

Original article:

**COMBINED 3D-QSAR MODELING AND MOLECULAR DOCKING
STUDY ON MULTI-ACTING QUINAZOLINE DERIVATIVES AS
HER2 KINASE INHIBITORS**

Sako Mirzaie^{1*}, Majid Monajjemi², Mohammad Saeed Hakhamaneshi³, Fardin Fathi⁴,
Mostafa Jamalani¹

¹ Department of Biology, Faculty of Science, Science and Research Branch, Islamic Azad University, Tehran, Iran

² Department of Chemistry, Faculty of Science, Science and Research Branch, Islamic Azad University, Tehran, Iran

³ Department of Biochemistry, Faculty of Medicine, Kurdistan University of Medical Science, Sanandaj, Iran

⁴ Cellular and Molecular Research Center, Kurdistan University of Medical Sciences, Sanandaj, Iran

* corresponding author: sako.biochem@gmail.com

ABSTRACT

A series of new quinazoline derivatives has been recently reported as potent multi-acting *histone deacetylase* (HDAC), epidermal growth factor receptor (EGFR), and human epidermal growth factor receptor 2 (HER2) inhibitors. HER2 is one of the major targets for the treatment of breast cancer and other carcinomas. Three-dimensional structure-activity relationship (3D-QSAR) is a well-known technique, which is used to drug design and development. This technique is used for quantitatively predicting the interaction between a molecule and the active site of a specific target. For each 3D-QSAR study, a three-dimensional model is created from a large curve fit to find a fitting between computational descriptors and biological activity. This model could be used as a predictive tool in drug design. The best model has the highest correlation between theoretical and experimental data. Self-Organizing Molecular Field Analysis (SOMFA), a grid-based and alignment-dependent 3D-QSAR method, is employed to study the correlation between the molecular properties and HER2 inhibitory potency of the quinazoline derivatives. Before presentation of inhibitor structures to SOMFA study, conformation of inhibitors was determined by AutoDock4, HyperChem and AutoDock Vina, separately. Overall, six independent models were produced and evaluated by the statistical partial least square (PLS) analysis. Among the several generated 3D-QSARs, the best model was selected on the basis of its statistical significance and predictive potential. The model derived from the superposition of docked conformation with AutoDock Vina with reasonable cross-validated q^2 (0.767), non cross-validated r^2 (0.815) and F-test (97.22) values showed a desirable predictive capability. Analysis of SOMFA model could provide some useful information in the design of novel HER2 kinase inhibitors with better spectrum of activity.

Keywords: quinazoline, human epidermal growth factor receptor 2, 3D-QSAR, self-organizing molecular field analysis, AutoDock

INTRODUCTION

The epidermal growth factor receptor (EGFR) is the cell-surface receptor of the

epidermal growth factor family (Herbst, 2004). This family comprises four structurally related tyrosine kinase receptors, name-

ly, epidermal growth factor receptor 1 (alternatively named EGFR, ErbB1) or human epidermal growth factor receptor 1 (HER1), HER2 (ErbB2), HER3 (ErbB3) and HER4 (ErbB4) (Cho et al., 2003; Satyanarayanan et al., 2009; Woodburn, 1999; Yarden et al., 2004). ErbB receptors are made up of an *extracellular* region or ectodomain that provides a ligand-binding site, a single *transmembrane*-spanning region, and a cytoplasmic *tyrosine kinase* domain (Garrett et al., 2002). When a certain ligand binds to the extracellular domain of an ErbB receptor, its dimerization is triggered with other adjacent ErbB. There is not any known ligand for HER2 and instead this receptor acts as a co-receptor for each of the other three (Linggi and Carpenter, 2006). Dimerization leads to a rapid activation of receptor's cytoplasmic kinase domain (Satyanarayanan et al., 2009). Activated form of the ErbB receptors plays a critical role in the regulation of cell growth, proliferation and migration. Kinase domain subsequently catalyzes receptor autophosphorylation and tyrosine phosphorylation of ErbB substrates (Satyanarayanan et al., 2009; Pawson, 1995). It has been shown that HER2 overexpression is involved in colon, ovarian, non-small-lung cell and approximately 25-30 % of breast cancers (Shmeeda et al., 2009). Furthermore, amplification content is conversely correlated with the observed average survival time (Carter et al., 1992).

Breast cancer is one of the major health problems among developed community. It is the most common neoplasia among women, in which more than 1 million new cases occur every year, and it is the first cause of death in women aged 40–59 years old (Galvao et al., 2011). Just in 2010, 207,090 new cases are estimated to occur in United States with an expected mortality rate of 39,840 women (Galvao et al., 2011; Jemal et al., 2010). Obstruction of HER2-mediated multimerization leads to inhibition of phosphorylation and subsequently, uncontrolled cell growth and division will limit. So, blocking of HER2-mediated signaling has made it a key target of breast

cancer therapies (Satyanarayanan et al., 2009). Trastuzumab (Herceptin) is an FDA-approved humanized monoclonal antibody, which is used to treat the early stage of HER2 overexpressed breast cancer. It binds to the ectoplasmic domain of HER2 and inhibits any receptor dimerization (Burgess et al., 2003). Use of trastuzumab as an effective drug, because of its cost, is still controversial. In the U.S., 34-week treatments for every case (average duration of treatment) costs approximately 23,400 USD (Stebbing et al., 2000). Recent studies have also shown that cancer treatment by trastuzumab may be contributed in one kind of drug-induced cardiac dysfunction (Zeglinksi et al., 2011). In contrast with trastuzumab, chemical new inhibitors like erlotinib and lapatinib were designed and synthesized against cytoplasmic kinase domain of HER2 (Cai et al., 2010). Because of high identity and similarity between HER2 kinase domain and the other more than one hundred kinases in cells, a highly specific kinase domain inhibitor is challenging. On the other side, drug resistance could occur due to ATP-binding site mutations (Satyanarayanan et al., 2009). So, it seems that the need to design and synthesis of a more potent and specific HER2 kinase inhibitor is urgent.

Quantitative structure-activity relationships (QSAR) have played an influential role in the design of pharmaceuticals in medicinal chemistry (Akamatsu, 2002). In this technique, variation in biological activities is statistically analyzed using descriptors related to structural properties of compounds. The self-organizing molecular field analysis (SOMFA) is a simple three-dimensional quantitative structure-activity relationship (3D-QSAR) technique. In this method, the molecular shape and electrostatic potential are used to construct the QSAR models (Robinson et al., 1999). Recent study reports the first published crystal structure of the kinase domain of HER2, which could be a valuable target to design the new anti-HER2 drugs (Aertgeerts et al., 2011). The purpose of this paper is to de-

scribe the application of molecular docking studies and self-organizing molecular field analysis, SOMFA, on a series of quinazoline derivatives. The worthwhile molecular architecture required for designing new specific inhibitors against HER2 positive breast cancer is also discussed.

MATERIALS AND METHODS

Overall strategy

To produce the SOMFA model, minimized and correct geometrical conformations of quinazoline derivatives are needed. For this purpose, three independent categories of inhibitor conformations were produced by AutoDock 4.2 (Morris et al., 2009), HyperChem 8.0 (Froimowitz, 1993) and AutoDock Vina (Trott and Olson, 2010). For each category, compound 1 was selected as the reference compound and the other compounds were superimposed by atom based alignment technique (Figure 1). In this method, superposition of molecules is based on trying to minimize root-mean-squares (RMS) differences in the fitting of selected atoms with those of a reference molecule (Figure 2). Finally, for each category, related SOMFA models were pro-

duced by SOMFA software (Figure 3) (Robinson et al., 1999).

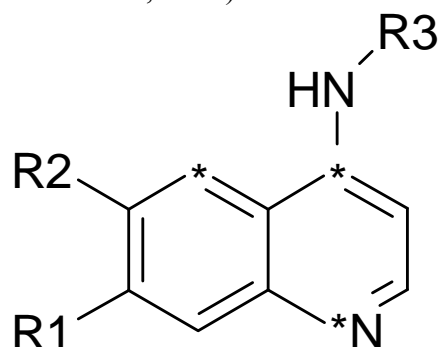


Figure 1: Atoms used for atom based alignment are specified by an asterisk data set.

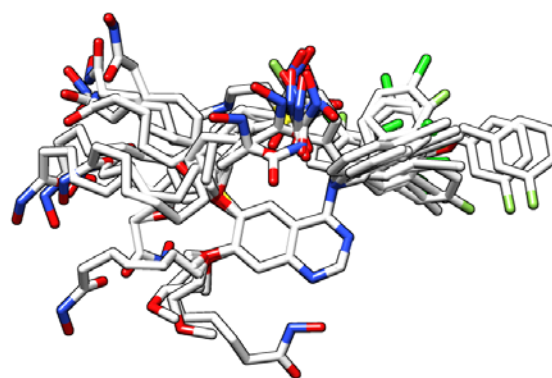


Figure 2: Superposition of all structures on reference compound. For ease of visualization, hydrogen atoms are not shown.

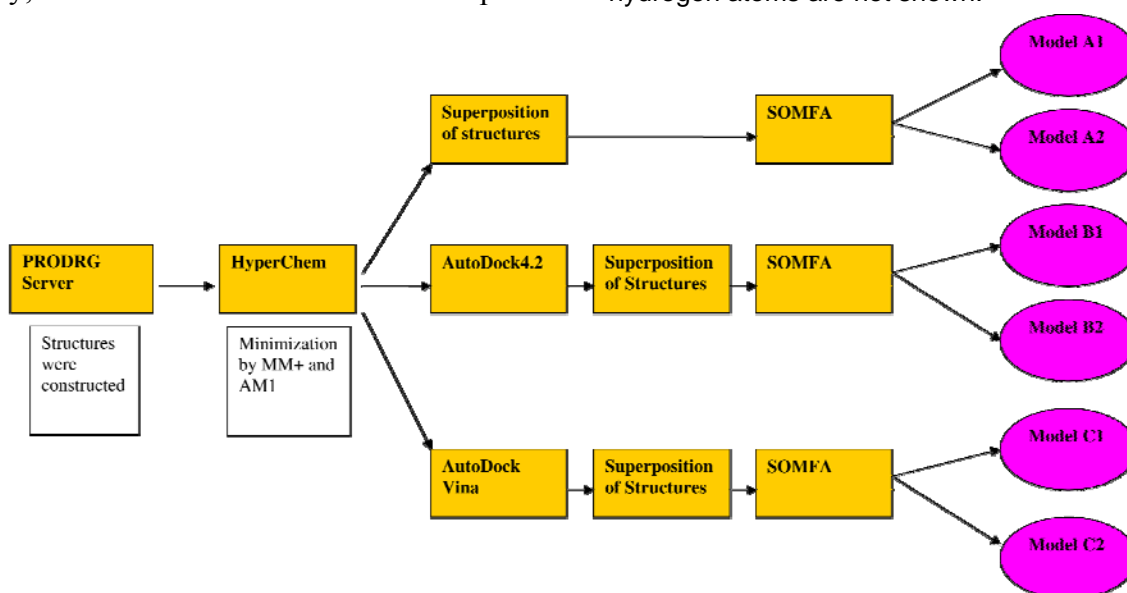
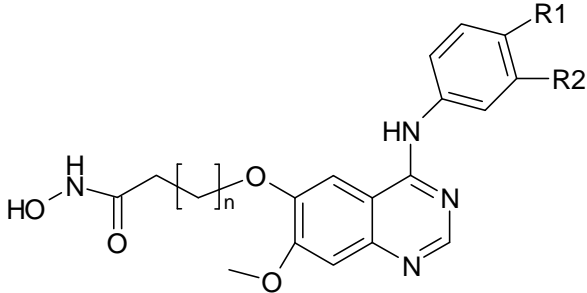
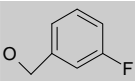


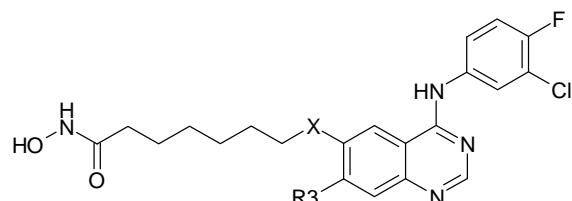
Figure 3: Flow chart of 3D-QSAR modeling procedure which is used in this study. Based on molecular docking and modeling software and SOMFA resolution utilized, for each category two models were generated. Models A1, B1 and C1 have resolution 1, whereas, resolution of 0.5 was assigned for model A2, B2 and C2.

In this study, a data set of 24 compounds belonging to the quinazoline derivatives as HER2 inhibitors were taken from the literature (Cai et al., 2010) and used to produce SOMFA model (Table 1). The chemical structures of mentioned compounds were generated by PRODRG server (Schuttelkopf and van Aalten, 2004). For optimization of constructed compounds, molecular mechanic (MM+) optimization followed by the semi-empirical AM1 (Austin Model 1) method implemented in HyperChem software were used. For both the above mentioned methods, molecular structures were optimized using the Polak-Ribiere (conjugate gradient) algorithm until the root mean square (RMS) gradient achieves a value smaller than $0.001 \text{ kcal mol}^{-1}$.

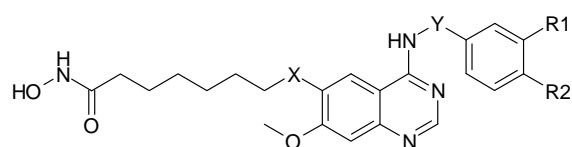
Table 1: Structural and activity data of the quinazoline derivatives studied (Cai et al., 2010)



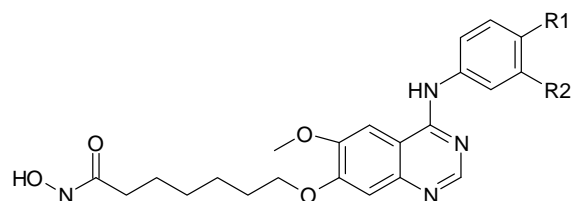
Compound	n	R1	R2	IC ₅₀ (nM) in enzyme assay*
1	1	F	Cl	10.6
2	2	F	Cl	12.4
3	3	F	Cl	15
4	3	H	C≡CH	38.4
5	4	F	Cl	20
6	4	H	C≡CH	38.2
7	5	F	Cl	19
8	5	H	C≡CH	15.7
9	5	F	C≡CH	76.5
10	5		Cl	10.1
erlotinib				134.5
lapatinib				10.2



Compound	X	R3	IC ₅₀ (nM) in enzyme assay*
11	O	H-	43.5
12	O	CH ₃ OCH ₂ CH ₂ O-	24
13	CONH	CH ₃ O-	20.1
14	S	CH ₃ O-	24.2



Compound	Y	R1	R2	IC ₅₀ (nM) in enzyme assay*
15	CH ₂	F	Cl	478.2
16	CH ₂	H	H	225.2
17	CH ₂	F	H	295.4
18	(R)-CHCH ₃	H	H	27.8
19	(R)-CHCH ₃	F	H	175
20	(R)-CHCH ₃	Cl	H	556.9



Compound	R1	R2	IC ₅₀ (nM) in enzyme assay*
21	F	Cl	478.2
22	H	C≡CH	225.2

*enzyme assay was performed by HTScan® HER2/ErbB2 Kinase Assay Kit.

Biological activity

The negative logarithm of the measured IC₅₀ (M) against human HER2 kinase domain as pIC₅₀ was used as dependent variable (Cai et al., 2010). HER2 kinase activity was measured using HTScan® HER2/ErbB2 Kinase Assay Kit (Cell Signaling Technology). It includes active HER2/

ErbB2 kinase (supplied as a GST fusion protein), a biotinylated peptide substrate and a phospho-tyrosine antibody for detection of the phosphorylated form of the substrate peptide. The enzyme assay was performed based on fluorescent immuno-detection approach (Cai et al., 2010).

Protein preparation

3D crystallographic structure of the kinase domain of human HER2 (ErbB2) with PDB code 3PPO (Aertgeerts et al., 2011) was obtained from the Brookhaven protein data bank (Berman et al., 2000). For use of kinase domain structure in docking simulation, all missing hydrogen atoms were added by Hbuild command implemented in CHARMM molecular dynamic package (Brooks et al., 2009). All water molecules were eliminated except for water ID 22, which is tightly bounded to the active site of HER2 kinase domain. After initial minimization by Adopted Basis Newton-Raphson (ABNR) and steepest descent (SD) methods, output structure was saved and used for molecular docking studies.

Molecular docking studies

AutoDock 4.2

Autodock is a flexible docking technique which is based on Monte Carlo simulated annealing search to find the optimal conformation of ligand in a macromolecule. Inhibitor structures minimized by HyperChem were used as input ligand files for AutoDock 4.2. Ligand preparation was done by AutoDockTools. The grid maps were calculated using AutoGrid (part of the AutoDock package). Because the active site of HER2 kinase domain is known (Aertgeerts et al., 2011), a grid map with $25 \times 25 \times 25$ points with a grid-point spacing of 1 Å (roughly a quarter of the length of a carbon-carbon single bond) were defined. For both protein and ligands, the partial atomic charges were calculated using the Gasteiger-Marsili method (Gasteiger and Marsili, 1980). For all dockings, 50 independent runs with the step sizes of 0.2 Å for translations and 5° for orientations and tor-

sions were performed. For each ligand, pose with the minimum free energy of binding was saved and used for SOMFA study.

AutoDock Vina

AutoDock Vina is a molecular docking software that profits of knowledge-based potentials and empirical scoring functions (Trott and Olson, 2010). Minimized inhibitor structures were defined as input files for AutoDock Vina. The docking site for inhibitors on 3PPO was defined by setting up a box with a grid map with $25 \times 25 \times 25$ points (grid-point spacing of 1 Å) at the geometrical center of the native ligand present in PDB structure. Software parameter "Exhaustiveness", which determines how comprehensively the program searches for the lowest energy conformation, was set to the default value, eight for all docking runs. For each docking calculation, structure with relative binding energy, based on the software cluster, was saved and employed to SOMFA study.

SOMFA 3D-QSAR study

To produce a 3D-QSAR model, SOMFA2 package software was used. This suite of programs is designed to enable the user to carry out an advanced 3D-QSAR analysis of a series of molecules. The results of the analysis enable the rationalization of the activities of known molecules together with the potential for reliably estimating the activity of molecules that have yet to be synthesized (Robinson et al., 1999). In our SOMFA study, a $40 \times 40 \times 40$ Å grid originating at (-20, -20, -20) with a two resolution (1 and 0.5 Å) were generated around the aligned compounds.

Six different models using different resolutions and aligned conformations of inhibitors have been produced and presented in Table 2. For all the studies, shape and electrostatic potential were produced as an independent variable. These two variables, consequently, combined to create the final model by utilizing the partial least square (PLS) algorithm in conjugation with leave one out (LOO) cross-validation method.

Based on equation 1, weighted average of shape and electrostatic were calculated:

$$\text{Activity}_{\text{ESP}} = c_1 \text{Activity}_{\text{shape}} + (1 - c_1) \text{Activity}_{\text{ESP}} \quad (1)$$

where c_1 is a mixing coefficient.

The predictive ability of the model is quantitated in terms of q^2 (r^2_{cv}) which is defined in equation 2:

$$q^2 = 1 - \text{PRESS}/\text{SSD} \quad (2)$$

where SSD is the sum of squared deviations between observed values and the mean of predicted property values. PRESS is the sum of squared deviations between observed and predicted property values:

$$\text{PRESS} = \sum (Y_{\text{observed}} - Y_{\text{predicted}})$$

q^2 takes up values in the range from 1 to less than 0. Low or negative values of q^2 indicate the errors of prediction are greater than the error from assigning each compound mean activity of the model. Higher values of q^2 is necessary but not sufficient for evaluating a QSAR model (Golbraikh and Tropsha, 2002). One of the several variance-related parameters that can be used as a criterion of the level of statistical significance of the regression model is Fischer statistic parameter (F-value). The larger value of F implies that a more significant

correlation has been reached (Kulkarni et al., 1999). Because the final equations are not very useful to efficiently denote the SOMFA models, 3D master grid maps of the best models are displayed by Grid-Visualizer program. These grids represent an area in space where steric and electrostatic field interactions are responsible for the observed variations of the biological activity.

RESULTS AND DISCUSSION

Numerous small organic compounds have been synthesized and evaluated as inhibitors of EGFR and HER2 tyrosine kinases (Traxler et al., 1997; de Bono and Rowinsky, 2002). Since, quinazoline derivatives could inhibit the HER2 and account for potent HER2/EGFR inhibitors (Cai et al., 2010), 3D quantitative structure activity relationship study with aid of molecular docking method was done to develop more influential HER2/EGFR mediated anti-cancer drugs. In the first phase of this study, to ensure the docking protocol is set up correctly; an internal validation phase was done.

Table 2: Statistical characteristics of the 3D QSAR models for selectivity

Model	Modeling software*	Resolution	r^2	s	F	c_1	c_2	q^2	r
A1	AutoDock4	1	0.722	0.289	56.999	0.7	0.3	0.685	0.849
A2	AutoDock4	0.5	0.702	0.3	51.745	0.7	0.3	0.658	0.838
B1	HyperChem	1	0.731	0.284	59.931	0.7	0.3	0.697	0.855
B2	HyperChem	0.5	0.722	0.289	57.124	0.7	0.3	0.686	0.850
C1	AutoDock Vina	1	0.813	0.237	95.838	0.6	0.4	0.765	0.902
C2	AutoDock Vina	0.5	0.815	0.236	97.22	0.6	0.4	0.767	0.903

*Modeling software which is used to produce the initial conformation of inhibitors as input files in SOMFA.

Resolution, utilized grid spacing in SOMFA software; r^2 , Non cross-validated correlation coefficient; s, standard error of estimate; F, F-test value; c_1 and c_2 , mixing coefficient of SOMFA model; q^2 , Cross-validated correlation coefficient; r, correlation coefficient. A model with high r^2 , q^2 and F-test values was selected as the best model. It should be noted that a high q^2 is only a necessary but not sufficient condition for a good prediction.

Internal validation phases

In this phase, SYR127063, which is the reference ligand of the crystal structure with PDB code 3PPO, was docked by AutoDock 4 and AutoDock Vina, separately (Aertgeerts et al., 2011). Superimposing the experimental and predicted conformations was expressed as root mean square deviation (RMSD). Calculated RMSD for SYR127063 were 3.56 and 0.65 Å for AutoDock4 and AutoDock Vina, respectively. The results showed that the molecular docking method by AutoDock Vina is robust and suitable for estimating the interactions of such ligand with kinase domain of HER2. Conversely, computed RMSD of 3.56 is accepted as poor docking method. Figure 4 shows the similarity between experimental and docked structures by two docking methods.

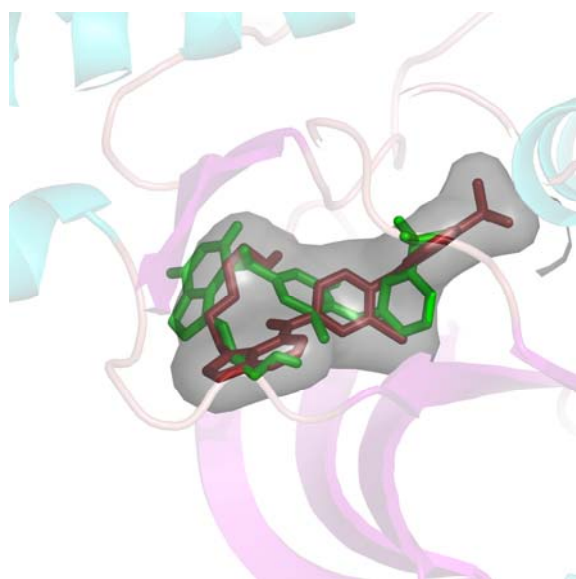


Figure 4a

Figure 4: Cartoon view of the active pocket of kinase domain of HER2 with its ligand

(a) Superposition of the docked structure and the crystal structure of SYR127063 which is calculated by AutoDock4

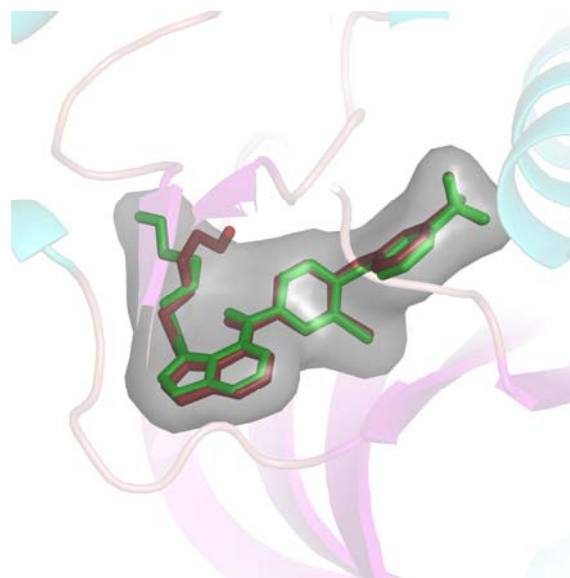


Figure 4b

(b) Superposition of the docked structure and the crystal structure of SYR127063 which is computed by AutoDock Vina. The crystal structure and docked conformation of SYR127063 are shown by red and green colors, respectively. For ease of visualization, hydrogen atoms are neglected.

Models produced by 3D-QSAR

In this section, as mentioned prior, six models were produced based on initial coordination of inhibitors. For predicted models, cross-validated r^2 or q^2 and also F-values were computed and served as a quantitative measure of the predictability of the SOMFA models. Statistical analyses of models are summarized in Table 2. Poor models are belonging to model A1 and A2 with the minimum values of q^2 and high quantity of standard error. As shown in Figure 4, the RMSD between docked and crystal structure is high, and it's not acceptable as a successful docking in molecular modeling studies. So, it can be postulated that the conformation of models, which are calculated by AutoDock4 have lower validity. When the three-dimensional structure of macromolecule target is not determined, or if it has not an acceptable resolution, or also when a critical amino acid of target active site is missed, use of semi-empirical methods to calculate the initial coordination in the 3D-QSAR study is common. Some studies apply this method

instead of molecular docking (Rajwade and Pande, 2008; Aggarwal et al., 2010). Model B1 and B2 have higher values of q^2 . The minimization of structures of these models was done by HyperChem. Among all of models, the best one which is predicted by using of AutoDock Vina and SOMFA software was C2. The mentioned model which has q^2 value of 0.76, r^2 value of 0.8, minimum standard error of 0.23 and F-test value of 97.22, demonstrates a satisfying statistical correlation and predictive ability. Based on our results, the models with the initial structures docked by AutoDock Vina are the best and those docked with AutoDock4 are poor and models B1 and B2 are moderate in view of validity. In molecular docking software, three parameters determine the success of docking study. The first is ligand and macromolecule representation, which is same in both AutoDock4 and Vina. Second is the scoring function of docking software. AutoDock4 uses five terms, including van der Waals, electrostatic, hydrogen bond, torsional penalty and desolvation parameter as its scoring function (Morris et al., 2009; Chang et al., 2010). Instead, AutoDock Vina employs just three score function terms comprising hydrophobic interaction (van der Waals), hydrogen bond and torsional penalty (Trott and Olson, 2010). Although, the above-mentioned docking packages use the same empirically-weighted scoring functions, but they have different empirical parameters and thereby, have a distinctive calculation. The third item that affects docking validity is the search algorithm. With regard to Figure 4 and Table 2, it seems that stochastic search algorithm of AutoDock4 has a poor function and efficiency compared to Vina. AutoDock Vina optimizes its local search application by calculation of derivatives of any ligand conformation energy, which is known as the gradient-based local search algorithm. Such method could increase speed and accuracy in AutoDock Vina. Trott and Olson (2010) showed that in redocking of 190 protein-ligand complexes with RMSD tolerance of 2 Å, success per-

centage of AutoDock4 and Vina was 49 % and 78 %, respectively (Trott and Olson, 2010). Statistical analysis of our data and also the calculated RMSD of SYR127063 in validation phase confirm such conclusion. In the course of SOMFA study, grid spacings of 1 and 0.5 Å were examined. Our data showed that in comparison with 1 Å, the 0.5 Å grid spacing generates a model with a good correlation (Table 2). For the other models, it seems that 1 Å grid spacing is suitable. It can be postulated if the structures with wrong or far from native conformation are used for grid-based 3D-QSAR, with decreasing of grid spacings, the amount of noise in the descriptor data is increased. Therefore, the quality of the model is reduced. In model C2, which has the best correlation with experimental data, the biological and predicted activities of the training and test sets are presented in Table 3. Figure 5 is depicted based on Table 3 and shows a good correlation ($r^2= 0.81$) between biological and predicted activity of HER2 inhibitors. Figure 6 shows the plot of the predicted versus biological activity values of test set, demonstrating that the most compounds are predicted satisfactorily.

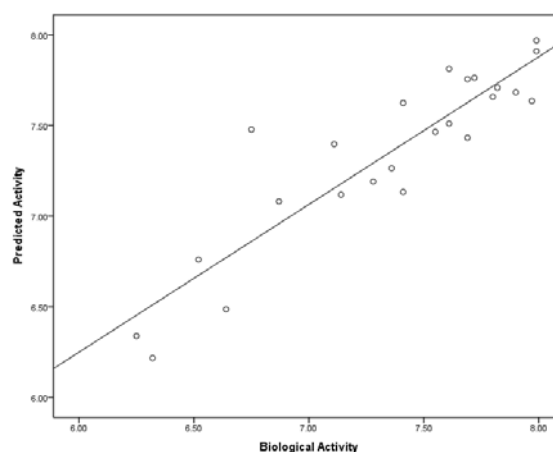


Figure 5: Predicted versus biological activity of data set from the best predictive SOMFA model

Table 3: Biological and predicted activities of training and test sets from the Model C2

Compound	Biological activity (pIC ₅₀)*	Predicted activity	Residual activity
1	7.970	7.636	0.334
2	7.900	7.684	0.216
3	7.820	7.709	0.111
4	7.410	7.133	0.277
5	7.690	7.756	-0.066
6	7.410	7.625	-0.215
7	7.720	7.764	-0.044
8	7.800	7.659	0.141
9	7.110	7.397	-0.287
10	7.990	7.970	0.020
11 ^T	7.360	7.264	0.096
12 ^T	7.610	7.511	0.099
13 ^T	7.690	7.433	0.257
14 ^T	7.610	7.814	-0.204
15	6.320	6.216	0.104
16	6.640	6.486	0.154
17	6.520	6.760	-0.240
18	7.550	7.465	0.085
19	6.750	7.478	-0.728
20	6.250	6.338	-0.088
21	7.140	7.119	0.021
22	7.280	7.191	0.089
erlotinib ^T	6.870	7.081	-0.211
lapatinib	7.990	7.910	0.080

T: test set molecules

*negative logarithm of IC₅₀, which is measured by HTScan® HER2/ErbB2 Kinase Assay Kit (Cai et al., 2010).

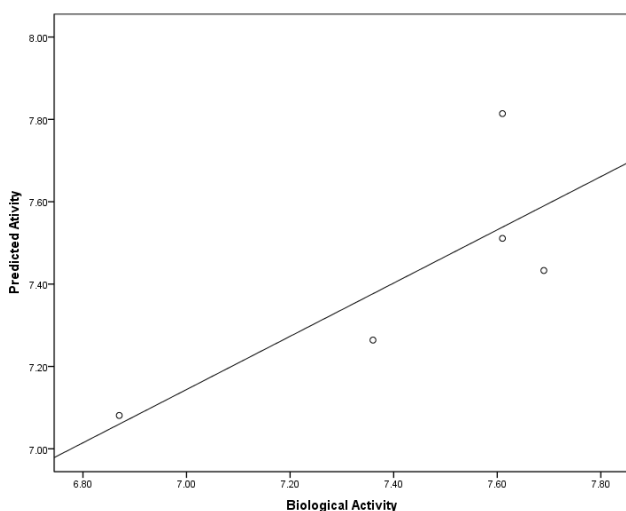


Figure 6: Predicted versus biological activity of test set from the best predictive SOMFA model

Electrostatic and shape contribution in SOMFA model

SOMFA produces a model based on shape and electrostatic independent variables. Shape values are given a value of 1 inside the van der Waals envelope and 0 for outside. Electrostatic values calculated by use of the partial charges distributed across the atom centers (Robinson et al., 1999) the SOMFA models. The models that have been produced using AutoDock 4 or Vina, electrostatic contribution was decreased to 0, by use of Gasteiger charge (data not shown). However, by assigning the AM1 charge as a partial atom charge, this proportion was raised. The same deduction was presented by Li et al. (2011). It has been shown that the partial charge computation method can affect the validity of 3D-QSAR models in a statistically significant mode (Mittal et al., 2009). AM1 charges are a set of Mulliken-type charges derived from a semi-empirical quantum-mechanical calculation (Mittal et al., 2009). Semi-empirical quantum methods are a mediocre between the empirical and ab initio quantum chemical approaches in terms of accuracy and computational time. Instead, Gasteiger charge calculation is an empirical atomic partial charge method which utilizes experimental parameters of its data set (Gasteiger and Marsili, 1980). One of the reasons for superiority of AM1 than Gasteiger atomic charge method is that AM1 embeds some quantum mechanic equations, which help to calculate a more accurate atomic charge than Gasteiger procedure. Mittal et al. (2009) concluded that semi-empirical charges give better predictive certain 3D-QSAR models than using Gasteiger partial charge calculation methods. The second reason is that in molecular docking methods, which are based on AutoDock4 or Vina, just polar hydrogen atoms are assigned. So, partial atomic charges were not calculated for non-polar hydrogen atoms. It might affect the final SOMFA model and cause a negative deviation in computation of electrostatic contribution. The both electrostatic and shape potentials of the model

C2 were computed. The contribution of electrostatic field and shape field to QSAR equation is 40 % and 70 %, respectively. Our SOMFA analysis result indicates that the shape contribution is of higher importance than the electrostatic one.

SOMFA model and HER2 active site analysis

The SOMFA shape potential for the analysis is presented as master grid in Figure 7. In this figure, structure of compound 1 is shown as reference compound. Red spots indicate that steric bulk enhances activity in this region. Contrariwise, cyan spots show that steric bulk detracts from activity in this region. Favorable and unfavorable electrostatic regions are depicted in Figure 8. In this figure, red spots indicate that positive charge is favored in this region. Alternatively, negative charge is disfavored in this region. Cyan spots demonstrate that negative charge is favored in this region. In other words, positive charge is disfavored in this region. Based on Figure 7, ring substitution in the position R₁ and R₂ in compounds 1-10 can moderately affect the HER2 inhibitory potency. Compound 1 has two bulky chlorine and fluorine on its phenyl ring, which could make a hydrophobic interaction with Val 734, Thr 798, Ala 751, Leu 852 and Leu 726 (Figure 9). Since, water ID 22 acts as a portion of the active site of HER2 and it's essential for catalysis (Aertgeerts et al., 2011), it was assigned in molecular docking studies. Based on Figure 9, hydrogen bond is formed between compound 1 and water ID 22. The most potent compound (compound 10) has the minimum IC₅₀. The bulky fluorobenzene moiety could interact with Leu 785, Thr 798, Val 734 and Leu 852 via hydrophobic interaction (Figure 10). These amino acids form a hydrophobic pocket that fluorobenzene moiety could lie on it. This hydrophobic moiety decreases the IC₅₀, which is in accordance with our shape SOMFA model (Figure 7). Based on Figure 7, number of carbons in the hydrocarbon tail might affect the HER2 kinase inhibitory

potency. Although, compound 1 with n=1 (number of carbon in the hydroxamic tail) has an IC₅₀ equal to 10.6, but it seems that optimal n is 5 (Cai et al., 2010). Hydrocarbon tail (n=5) of hydroxamic acid lets this moiety to form a hydrogen bond with Arg 849 and Asp 863. NH of Met 801, which is an important amino acid in HER2 active site (Aertgeerts et al., 2011), makes a hydrogen bond with one of the adenine moiety nitrogens. The remaining nitrogen also forms a hydrogen bond with water 22. With regard to Figure 7, C-7 substitution (R3) with long chain at quinazoline ring can reduce the inhibitory activities. High IC₅₀ of compounds 21 and 22 (478.2 for compound 21 and 225.2 for compound 22) correlate with our results (Cai et al., 2010). Figure 11 shows the 2-dimensional position of compound 22 in HER2 kinase active site. In spite of the fact that compound 22 contributes in several hydrogen bonds, polar Lys 753 with the positive charges in position with hydrocarbon tail might decrease the HER2 inhibitory potency. Furthermore, because of the long chain in compounds 21 and 22, conformational change of chlorofluorophenyl moiety could not let this group interact with any non-polar residues via hydrophobic interaction. Our data shows that hydrophobic interaction of the moiety mentioned above could influentially increase the HER2 inhibitory potency. Table 1 shows that addition of a bulky group in position R1, R2 and Y simultaneously leads to decrease in HER2 inhibitory activity. Compound 20 has chlorine in position R1 and ethyl in position Y. IC₅₀ of this compound is 556.9. With exchange of chlorine with fluorine, IC₅₀ was sharply decreased (compound 19). Compound 15 possesses CH₂ in position Y and fluorine and chlorine in positions R1 and R2, respectively. IC₅₀ of this compound is slightly lesser than that of compound 20. Based on our SOMFA model (Figure 7), phenyl moiety of compounds 15, 16 and 20 are located on cyan spots area. These results have a good agreement with experimentally obtained IC₅₀ values (Cai et al., 2010).

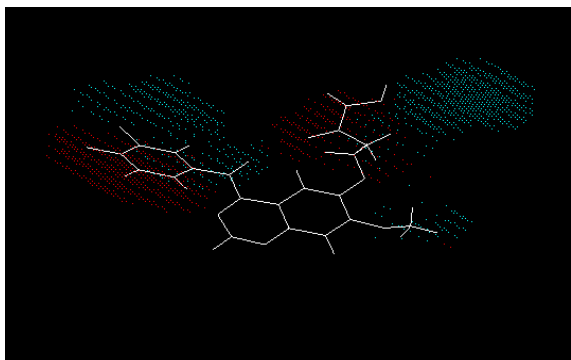


Figure 7: The shape potential master grid of model C2 with compound 1. Red spots portray areas of desirable steric interactions. Blue spots represent areas of undesirable steric interactions.

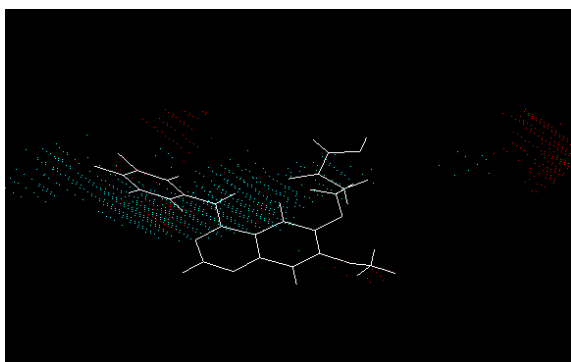


Figure 8: The electrostatic potential master grid of model C2 with compound 1. Red portrays areas where positive potential is favorable, or negative charge is unfavorable. Blue demonstrates areas where negative potential is desirable, or positive charge is undesirable.

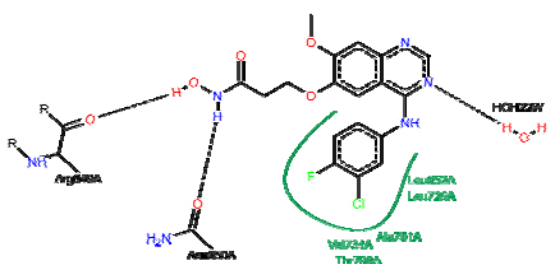


Figure 9: Schematic two-dimensional representations of the binding interactions between compound 1 and active site of HER2 kinase domain. Black dashed lines are hydrogen bonds. Hydrophobic interactions are shown by solid line green. HOH22W represents water molecule.

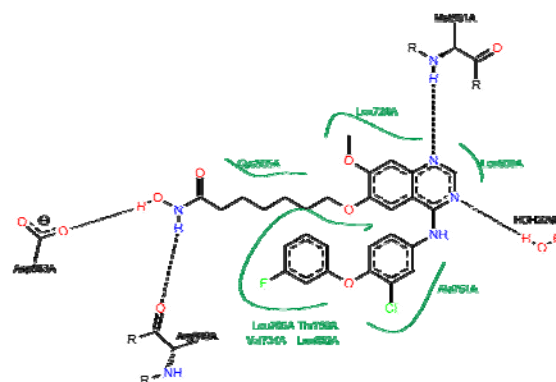


Figure 10: Schematic two-dimensional representations of the binding interactions between compound 10 and active site of HER2 kinase domain. Black dashed lines are hydrogen bonds. Hydrophobic interactions are shown by solid line green. HOH22W represents water molecule.

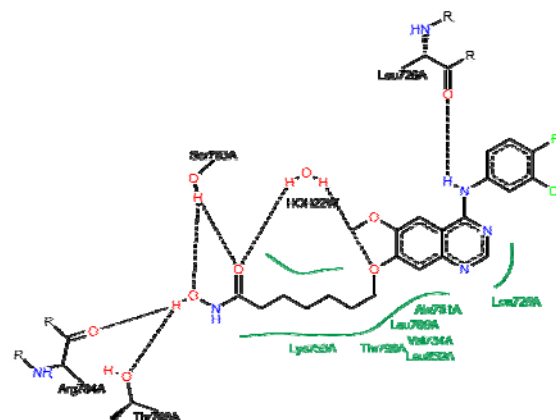


Figure 11: Schematic two-dimensional representations of the binding interactions between compound 22 and active site of HER2 kinase domain. Black dashed lines are hydrogen bonds. Hydrophobic interactions are shown by solid line green. HOH22W represents water molecule.

The authors suggest that based on SOMFA analysis presented in this study, more potent inhibitors of kinase domain of HER2 could be designed and synthesized.

CONCLUSION

In the last decade, structure-based methods have become great tools in drug design, including lead discovery and optimization. It has also been shown that structure-based methods are now able to predict, with an acceptable degree of exactness, the position of a ligand in its binding site.

Combining the molecular docking and 3D-QSAR studies could be employed to design the more potent HER2 chemical inhibitors. By the aid of molecular docking, the derived 3D-QSAR models are also able to indicate which interaction sites in the binding pocket might be responsible for the variance in biological activities. Active site analysis of kinase domain of HER2 also lets us interpret and validate our QSAR model. In this study, calculated SOMFA 3D-QSAR models for quinazoline derivatives were expanded. Because of the nature of HER2 active site, which is covered by hydrophobic residues, shape contribution in our SOMFA model is a dominant factor. The number of hydrophobic interactions and hydrogen bonding between chemical inhibitors and HER2 active site could increase the inhibitory potency. These features provide some useful information that might be used in new quinazoline derivatives with a high and influential HER2 inhibitory activity.

Conflict of interest statement

The authors declare that they have no conflict of interest.

REFERENCES

- Aertgeerts K, Skene R, Yano J, Sang BC, Zou H, Snell G et al. Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. *J Biol Chem* 2011;286:18756-65.
- Aggarwal S, Thareja S, Bhardwaj TR, Kumar M. Self-organizing molecular field analysis on pregnane derivatives as human steroidal 5 α -reductase inhibitors. *Steroids* 2010;75:411-8.
- Akamatsu M. Current state and perspectives of 3D-QSAR. *Curr Topics Medicinal Chem* 2002;2:1381-94.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H et al. The Protein Data Bank. *Nucl Acids Res* 2000;28:235-42.
- Brooks BR, Brooks CL 3rd, Mackerell AD Jr., Nilsson L, Petrella RJ, Roux B et al. CHARMM: the biomolecular simulation program. *J Comput Chem* 2009;30:1545-614.
- Burgess AW, Cho HS, Eigenbrot C, Ferguson KM, Garrett TP, Leahy DJ et al. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol Cell* 2003;12:541-52.
- Cai X, Zhai HX, Wang J, Forrester J, Qu H, Yin L et al. Discovery of 7-(4-(3-ethynylphenylamino)-7-methoxyquinazolin-6-yloxy)-N-hydroxyheptanamide (CUDc-101) as a potent multi-acting HDAC, EGFR, and HER2 inhibitor for the treatment of cancer. *J Medicinal Chem* 2010;53:2000-9.
- Carter P, Presta L, Gorman CM, Ridgway JB, Henner D, Wong WL et al. Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc Natl Acad Sci USA* 1992;89:4285-9.
- Chang MW, Ayeni C, Breuer S, Torbett BE. Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina. *PLoS One* 2010;5(8):e11955.
- Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW Jr. et al. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* 2003;421(6924):756-60.
- de Bono JS, Rowinsky EK. The ErbB receptor family: a therapeutic target for cancer. *Trends Mol Med* 2002;8(4 Suppl):S19-26.
- Froimowitz M. HyperChem: a software package for computational chemistry and molecular modeling. *BioTechniques* 1993;14:1010-3.
- Galvao ER, Martins LM, Ibiapina JO, Andrade HM, Monte SJ. Breast cancer proteomics: a review for clinicians. *J Cancer Res Clin Oncol* 2011;137:915-25.

- Garrett TP, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO et al. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* 2002;110:763-73.
- Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. *Tetrahedron* 1980;36:3219-28.
- Golbraikh A, Tropsha A. Beware of q²! *J Mol Graph Modelling* 2002;20:269-76.
- Herbst RS. Review of epidermal growth factor receptor biology. *Int J Radiat Oncol Biol Phys* 2004;59(2 Suppl):21-6.
- Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA: a cancer journal for clinicians* 2010;60:277-300.
- Kulkarni SS, Gediya LK, Kulkarni VM. Three-dimensional quantitative structure activity relationships (3-D-QSAR) of anti-hyperglycemic agents. *Bioorg Medicinal Chem* 1999;7:1475-85.
- Li Z, Zhou M, Wu F, Li R, Ding Z. Self-organizing molecular field analysis on human beta-secretase nonpeptide inhibitors: 5, 5-disubstituted aminohydantoins. *Eur J Medicinal Chem* 2011;46:58-64.
- Linggi B, Carpenter G. ErbB receptors: new insights on mechanisms and biology. *Trends Cell Biol* 2006;16:649-56.
- Mittal RR, Harris L, McKinnon RA, Sorich MJ. Partial charge calculation method affects CoMFA QSAR prediction accuracy. *J Chem Inf Model* 2009;49:704-9.
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 2009;30:2785-91.
- Pawson T. Protein modules and signalling networks. *Nature* 1995;373(6515):573-80.
- Rajwade RP, Pande R. Evaluation of lipophilicity of N-arylhydroxamic acids by reversed phase-high performance liquid chromatographic method and self-organizing molecular field analysis. *Anal Chim Acta* 2008;630:205-10.
- Robinson DD, Winn PJ, Lyne PD, Richards WG. Self-organizing molecular field analysis: a tool for structure-activity studies. *J Medicinal Chem* 1999;42:573-83.
- Satyanarayanajois S, Villalba S, Jianchao L, Lin GM. Design, synthesis, and docking studies of peptidomimetics based on HER2-herceptin binding site with potential anti-proliferative activity against breast cancer cell lines. *Chem Biol Drug Design* 2009;74:246-57.
- Schuttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D: Biol Crystallogr* 2004;60:1355-63.
- Shmeeda H, Tzemach D, Mak L, Gabizon A. Her2-targeted pegylated liposomal doxorubicin: retention of target-specific binding and cytotoxicity after in vivo passage. *J Contr Release: Off J Contr Release Soc* 2009;136:155-60.
- Stebbing J, Copson E, O'Reilly S. Herceptin (trastuzumab) in advanced breast cancer. *Cancer Treat Rev* 2000;26:287-90.
- Traxler P, Furet P, Mett H, Buchdunger E, Meyer T, Lydon N. Design and synthesis of novel tyrosine kinase inhibitors using a pharmacophore model of the ATP-binding site of the EGF-R. *J Pharm Belg* 1997;52:88-96.
- Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;31:455-61.

Woodburn JR. The epidermal growth factor receptor and its inhibition in cancer therapy. *Pharmacol Ther* 1999;82):241-50.

Yarden Y, Baselga J, Miles D. Molecular approach to breast cancer treatment. *Sem Oncol* 2004;31(Suppl 10):6-13.

Zeglinski M, Ludke A, Jassal DS, Singal PK. Trastuzumab-induced cardiac dysfunction: A 'dual-hit'. *Exp Clin Cardiol* 2011;16:70-4.