where S.D. is defined as the sum of the squared deviations between the biological activities of the test set and mean activities of the training set molecules and PRESS is the sum of the squared deviation between predicted and actual activity values for every molecule in the test set.

4. Results and discussion

4.1. CoMFA model A

The CoMFA analysis was performed to explore the structure–activity relationship of C₅ substituted pyrrolotriazines dual inhibitors of HER2 and EGFR protein tyrosine kinases. The dataset consisting of 32 molecules was divided into training set of 24 molecules and test set of 8 molecules by considering the structural diversity and activity range. During the rigorous cycle of model development and validation, we found two molecules **9** and **12** as outlier (OL) not to fit either of the training set or test set. These molecules (**9** and **12**) were removed from the dataset to construct final CoMFA model A. Initially, CoMFA analysis was performed on a dataset of 32 molecules with inclusion of molecules **9** and **12**. Many models were generated by taking various combination of training and test set molecules. It was observed that inclusion of molecules **9** and **12** in training set drastically decreases r_{cv}^2 and which in turn

decreases robustness of the model. Their inclusion in the test set showed increase in r_{cv}^2 , but model showed poor predictive power with high residual for molecules **9** and **12** (Table 1). Possible reason for such behavior is that molecule **9** an ether analogue and molecule **12** a neutral ester, which are different from the rest of the molecules included in the training and test set and is also weak HER2 inhibitors. OL generally occurs when their mode of action is different, though the structure in this case is not that unique with the structure of other molecules in the series or also depend upon the quality of the biological data for QSAR analysis.

The statistical results of CoMFA PLS analysis of model A is presented in Table 2. The constructed CoMFA model A is robust with r_{cv}^2 of 0.627 and conventional r^2 of 0.992 having 7 optimum number of components (NOC). The r_{cv}^2 represents goodness of internal prediction whereas r^2 represent the goodness of fit of a QSAR model. The external predictive capability of a QSAR model is generally checked using test set molecules. All other procedures including molecule geometry optimization, charge calculation, interaction energy calculation and alignment of the test set molecules were done in manner analogues to the training set molecules. The predictive power (r_{pred}^2) of developed CoMFA model A is good having 0.650 value. Further the statistical validity and stability of the CoMFA model A was assessed by running bootstrap analysis for 100 runs. The higher r_{bs}^2 value of 0.996 obtained after 100 runs of bootstrapping supports this analysis [13]. In addition to this low value of PRESS (0.519) and standard error of estimates (0.051) supports the significance of the developed model. The steric to electrostatic fields were found to be 57:43 which implies that the contribution of steric part is predominant for interaction of these inhibitors to HER2.

The scatter plots for actual versus predicted activities of training and test set molecules were shown in Fig. 2a and b, respectively. Table 1 shows the structures and the corresponding IC₅₀ (μM), actual and predicted pIC₅₀ values of the

Table 2
PLS result summary for CoMFA models A and C

	Model A	Model C
PRESS	0.519	0.660
r_{cv}^2	0.627	0.609
NOC	7	7
SEE	0.051	0.047
r^2	0.992	0.993
F-test	284.99	325.84
r_{bs}^2	0.996	0.996
S.D. _{bs}	0.004	0.004
r_{pred}^2	0.650	0.532
Fraction of field contributions		
Steric		0.57
Electrostatic		0.43

PRESS = predictive residual sum of squares for the training set; r_{cv}^2 = cross-validated correlation coefficient by PLS LOO method; NOC = optimum number of components as determined by PLS LOO cross-validation study; SEE = standard error of estimate; r^2 = conventional correlation coefficient; r_{bs}^2 = correlation coefficient after 100 runs of bootstrapping; S.D._{bs} = standard deviation from 100 runs of bootstrapping; r_{pred}^2 = predictive correlation coefficient.

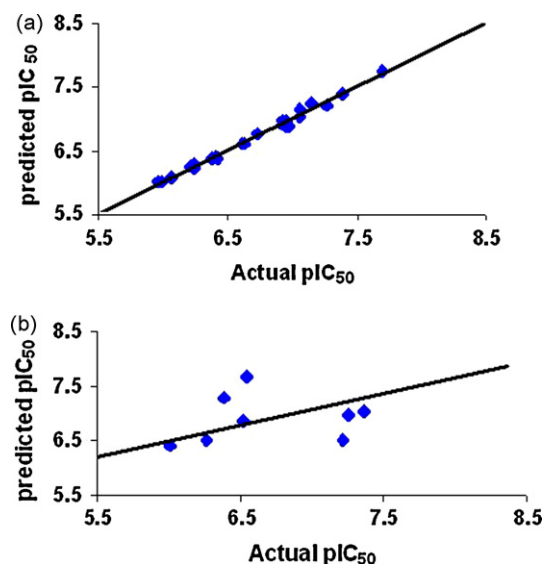


Fig. 2. Scatter plot of actual versus predicted pIC_{50} of the training set (a) and test set (b) molecules obtained using model A.

molecules under study for CoMFA model development. All the aligned molecules are shown in Fig. 1.

4.2. CoMFA model B

We have developed CoMFA model B by including compounds **9** and **12** in the training set of model A, to monitor the corresponding variation in its PLS statistics. The CoMFA model B showed poor predictive power with r^2_{cv} of 0.104, high standard error of estimation 0.362, r^2 value of 0.404 and F -value of 16.29.

4.3. CoMFA model C

Third model was created by giving special emphasis on the conformation of the side chain of the compounds **9** and **12** in order to check the sensitivity of the CoMFA model with respect to its molecular conformation. As mentioned earlier compounds **9** and **12** behaved as outlier for model building (CoMFA models A and B), and we thought it to be challenging to get a scientific insight as whether this C_5 side chain do play an important role in predicting activity of compounds. A systematic conformation approach in search for proper C_5 side chain conformation of the molecules **9** and **12** were performed, and the generated conformation was selected for model building. The molecules **9** and **12** with the latest conformation generated using systematic conformation approach was tested by including in the training set of CoMFA model A. This intermediate model showed improvement in PLS statistics in comparison to that of model B, but was not as robust like model A. Therefore we decided to analyse these molecules **9** and **12** in the test set, which lead to meaningful prediction with its actual activity value showing less residual (Table 1). Therefore, CoMFA model C is developed and its PLS statistics is presented in Table 1, viz. r^2_{cv} of 0.609 and conventional r^2 of 0.993 having 7 optimum number of components. The predictive power of developed CoMFA model

C is good, showing r^2_{pred} of 0.532, and also showed good prediction values for **9** and **12** along with other test set molecules (Table 2). The higher r^2_{bs} value of 0.996 obtained after 100 runs of bootstrapping supports this analysis. In addition to this low value of PRESS (0.660) and standard error of estimates (0.047) supports the significance of the developed model. The counter map for CoMFA model C is the same as that of model A. In addition, there is only minor variation in training set pIC_{50} values of models A and C (training set predicted pIC_{50} values for model C not shown). From the above three CoMFA models A, B, and C it was observed that molecules **9** and **12** add noise to model A when included in training set (model B), while there inclusion in test set after systematic conformation search (model C), showed good prediction, without any influence on counter map, as expected. Both models A and C showed good predictive power and model C can be used as alternative to model A. The scatter plots for actual versus predicted activity of training and test set molecules using model C were presented as Supplementary Fig. S2a and b, respectively.

4.4. External data set evaluation of CoMFA models A and C

In order to establish further the robustness of the CoMFA models A and C, 20 novel C-5 aminomethyl pyrrolotriazine dual inhibitors of EGFR and HER2 protein tyrosine kinases, reported recently by Mastalerz et al. [14] were evaluated. Structures with actual and predicted pIC_{50} values for these external dataset of 20 molecules (**33–52**) are shown in Table 3. The models A and C showed good predictivity with low residuals for these 20 molecules (Table 3 and Supplementary Table S1), which further support reliability of our CoMFA models for different classes of HER2 inhibitors.

4.5. Contour map analysis

CoMFA model is usually represented as 3D coefficient contour maps which surround all lattice point where QSAR is found to strongly deals with the change in interaction energy or binding affinity with respect to the structural changes. In simple words, contour maps shows how the variation of steric or electrostatic property in structural features of molecules contained in the training set leads to increase or decrease in activity. The polyhedra produced surround lattice points where the scalar products of the associated QSAR coefficient and the standard deviation of all values in the corresponding column of the data matrix are higher or lower than a user-specified value. SYBYL setting used different color to represent electrostatic contour (blue and red) and steric (yellow and green) contour. In case of electrostatic contour map, region where negative charge leads to increase in activity is represented by red color, while region with favorable positive charge is represented as blue color. On the other hand, in case of steric contour, yellow color region represents sterically disfavored point, and addition of bulky group at that point leads to decrease in activity. While green color regions represent sterically favored point, where the bulky group substitution at this region could lead to increase in the activity of the molecule.

Table 3

External dataset molecules (C₅ aminomethyl analogs) with their actual and predicted pIC₅₀ value obtained using CoMFA model A

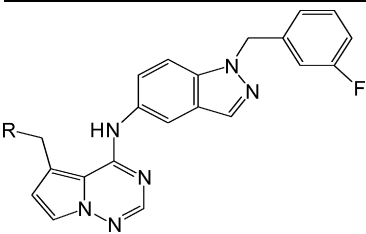
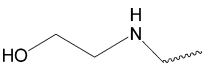
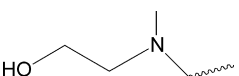
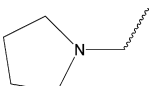
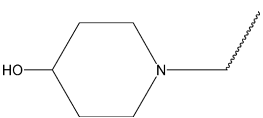
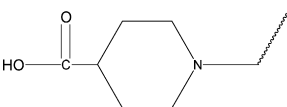
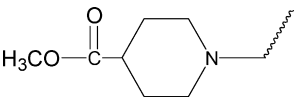
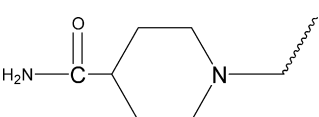
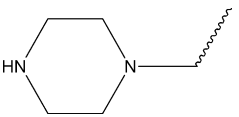
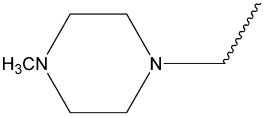
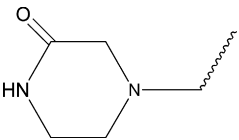
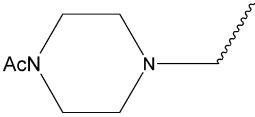
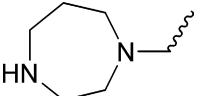
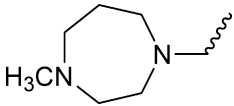
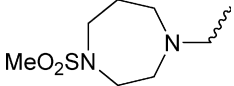
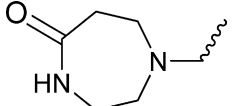
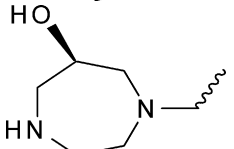
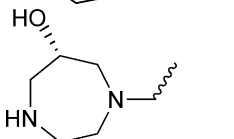
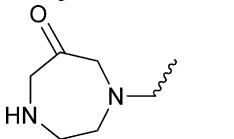
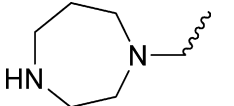
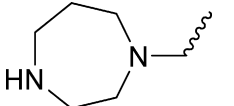
S. No.	R	^a IC ₅₀ (μM)	Actual pIC ₅₀	Predicted pIC ₅₀	Residual
					
33		0.047	7.33	7.26	−0.07
34		0.11	6.96	7.29	0.34
35		0.36	6.44	7.18	0.74
36		0.06	7.20	6.75	−0.45
37		0.11	6.96	6.28	−0.68
38		1.00	6.00	6.19	0.19
39		0.31	6.51	6.94	0.43
40		0.081	7.09	7.07	−0.02
41		0.12	6.92	7.21	0.29
42		0.20	6.70	7.25	0.55
43		0.61	6.21	6.30	0.09
44		0.027	7.51	7.21	−0.30

Table 3 (Continued)

S. No.	R	^a IC ₅₀ (μM)	Actual pIC ₅₀	Predicted pIC ₅₀	Residual
45		0.12	6.92	7.20	0.28
46		0.12	6.92	7.25	0.33
47		0.075	7.12	7.05	−0.07
48		0.016	7.77	7.10	−0.67
49		0.076	7.49	7.12	−0.37
50		0.039	7.11	6.97	−0.14
51		0.17	6.77	7.29	0.52
52		0.034	7.13	7.21	0.08

^a From Ref.: [14].

The electrostatic and steric contours for most active molecule (molecule **11**) are displayed in Fig. 3a and b, respectively. The electrostatic contour map shown in Fig. 3a displayed one blue and two red polyhedrons. The small red polyhedrons on carbonyl oxygen of C5 solubilizing group suggest that electronegative group is favorable at this point. The blue polyhedron on terminal oxygen depicts that electropositive group with optimal bulk is required at this position to show optimal activity. This explains good activity of molecule **21** with morpholine group. The basic electropositive nitrogen in morpholine ring is known to be first protonated and then forms a hydrogen bond with Cys805 (Cys773 in EGFR) and Asp808 (Asp776 in EGFR) in ribose phosphate pocket of HER2 protein, which is in accordance with the fact that active site residues in HER2 and EGFR are conserved. The sequence alignment of HER2 and EGFR showed full conservation of active site residues (having only difference in the residue numbering) [15] so we considered active site of HER2 in accordance with EGFR for further discussion. In addition to this the C₅ side chain of quinazoline HER2 and EGFR kinase inhibitors have recently been hypothesized to similarly extend into the ribose portion of the ATP binding pocket [9].

The most active molecule in the series, molecule **11**, is acid and show good kinase inhibition and docking study suggested that acid group could wither interact with the active site of EGFR i.e. Lys745 (Lys753 in HER2) normally involved in binding the phosphate group of ATP and Cys773 (Cys805 in HER2) of the solvent exposed hinge region. Molecule **12** which is neutral ester and molecules **9** and **10**, an ether analog, have methyl substitution on the terminal oxygen which is unable to form hydrogen bonds with the ribose phosphate pocket of HER2 and this would lead to its low activity. There is presence of small red contour near fluorine (Fig. 3a) which shows that substitution of electronegative group at that point leads to an increase in activity. The steric contour plot presented in Fig. 3b showed a big green polyhedron that surrounds metafluoro benzyl group, indicating sterically favorable region. This accounts for better activity of compound **11** than that of molecules **1** and **2**. In case of molecule **11**, the bulky benzyl indazole group is known to be extended into deep hydrophobic pocket showing an increase in activity, while in case of molecules **1** and **2** there is complete absence of substituted benzyl group, which leads to its drastic decrease in activity. In case of molecules **3–7**, it was observed that as there is substitution of bulkier benzyl indazole

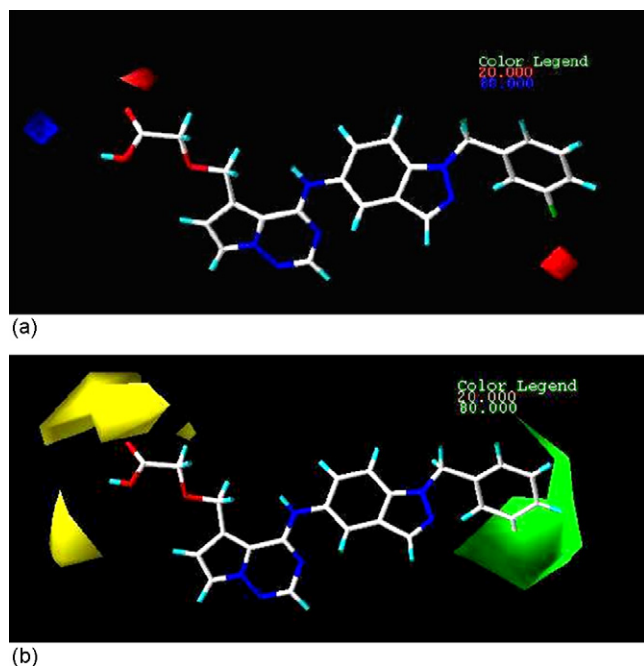


Fig. 3. (a) CoMFA electrostatic contour map together with molecule 11. Red contours indicate negative charge increase binding affinity, whereas blue contours indicate positive charge favoring binding affinity. (b) CoMFA steric contour map together with molecule 11. Green contours show favorable bulky group substitution at that point while yellow region shows unfavorable bulky group for activity.

group in the molecule and showed relative increase in activity of the order of $3 < 5 < 6 < 4 < 7$, due to the above mentioned favorable interactions. In addition, it also showed yellow color regions at C₅ substituted group. The large yellow polyhedron on carbonyl oxygen and small yellow polyhedron on methylene previous to it showed sterically disfavored point. The terminal hydrogen attached to 'oxygen' in molecule 11 also showed small yellow region, which would suffice that substitution of methyl group in place of hydrogen leads to decrease in activity of molecules 9, 12 and 17.

5. Conclusions

To explore the structure–activity relationships of C₅ substituted pyrrolotriazines for their HER2 inhibitory activity, and to build statistically significant model with good predictive power, 3D-QSAR CoMFA study was performed. The developed CoMFA model showed good predictive power with low residuals for test set molecules, as well as for recently synthesized external dataset of 20 molecules. Counter map analysis showed that addition of lipophilic group at C₄ position is favorable for activity, which is known to be interacting with hydrophobic pocket in receptor. Also these C₅ substituted analogues showed different range of activity, viz., the ester and ether analogues showed less activity, while acid and morpholine ring substituted analogues showed good activity, which is associated with the ability of side chain substituent to hydrogen bond with the ribose phosphate pocket of the protein. Thus, CoMFA contour analysis has provided useful information about the structural requirements for the observed HER2 inhibitory

activity. This *in silico* analysis could broadly support in the rational design of potential drug candidates with an improved inhibitory activity.

Acknowledgements

One of the author CGM acknowledges DST, New Delhi and Department of Biotechnology (IFD-Dy. No. 102/IFD/SAN/884/2006-2007), New Delhi, India for financial support on this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2007.10.008.

References

- [1] S. Menard, P. Casalini, M. Campiglio, S.M. Pupa, E. Tagliabue, Role of HER2/neu in tumor progression and therapy, *Cell Mol. Life Sci.* 61 (2004) 2965–2978.
- [2] A. Jimeno, M. Hidalgo, Pharmacogenomics of epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors, *Biochim. Biophys. Acta (BBA)-Rev. Cancer* 1766 (2006) 217–229.
- [3] R. Worthylake, L.K. Opresko, H.S. Wiley, ErbB-2 amplification inhibits down-regulation and induces constitutive activation of both ErbB-2 and epidermal growth factor receptors, *J. Biol. Chem.* 274 (1999) 8865–8874.
- [4] M. Van de Vijver, Genetic alterations in intraductal and invasive breast cancer, *Eur. J. Cancer* 37 (2001) 12–112.
- [5] Y. Yarden, M.X. Sliwkowski, Untangling the ErbB signalling network, *Nat. Rev. Mol. Cell Biol.* 2 (2001) 127–137.
- [6] J.T. Hunt, T. Mitt, R. Borzilleri, J. Gullo-Brown, J. Fargnoli, B. Fink, W.C. Han, S. Mortillo, G. Vite, B. Wautlet, Discovery of the pyrrolo [2,1-f][1,2,4] triazine nucleus as a new kinase inhibitor template, *J. Med. Chem.* 47 (2004) 4054–4059.
- [7] W.-m. Shi, Q. Shen, W. Kong, B.-x. Ye, QSAR analysis of tyrosine kinase inhibitor using modified ant colony optimization and multiple linear regression, *Eur. J. Med. Chem.* 42 (2007) 81–86.
- [8] S. Kamath, J.K. Buolamwini, Receptor-guided alignment-based comparative 3D-QSAR studies of benzylidene malonitrile tyrphostins as EGFR and HER-2 kinase inhibitors, *J. Med. Chem.* 46 (2003) 4657–4668.
- [9] H. Mastalerz, M. Chang, P. Chen, P. Dextraze, B.E. Fink, A. Gavai, B. Goyal, W.C. Han, W. Johnson, D. Langley, F.Y. Lee, P. Marathe, A. Mathur, S. Oppenheimer, E. Ruediger, J. Tarrant, J.S. Tokarski, G.D. Vite, D.M. Vyas, H. Wong, T.W. Wong, H. Zhang, G. Zhang, New C-5 substituted pyrrolotriazine dual inhibitors of EGFR and HER2 protein tyrosine kinases, *Bioorg. Med. Chem. Lett.* 17 (2007) 2036–2042.
- [10] R.D. Cramer III, D.E. Patterson, J.D. Bunce, Effect of shape on binding of steroids to carrier proteins, *J. Am. Chem. Soc.* 110 (1988) 5959–5967.
- [11] SYBYL7.1, Tripos Inc., St. Louis, MO 63144, USA.
- [12] S. Wold, A. Johansson, M. Cochi, PLS-partial least square projection to latent structures, in: H. Kubinyi (Ed.), *3D-QSAR in Drug Design: Theory, Methods and Applications*, ESCOM, Leiden, 1993, pp. 523–550.
- [13] R.D. Cramer III, J.D. Bunce, D.E. Paterson, I.E. Frank, Cross-validation, bootstrapping and partial least squares compared with multiple regression in conventional QSAR studies, *Quant. Struct. -Act. Relat.* 7 (1988) 18–25.
- [14] H. Mastalerz, M. Chang, A. Gavai, W. Johnson, D. Langley, F.Y. Lee, P. Marathe, A. Mathur, S. Oppenheimer, J. Tarrant, Novel C-5 aminomethyl pyrrolotriazine dual inhibitors of EGFR and HER2 protein tyrosine kinases, *Bioorg. Med. Chem. Lett.* 17 (2007) 2828–2833.
- [15] J. Singh, E.M. Dobrusin, D.W. Fry, T. Haske, A. Whitty, D.J. McNamara, Structure-based design of a potent, selective and irreversible inhibitors of the catalytic domain of erbB receptor subfamily of protein tyrosine kinases, *J. Med. Chem.* 40 (1997) 1130–1135.