Molecular docking, MM/GBSA and 3D-QSAR studies on EGFR inhibitors

RAJU BATHINI, SREE KANTH SIVAN, SABIHA FATIMA and VIJJULATHA MANGA*

Molecular Modeling and Medicinal Chemistry Group, Department of Chemistry, University College of Science, Osmania University, Hyderabad 500 007, India
e-mail: vijjulathamanga@gmail.com

Abstract. Epidermal growth factor receptor (EGFR) is the first growth factor receptor proposed as a target for cancer therapy. Molecular modeling protocols like molecular docking, molecular mechanics/generalized born surface area (MM/GBSA) calculations and three dimensional-quantitative structure activity relationship (3D-QSAR) studies were performed on 45 molecules to understand the structural requirements for EGFR tyrosine kinase inhibitors. Conformation for all the molecules obtained from molecular docking were used as is for 3D-QSAR analysis. Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) models were obtained by performing partial least square analysis on 35 training molecules and these models were validated using 10 test molecules. The models showed good statistical results in terms of r^2, q^2_{loo} and r^2_{pred} values. Information rendered from 3D-QSAR model and sitemap analysis was used to optimize lead molecule to design prospective inhibitors. Improvement in EGFR binding affinity can be achieved by substitutional modification on phenyl ring attached to alkynyl group with bulkier hydrogen bond donor and acceptor substituents that can increase favourable interaction with the receptor.

Keywords. Epidermal growth factor receptor (EGFR); Extra precision (XP) docking; 3D-QSAR; Comparative molecular field analysis (CoMFA); Comparative molecular similarity indices analysis (CoMSIA); Molecular mechanics; generalized born surface area (GBSA).

1. Introduction

Epidermal growth factor receptor (EGFR) is a transmembrane glycoprotein consisting of a single polypeptide chain of 1186 amino acids.[1] It belongs to a family of receptor tyrosine kinases (TKs) that includes EGFR, ErbB-2, ErbB-3 and ErbB-4 (also known as HER-1, HER-2, HER-3, and HER-4).[2] The EGFR and its family members are composed of an extracellular ligand-binding domain, a transmembrane domain, and an intracellular domain. When epidermal growth factor (EGF), binds to the ligand-binding domain, the EGF receptor forms a homodimer or heterodimer with other members of EGFR family. This activates the intrinsic kinase domain that in turn triggers autophosphorylation events on specific tyrosine residues within the cytoplasmic tail activity.[3,4] Downstream single proteins then initiate several signal transduction cascades, leading to DNA synthesis, and cell proliferation.[5] EGFR is a crucial regulator in several cell phenotypes, such as cell migration, adhesion, proliferation, and immune response in the human skin making it a key target for anti-tumor strategy.

Early studies on EGFR inhibition have reported many compounds having inhibitory activity, these included 4-anilinoquinazoline derivatives,[6–8] arylaminopyrimidines[9,10] and pyrrole triazines derivatives.[11,12] Lapatinib, which is a potent dual EGFR/ErbB2 inhibitor, approved by FDA for the treatment of breast cancer (figure 1), is a modified 4-Anilinoquinazoline derivative.[13] Recently, pyrimidine derivatives like anilinopyrimidines, thienopyrimidines, and alkenyl-methylpyrimidine were reported as EGFR inhibitors.[14–18] All these skeletons possess a pyrimidine ring portion that bind with EGFR. The crystal structures of EGFR bound to inhibitors have two hydrogen bond interactions specifically with the active site residues Met793 and Asp800. Met793 played a key role in the inhibitory activity of EGFR. In an ongoing effort, we report receptor based 3D-QSAR studies using comparative molecular field analysis (CoMFA)[19,20] and comparative molecular similarity indices analysis (CoMSIA)[21] methodology on EGFR inhibitors. Partial least square (PLS)[22] based statistical analysis was carried out on 45 molecules to identify the correlation. Contour maps generated enabled us to explain the observed variation in activity and guided us to design new molecules.
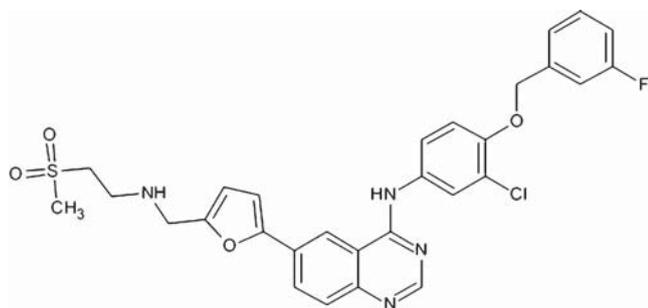
*For correspondence

Figure 1. Lapatinib – Dual EGFR/ErbB-2 inhibitor.

2. Methodology

2.1 *Ligand preparation*

All molecular modeling calculations were performed on a Linux operating system. A total of 45[14–18] molecules were available with reported IC_{50} values for inhibition of EGFR tyrosine kinase. The kinase inhibition assay IC_{50} determinations were measured with the homogeneous time resolved fluorescence (HTRF) KinEASE-TK assay. Three dimensional structures of these inhibitors were sketched using Mastreo build panel in Schrödinger 2010 suite, and the IC_{50} values were converted to pIC_{50} values ($pIC_{50} = -\log IC_{50}$) (table 1).

The ligprep module of Schrodinger was used to generate all possible states at a physiological pH range 7 ± 2 and produces lowest potential energy conformer of the ligand using an OPLS_2005 force field. All these ligands post preparation was used for molecular docking studies and ADME properties using Qikprop module of Schrodinger Suite.

2.2 *Protein preparation*

Crystal structure of EGFR tyrosine kinase bound with lapatinib with resolution of 2.4 Å was downloaded from RCSB protein data bank.(http://www.rcsb.org/pdb/) (PDB id 1xkk). Protein was prepared using *protein preparation wizard module of* Schrodinger Suite. All the water molecules and heteroatoms except lapatinib (FMM) were deleted and hydrogen atoms were added to the crystal structure. The complex was submitted to minimization to relieve steric clashes using the OPLS 2005 force field. Sitemap module was used to characterise features of binding sites. Hydrophobic and hydrophilic maps are generated which are dispersed as donor, acceptor, and hydrophobic regions.

2.3 *Molecular docking*

Active site for docking ligands was defined using Receptor Grid Generation module in Glide 5.6.[23]

Co-crystallied ligand was selected and grid was generated around the active site of EGFR intrinsic domain, with a receptor van der Waals scaling for non-polar atoms as 0.9.[24,25] Molecules selected from literature (figure 2) were prepared by LigPrep application. These molecules were docked into the grid generated from EGFR protein structure. Two different docking protocols, standard precision (SP) and extra precision (XP) were employed to analyze binding modes of the known inhibitors from the literature.

2.4 *Prime MM/GBSA calculations*

To estimate relative binding affinities of protein ligand complexes, Molecular mechanics with generalized born surface area (MM/GBSA) is employed. The relative binding free energy ΔG_{bind} was estimated according to the following equation:

$$\Delta G_{bind} = E_{complex}(minimized) \\ -[E_{ligand}(unbound, minimized) \\ +E_{receptor}(unbound, minimized)]$$

Where ΔG_{bind} is the calculated relative free energy that includes both ligand and receptor strain energy. $E_{complex}$(minimized) is the MM/GBSA energy of the minimized complex, and E_{ligand} (unbound, minimized) is the MM/GBSA energy of the ligand after removing it from the complex and allowing it to relax. $E_{receptor}$ (unbound, minimized) is the MM/GBSA energy of protein after separating it from the ligand.[26,27] Prime MM/GMSA modules was used for calculating the relative binding free energy for each molecule. The protein ligand complexes obtained from XP docking were subjected to MM/GBSA calculations.

2.5 *CoMFA and CoMSIA studies*

CoMFA and CoMSIA studies were carried out as described by Carmer *et al.,*[19,20] and Fatima *et al.*[28] The process is briefly described here. Dock pose of each ligand that showed two hydrogen bond interactions with the receptor were imported into SYBYLX-2.1 molecular modeling program package[29] and Gasteiger–Huckel[30] charges were assigned. Standard Tripos force fields were employed for CoMFA and CoMSIA analysis. A 3D cubic lattice of dimension 4 Å in each direction with each lattice intersection of a regularly spaced grid of 2.0 Å was created. Steric and electrostatic parameters were calculated in CoMFA fields while hydrophobic, acceptor and donor parameters in addition to steric and electrostatic were calculated in CoMSIA fields at each lattice. The sp^3 carbon atom was used as a

Table 1. Experimental and predicted activity of EGFR tyrosine kinase inhibitors along with Dock score (Kcal/mol) and ΔG_{bind}(Kcal/mol).

Mol	$IC_{50(\mu M)}$	pIC_{50}	Predicted pIC_{50} CoMFA	Predicted pIC_{50} CoMSIA	SP Dock Score (Kcal/mol)	XP Dock Score (Kcal/mol)	ΔG_{bind} (Kcal/mol)
1	0.650	6.190	6.050	6.120	−11.260	−11.326	−121.746
2	0.022	7.660	8.040	7.910	−12.700	−12.128	−132.827
3	0.003	8.520	7.490	8.240	−11.740	−12.197	−133.303
4	0.009	8.050	7.720	7.820	−11.500	−11.572	−135.183
5	0.190	6.720	7.000	7.470	−11.940	−12.147	−125.868
6	0.079	7.100	7.370	7.380	−11.700	−11.527	−140.538
7*	0.030	7.520	7.416	7.648	−11.440	−11.427	−143.336
8	0.055	7.260	7.120	7.220	−12.130	−12.309	−128.936
9*	0.094	7.030	6.750	6.970	−11.020	−12.369	−136.369
10*	0.920	6.040	5.920	5.700	−9.620	−9.702	−90.464
11*	1.100	5.960	5.789	5.903	−9.930	−10.085	−100.213
12	47.000	4.330	4.670	4.490	−9.520	−9.527	−96.164
13	22.000	4.660	5.110	5.000	−9.090	−8.096	−109.857
14	10.000	5.000	5.920	5.460	−10.310	−10.175	−108.267
15*	0.068	7.170	6.708	7.113	−10.810	−11.282	−125.437
16	0.065	7.190	7.190	7.230	−11.160	−11.902	−115.438
17	0.590	6.230	7.060	6.450	−11.460	−11.119	−117.256
18	0.034	7.470	7.190	7.140	−11.020	−11.571	−143.399
19	0.240	6.620	7.290	7.040	−11.700	−11.636	−135.572
20	0.028	7.550	7.320	7.480	−11.580	−12.491	−140.553
21*	0.018	7.740	6.971	6.808	−11.720	−12.462	−128.446
22	0.014	6.940	6.960	6.840	−11.430	−12.010	−132.853
23	0.003	8.520	8.110	8.280	−11.480	−11.976	−128.161
24	0.019	7.720	7.430	7.880	−11.070	−11.846	−130.904
25	0.156	6.810	7.210	7.160	−11.050	−12.239	−131.088
26	0.013	7.890	7.350	7.080	−10.930	−11.846	−127.077
27*	4.130	5.380	6.371	6.682	−11.210	−11.451	−128.013
28	0.038	7.420	7.270	7.370	−11.380	−11.690	−127.590
29	0.011	7.960	7.740	7.360	−10.620	−11.284	−139.209
30*	0.009	8.040	8.279	8.235	−11.740	−11.722	−137.283
31	9.700	5.010	4.700	4.990	−9.540	−9.584	−108.649
32	5.600	5.250	5.050	5.480	−9.360	−8.639	−89.460
33	72.400	4.140	3.980	3.810	−9.020	−7.540	−84.807
34	3.500	5.460	4.980	5.710	−9.020	−8.215	−93.839
35	26.600	4.580	4.230	3.990	−10.800	−10.727	−100.470
36	8.600	5.070	5.410	5.410	−8.620	−7.755	−102.330
37	0.009	8.050	8.160	8.220	−13.310	−13.321	−137.448
38*	0.080	7.100	6.113	6.429	−11.410	−9.700	−138.488
39	0.180	6.740	6.720	6.530	−11.550	−12.447	−129..403
40	0.108	6.970	6.630	6.600	−10.800	−11.434	−115.789
41	0.129	6.890	7.010	6.840	−11.290	−11.592	−129.656
42*	0.048	7.320	6.626	6.551	−10.620	−11.936	−112.935
43	0.627	6.200	6.610	6.180	−11.430	−12.536	−124.355
44	0.821	6.090	6.670	6.190	−11.190	−12.043	−125.053
45	2.138	6.140	5.660	6.030	−9.830	−11.587	−107.014

* represents test set molecules.

probe atom to generate steric (Lennard–Jones potential) field energies and a charge of +1 to generate electrostatic (Coulombic potential) field energies. A distance dependent dielectric constant of 1.00 was used. Steric and electrostatic fields were condensed to +30.00 Kcal/mol.

PLS analysis was performed to correlate computed fields with biological activity ($pIC_{50} = -\log IC_{50}$)

for the data set of ligands. Molecules were randomly divided into training and test set of 35 and 10, respectively, as 3:1 ratio. The pIC_{50} difference between most active and least active molecule in the dataset is 4.38 log units. The training set was used to generate the QSAR model and test set was used to validate the generated model. In the PLS statistical analysis, biological activity values of ligands were used as dependent

variables. To improve the signal-to-noise ratio column filtering value (s) was set to 2.0 Kcal/mol by omitting lattice points whose energy variations were below the threshold. Leave-one-out (LOO) cross-validations were performed to determine the optimum number of components (ONC) and coefficient q^2_{loo}. The obtained ONC was applied to derive the final QSAR model. Validation of CoMFA and CoMSIA derived model was performed by predicting the activity of test set compounds.

3. Results and Discussion

The accuracy of molecular conformation for Quantitative Structure Activity Relationship (QSAR) studies
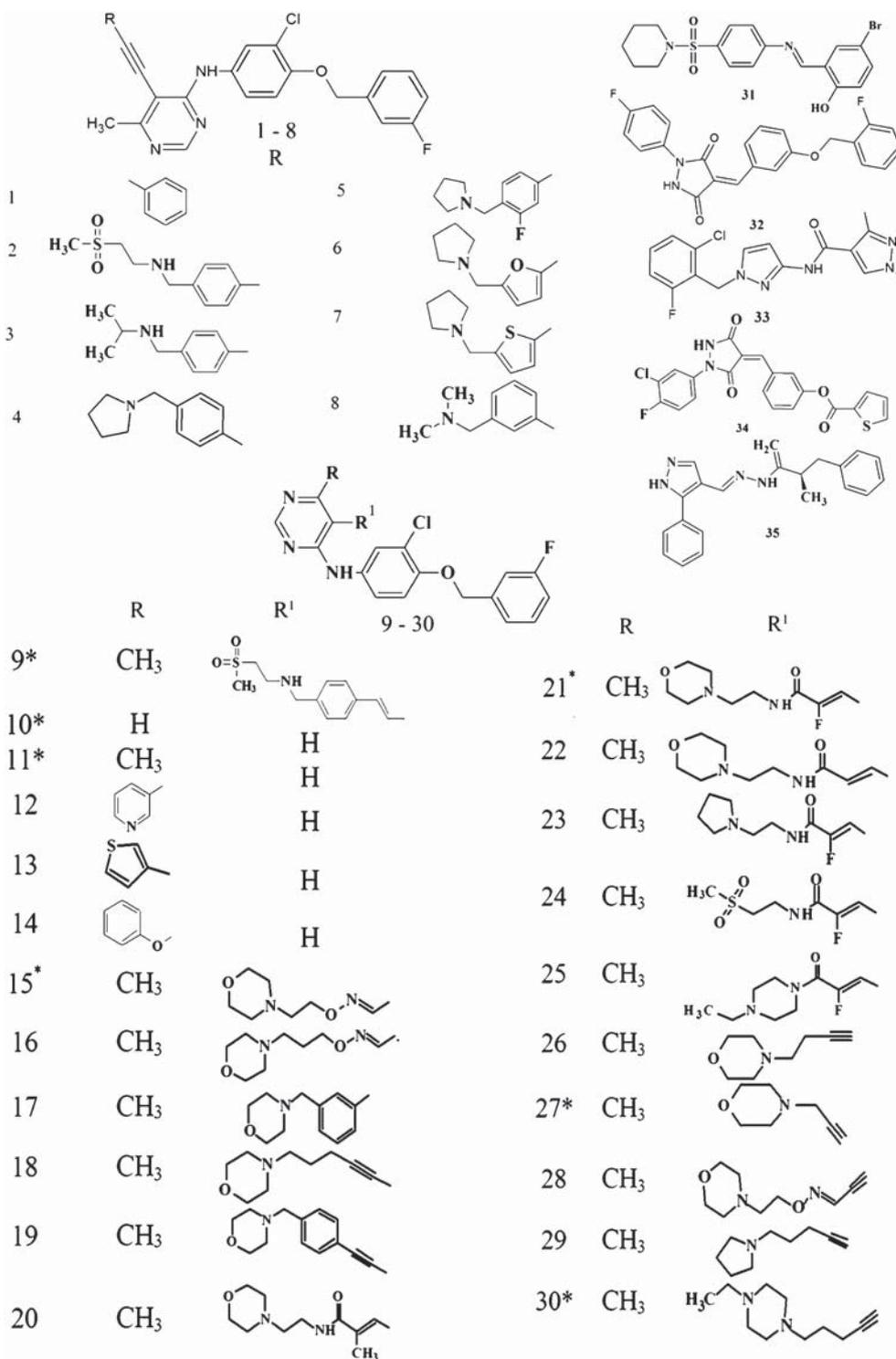
Figure 2. Structures of EGFR Inhibitors. *indicates test molecules.
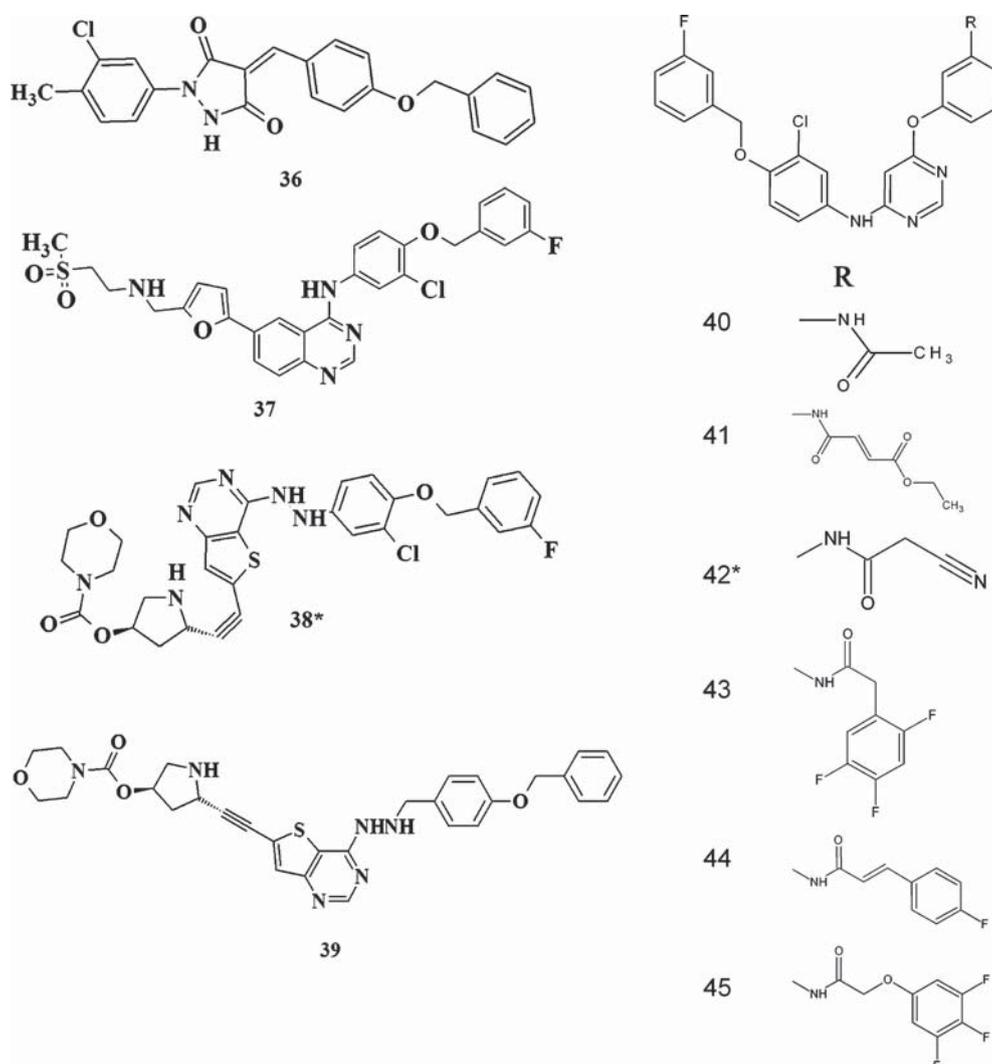
Figure 2. (*Continued*).

is a most important criteria, and the most favorable bioactive conformer selection is a tough task. Molecular docking of active compounds into the active site of the protein is one of the methods used for obtaining the active conformation. These molecules were docked into the active site (lapatinib binding site) of EGFR to obtain the receptor based conformations (figure 2). The RMSD between the predicted conformation and the observed X-ray Crystallographic conformation of the compound (Lapatinib) equaled 1.92 Å, a value that suggests the reliability of docking protocol in producing the experimentally observed binding mode for EGFR inhibitors. Dock poses analysis of all the molecules showed similar hydrogen bond interactions with active site residues. Molecules were having specific hydrogen bond interactions between N1 of the pyrimidine ring and amine in the tail region with hinge region amino acid Met793 and Asp800, respectively. These hydrogen bond interactions were also observed for the co-crystalized ligand (lapatinib), and it also showed a halogen interaction

with gate keeper amino acid Thr790. DFG binding motif is also a crucial part for inhibitor binding to EGFR, lapatinib and most of the pyrimidine derivative showed a hydrophobic ($\pi - \pi$) interaction with Phe856.

Molecular mechanics/generalized born surface area (MM/GBSA) calculations were performed for best ranking molecules in XP mode. MM/GBSA calculations were performed to estimate relative binding affinity of ligands to the receptor. In principle, MM/GBSA is used for free energy based ranking of ligands belonging to a congeneric series. MM/GBSA and docking protocols employed in the present work for assessing ligand affinities to EGFR allow protein flexibility and therefore, give more reliable results.

The list of inhibitors along with their pIC$_{50}$, dock scores from SP, XP and ΔG$_{bind}$ values from Prime have been tabulated in table 1. Dock score (XP dock score) vs. Experimental pIC$_{50}$ and ΔG$_{bind}$ vs. Experimental pIC$_{50}$ gave a correlation coefficient value (r) of 0.728 and 0.782, respectively, which shows a significant

relation between biological activity vs. docking and free energy calculations. The high correlation depicts the accuracy of the docking protocol. The correlation of MM/GBSA G_{bind} values and experimental values is in acceptable range of 78% (r = 0.78 and r^2 = 0.61). Even though a few molecules showed higher MM/GBSA values than best active molecule 3 and 23. The variation in binding affinity with respect to experimental activity can be correlated to lesser aqueous solubility of these molecules (ADME properties in supplementary data). Molecule 37 (lapatinib) is the co-crystallized ligand in the protein structure used for docking studies (pdb id 1xxk) hence its docking and binding values are higher compared to the best active molecule, other molecules had bulkier hydrophobic group (heterocyclic saturated groups like morpholine, pyrrolidine) that showed favourable interaction with hydrophobic amino acids residues (Leu 718, Cys 797, Leu 1001 and Met 1002) and also an electrostatic interaction with Asp 800. Scatter plot of pIC_{50} versus XP dock score and ΔG_{bind} values are shown in figure 3. Usually in 3D QSAR, all the molecules in the dataset are superimposed onto the best active molecule. Instead, the most promising poses (best GLIDE Score along with most physiologically similar positions) of each molecule has been considered for 3D QSAR analysis. Overlay of dock pose of each ligand is shown in figure 4.

CoMFA and CoMSIA analysis were carried out on these reported inhibitors. The molecules were divided into a training set consisting of 35 molecules for the derivation of QSAR model and a test set of 10 molecules for validation of derived model. CoMFA and CoMSIA statistical analysis are summarized in table 2. Statistical data shows q_{loo}^2 of 0.638 and 0.672, r^2 0.88 and 0.929, respectively, for CoMFA and CoMSIA. The q_{loo}^2 and r^2 values indicate a good internal predictive ability of this model. The SEE value of 0.448 and 0.345,

respectively, for CoMFA and CoMSIA indicate error in prediction of activity from the obtained model, as the model is derived from a training set of 35 molecules the
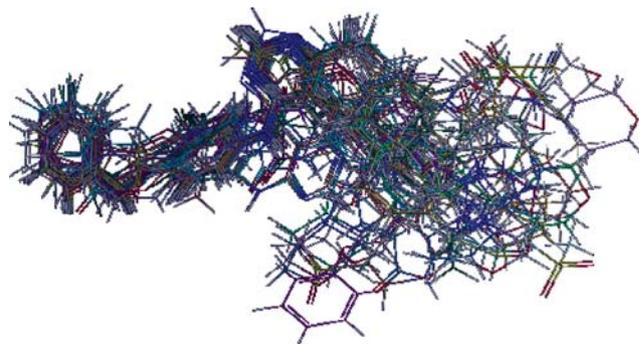

Figure 4. Dock pose alignment of EGFR inhibitors.

Table 2. PLS results in summary.

Statistical parameters	CoMFA	CoMSIA
Training set	35	35
Test set	10	10
q_{loo}^2	0.638	0.672
ONC	3	3
SEE	0.448	0.345
r^2	0.880	0.929
F_{ration}	75.645	134.713
r_{pred}^2	0.530	0.520
Fraction of field contributions		
Steric	0.619	0.124
Electrostatic	0.381	0.219
Hydrophobic	–	0.178
Donor	–	0.214
Acceptor	–	0.265

F_{ratio} = Fisher test value; q_{loo}^2 = cross-validated correlation coefficient by leave one out method; r^2 = conventional correlation coefficient; r_{pred}^2 = cross-validated correlation coefficient on test set; ONC, optimum number of components; SEE, standard error of estimate.
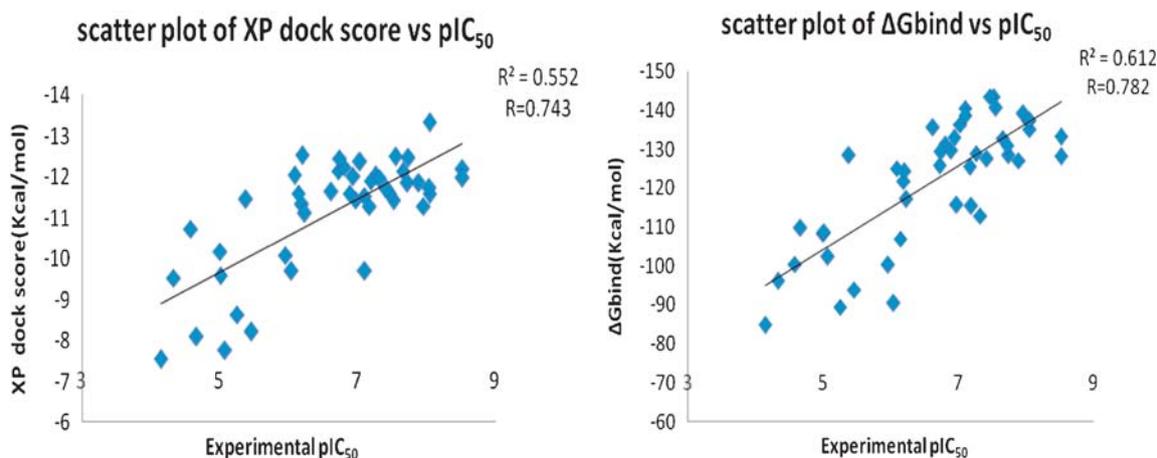

Figure 3. Scatter plot of pIC_{50} *vs*. XP dock score and ΔG_{bind} values.

value seem to be optimum with respect to the r^2 values obtained for the models, the error estimate would have been less for a larger set. Test set molecules excluded from the model generation were used to predictive ability of the model. Predictive correlation coefficient r^2_{pred} 0.53 for CoMFA and 0.52 for CoMSIA, respectively, wich indicates good external predictive ability of the model.

Scatter plots for actual and predicted pIC_{50} values for training and test sets of CoMFA and CoMSIA studies are shown in figure 5. Observed and predictive pIC_{50} values of these molecules are provided in table 1.

CoMFA and CoMSIA contour maps were generated to visualize information content of the 3D QSAR models. Modifications required to design new molecules are suggested by molecular fields (contour maps) that define favorable or unfavorable interaction energies of aligned molecules. Contour maps of CoMFA denote the region in the space where molecules would favorably or unfavorably interact with the receptor while CoMSIA contour maps symbolize specified regions where the

presence of a particular group with certain physico-chemical property binds to the receptor. CoMFA and CoMSIA results were graphically interpreted by field contribution maps using "STDEVCOEFF" field type. All contours represented default 80 and 20% level contribution for favored and unfavored regions.[19,20] Molecule 3, the most potent inhibitor within the series was incorporated on the map to aid visualization. Variation in activity is discussed based on the contour maps analyzes the characterstic features of binding sites.

Steric contour maps of CoMFA and CoMSIA models are similar as seen in figures 6a and 6b, respectively; the green and yellow areas represent favored and disfavored regions, respectively. Large green contour is seen near the substituted benzylamine ring, suggesting substitution with a bulky group at this position will increase the activity. The most active molecule (3) shows substituted benzylamine ring incorporated into the green region. A yellow contour is seen around the pyrimdine ring, suggesting more bulky substituents in this area will significantly decrease the biological activity. Figure 7
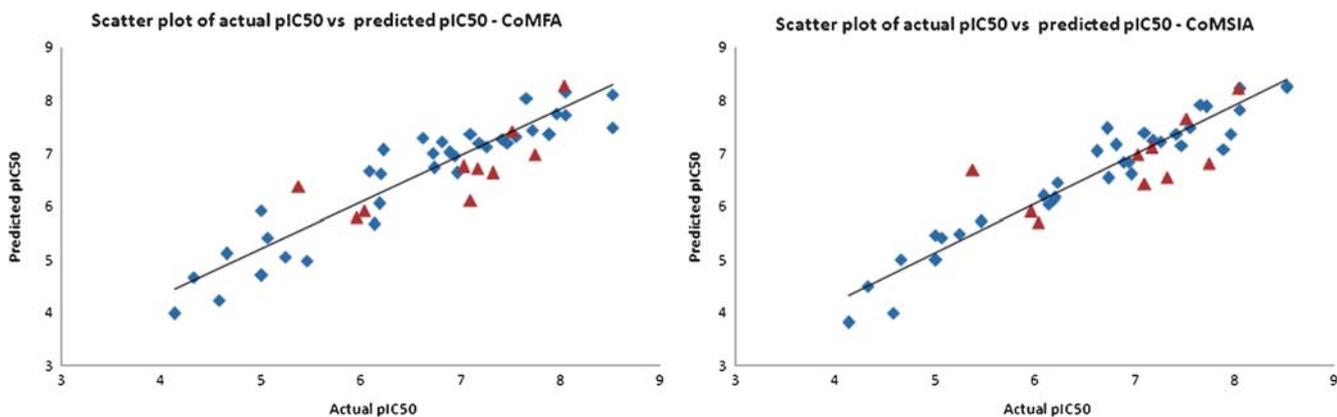

Figure 5. Scatter plot for the actual and predicted pIC_{50} values, the training set is represented in squares and test set in triangles.
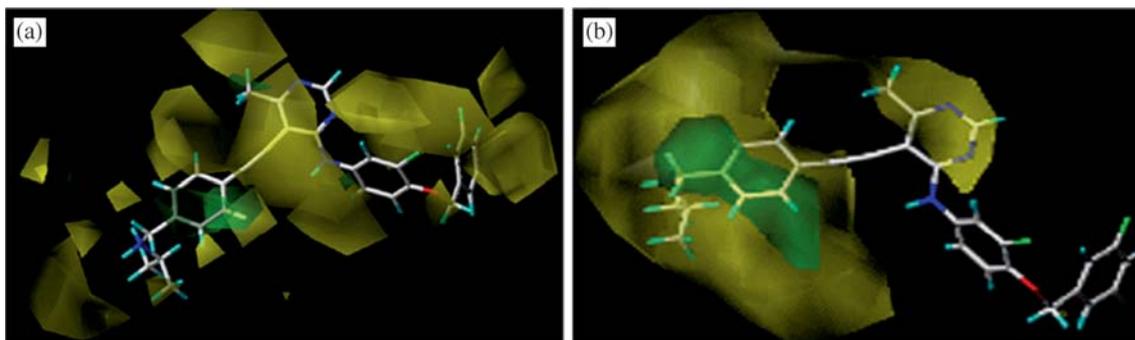

Figure 6. CoMFA and CoMSIA steric fields. (a) CoMFA and (b) CoMSIA standard deviation (SD* coefficient) contour map illustrating steric features in combination with molecule 3. Green contours represent favorable bulky group substitution at that point while yellow regions are disfavorable for activity.

shows electrostatic contour maps where the red contour represents electronegative charge favoring and blue contour represents the electropositive charge favoring, respectively; blue contour on the side chain of substituted benzylamine ring indicates lesser electronegative substituents may increase the activity. Red contour on benzyloxy phenyl ring suggests that high electronegative substituents like flourine at this position are essential for biological activity as in the case of the most active molecule 3.

Hydrophobic contour are embodied in figure 8a, yellow and white areas represent hydrophobic and hydrophilic preferred regions. Three small yellow contours are seen, one at phenyl ring, one at methyl substituent on pyrimidine ring and third on benzyloxy attached phenyl ring. Bulkier hydrophobic group substitution at this position may increase the activity of the molecules since this region is also steric favored. A white contour is observed at the phenyl ring

suggesting hydrophilic groups in this region will increase activity.

Acceptor and donor contour maps are shown in figure 8b, magenta and cyan fields indicate the hydrogen bond acceptor and donor favored regions whereas red and purple fields imply disfavored regions. In molecule 3, a donor disfavored purple contour is seen in the vicinity of the pyrimidine ring. Magenta contours are away from the molecule, and a small cyan contour is observed near the NH of the side chain phenyl ring. Site map analysis of the receptor showed similar hydrophobic, hydrogen bond donor and hydrogen bond acceptor regions as obtained from the CoMSIA analysis. These analyses of contour maps and sitemap assisted us to identify structural requirements for receptor binding and inhibition. New molecules have been designed based on these structural requirements as shown in figure 9. The dock pose of molecule 3 shown in figure 10a shows yellow hydrophobic region covering the phenyl rings that has
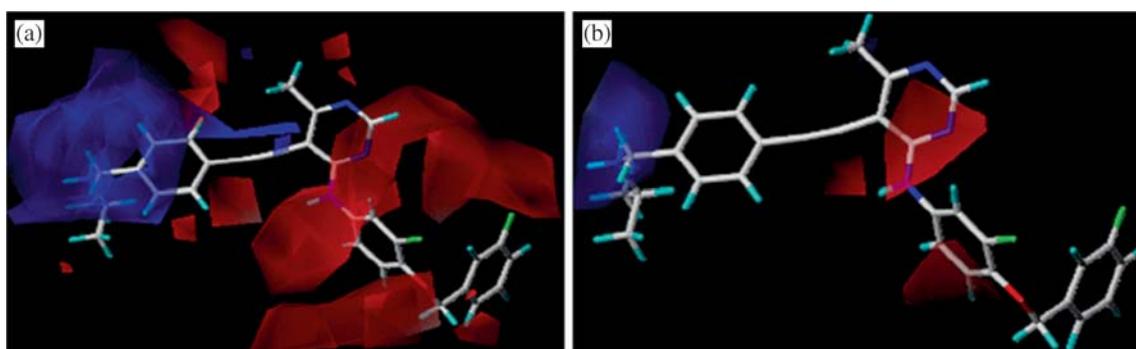

Figure 7. CoMFA and CoMSIA electrostatic fields. (a) CoMFA and (b) CoMSIA electrostatic standard deviation (SD* coefficient) contour maps illustrating electrostatic features in combination with molecule 3. Red contours indicate negative charge favoring activity, whereas blue contours indicate positive charge favoring activity.
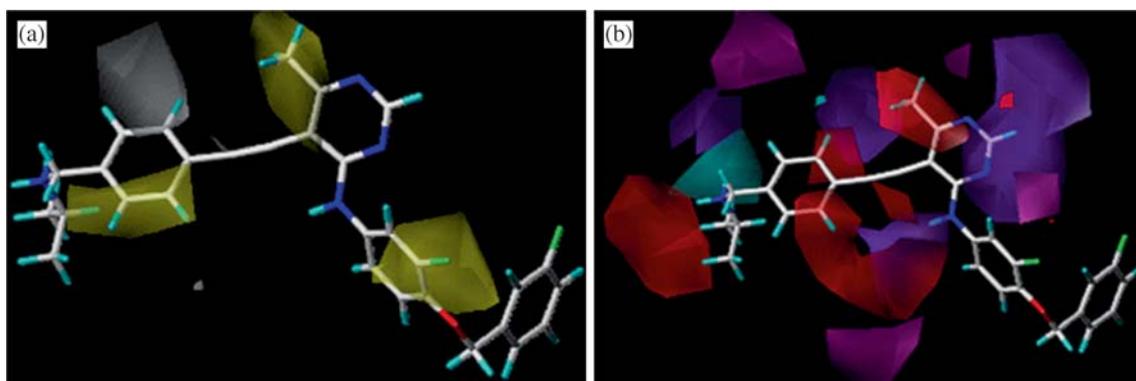

Figure 8. CoMSIA SD* coefficient contour maps illustrating, (A) hydrophobic, and (B) donor and acceptor features in combination with molecule 3. Yellow contour represents the hydrophobic favored region, white indicates the hydrophilic favored regions. The purple contour represents H-bond donor disfavored regions while cyan indicates H-bond donor favored region. Magenta and red contour represent H-bond acceptor favored and disfavored regions, respectively.
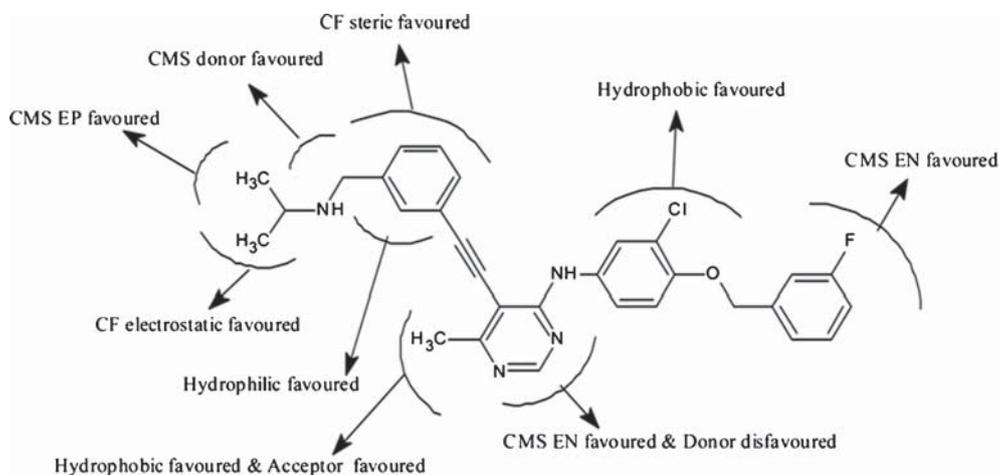
Figure 9. Structural requirements for EGFR inhibitors obtained from CoMFA (CF) and CoMSIA (CMS) contour map and sitemap analysis.
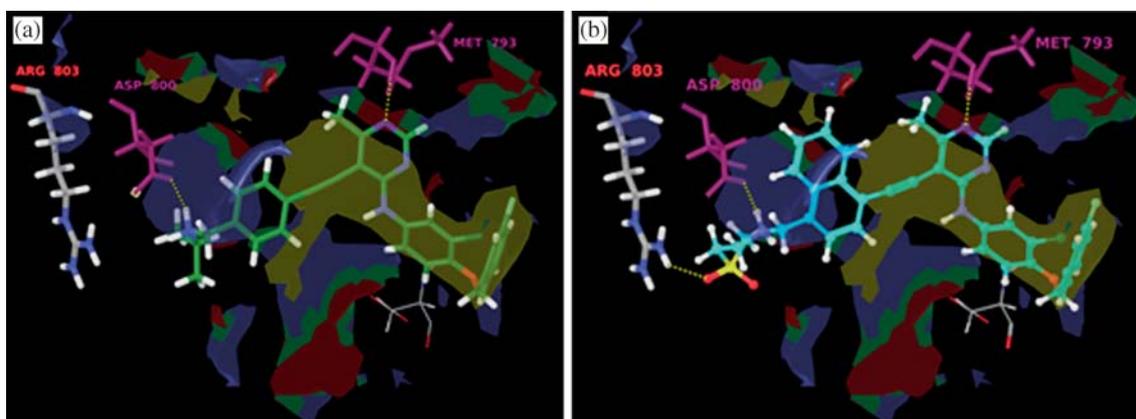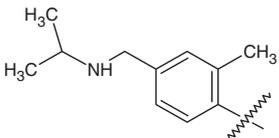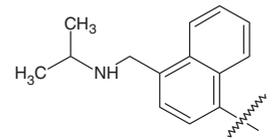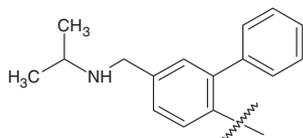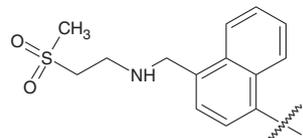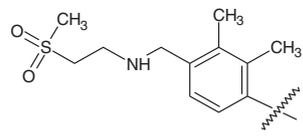

Figure 10. Sitemap analysis of, (a) most active molecule 3 of the data set, (b) dock pose of newly designed molecule N4 showing increased number of hydrogen bond interaction with active site amino acids. Yellow region represents hydrophobic, blue represents hydrogen bond donor, red represents hydrogen bond acceptor, and green represents hydrophilic.

$\pi - \pi$ interaction with Phe856, a hydrogen bond acceptor red region is seen over the nitrogen of pyrimidine group that interacts with Met793. Blue hydrogen bond donor region is seen near amine group of the tail region having salt bridge interaction with Asp800.

Bulky groups with hydrophilic character were substituted on the phenyl ring attached to alkynyl group. Groups like sulfonyl and amine were substituted to the side chain of phenyl ring resulting in more hydrogen bond interaction between the receptor and ligand (figure 10b). Designed molecules were docked into the active site, and they showed similar interactions with comparable dock score and predicted activity with respect to the the most active compound 3. Figure 10 shows dock poses of the most potent molecule 3 and newly designed molecule N4. These molecules are embedded in the protein along with

hydrogen bond interaction with active site amino acids. N4 showed one extra hydrogen bond interaction with Arg803. Methylsulphonyl side chain attached to the amine allows it to occupy the cavity having increased hydrogen bond interactions. Predicted pIC_{50} values were calculated and found to be better. Their structures and predicted pIC_{50} are given in table 3. ADME properties of the newly designed molecules were calculated and found to be in range as that of pyrimidine derivatives used in the present study. The toxicity properties of these molecules were calculated using pKCSM online server.[31] The new molecules showed similar properties as data set molecules, namely AMES toxicity, CYP and hERG inhibition, etc. The ADMET properties have been calculated and provided as Supplementary Information for both the new molecules and data set molecules.

Table 3. Structures of newly designed molecules with predicted pIC$_{50}$ values, Dock score(Kcal/mol) and ΔG_{bind}(Kcal/mol).
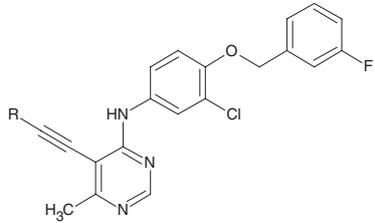
Mol	R	Pred pIC$_{50}$ CoMFA	Pred pIC$_{50}$ CoMSIA	SP Dock score (Kcal/mol)	XP Dock score (Kcal/mol)	ΔG_{bind} (Kcal/mol)
N1		7.855	7.387	−11.801	−12.166	−141.336
N2		7.704	8.229	−12.821	−13.368	−126.719
N3		7.884	7.595	−12.453	−12.889	−146.801
N4		7.503	7.285	−13.176	−13.242	−150.500
N5		7.681	7.288	−12.524	−13.237	−144.551

4. Conclusions

To understand the binding mode of molecules to receptors and rationalize structural requirements for inhibiton, molecular modeling tools like docking, MM/GBSA and 3D-QSAR studies are helpful. In the present study, molecular docking (SP and XP) and prime MM/GBSA calculations were performed to understand the binding mode of inhibitors with EGFR. CoMFA and CoMSIA methodologies were used to build a model for obtaining structural requirement for EGFR inhibitors. The generated model showed good statistical results in terms of high r^2 and q^2 values for all the models, this indicates the ablility of QSAR model to predict potent inhibitors from designed molecules. Based on detailed analysis of contour maps and sitemap, it is evident that improvement in EGFR binding affinity can be achieved by substitutional modification on phenyl ring attached to alkynyl group with bulky hydrophilic substituents. Maintaining hydrophobic character at phenyloxy group attached to pyrimidine ring of the best active molecule 3 has enhanced the binding affinity.

Supplementary Information (SI)

ADMET properties of newly designed molecules, Tables and Figures of Superimposition of structure pose and docked pose of co-crystalized ligand with RMS deviation of 1.92 Å, have been incorporated in SI, available at www.ias.ac.in/chemsci.

Acknowledgements

References

1. Ciardiello F and Tortora G 2008 *N. Engl. J. Med.* **358** 1160
2. Pao W and Chmielecki J 2010 *Nat. Rev. Cancer* **10** 760
3. Citri A and Yarden Y 2006 *Nat. Rev. Mol. Cell. Biol.* **7** 5
4. Downward J, Yarden Y, Mayes E, Scrace G, Totty N, Stockwell P, Ullrich A, Schlessinger J and Waterfield M D 1984 *Nature* **307** 521
5. Oda K, Matsuoka Y, Funahashi A and Kitano H 2005 *Mol. Syst. Biol.* **1** 200
6. Smaill J B, Palmer B D, Rewcastle G W, Denny W A, McNamara D J, Dobrusin E M, Bridges A J, Zhou H, Showalter H D H, Winters R T, Leopold W R, Fry D W, Nelson J M, Slintak V, Elliot W L, Roberts B J, Vincent P W and Patmore S J 1999 *J. Med. Chem.* **42** 1803
7. Tsou H-R, Mamuya N, Johnson B D, Reich M F, Gruber B C, Ye F, Nilakantan R, Shen R, Discafani C, DeBlanc R, Davis R, Koehn F E, Greenberger L M, Wang Y-F and Wissner A 2001 *J. Med. Chem.* **44** 2719
8. Wissner A, Berger D M, Boschelli D H, Floyd M B, Greenberger L M, Gruber B C, Johnson B D, Mamuya N, Nilakantan R, Reich M F, Shen R, Tsou H-R, Upeslacis E, Wang Y F, Wu B, Ye F and Zhang N 2000 *J. Med. Chem.* **43** 3244
9. Xu G, Abad M C, Connolly P J, Neeper M P, Struble G T, Springer B A, Emanuel S L, Pandey N, Gruninger R H, Adams M, Moreno-Mazza S, Fuentes-Pesquera A R and Middleton S A 2008 *Bioorg. Med. Chem. Lett.* **18** 4615
10. Xu G, Searle L L, Hughes T V, Beck A K, Connolly P J, Abad M C, Neeper M P, Struble G T, Springer B A, Emanuel S L, Gruninger R H, Pandey N, Adams M, Moreno-Mazza S, Fuentes-Pesquera A R, Middleton S A and Greenberger L M 2008 *Bioorg. Med. Chem. Lett.* **18** 3495
11. Mastalerz H, Chang M, Chen P, Fink B E, Gavai A, Han W-C, Johnson W, Langley D, Lee F Y, Leavitt K, Marathe P, Norris D, Oppenheimer S, Sleczka B, Tarrant J, Tokarski J S, Vite G D, Vyas D M, Wong H, Wong T W, Zhang H and Zhang G 2007 *Bioorg. Med. Chem. Lett.* **17** 4947
12. Mastalerz H, Chang M, Chen P, Fink B E, Gavai A, Han W -C, Johnson W, Langley D, Lee F Y, Leavitt K, Marathe P, Norris D, Oppenheimer S, Sleczka B, Tarrant J, Tokarski J S, Vite G D, Vyas D M, Wong H, Wong T W, Zhang H and Zhang G 2007 *Bioorg. Med. Chem. Lett.* **17** 2828
13. Rusnak D W, Lackey K, Affleck K, Wood E R, Alligood K J, Rhodes N, Keith B R, Murray D M, Knight W B, Mullin R J and Gilmer T M 2001 *Mol. Cancer Ther.* **1** 85
14. Li S, Guo C, Zhao H, Tang Y and Lan M 2012 *Bioorg. Med. Chem.* **20** 877
15. Li S, Guo C, Zhao H, Tang Y and Lan M 2012 *Bioorg. Med. Chem.* **22** 4004
16. Waterson A G, Petrov K G, Hornberger K R, Hubbard R D, Sammond D M, Smith S C, Dickson H D, Caferro T R, Hinkle K W, Stevens K L, Dickerson S H, Rusnak D W, Spehar G M, Wood E R, Griffin R J and Uehling D E 2009 *Bioorg. Med. Chem. Lett.* **19** 1332
17. Suzuki N, Shiota T, Watanabe F, Haga N, Murashi T, Ohara T, Matsuo K, Omori N, Yari H, Dohi K, Inoue M, Iguchi M, Sentou J and Wada T 2011 *Bioorg. Med. Chem. Lett.* **21** 1601
18. Suzuki N, Shiota T, Watanabe F, Haga N, Murashi T, Ohara T, Matsuo K, Omori N, Yari H, Dohi K, Inoue M, Iguchi M, Sentou J and Wada T 2012 *Bioorg. Med. Chem. Lett.* **22** 456
19. Cramer R D, Patterson D E and Bunce J D 1988 *J. Am. Chem. Soc.* **110** 5959
20. Cramer R D, Patterson D E and Bunce J D 1988 *Quant. Struct. Act. Relat.* **7** 18
21. Klebe G, Abraham U and Mietzner T 1994 *J. Med. Chem.* **37** 4130
22. Wold S, Johansson A and Cochi M 1993 In 3D QSAR in Drug Design, H Kubinyi (Ed.) (Leiden: ESCOM) p. 523
23. Schrödinger L L C 2010 *New York Glide Version* **5** 6
24. Friesner R A, Banks J L, Murphy R B, Halgren T A, Klicic J J, Mainz D T, Repasky M P, Knoll E H, Shelley M, Perry J K, Shaw D E, Francis P and Shenkin P S 2004 *J. Med. Chem.* **47** 1739
25. Taverna D M and Goldstein R A 2002 *Proteins* **46** 105
26. Srivastava H K and Sastry G N 2012 *J. Chem. Inf. Model* **52** 3088
27. Munnaluri R, Sivan S K and Manga V 2015 *Med. Chem. Res.* **24** 829
28. Fatima S, Bathini R, Sivan S K and Manga V 2012 *J. Recept. Signal Transduction* **32** 214
29. Sybyl-X 2.1 2012 (Tripos Associates: St. Louis, MO)
30. Gasteiger J and Marsili M 1980 *Tetrahedron* **36** 3219
31. Douglas E V P, Blundell T L and David B A 2015 *J. Med. Chem.* **58** 4066