



Received on 17 October, 2011; received in revised form 24 January, 2012; accepted 28 January, 2012

## LINEAR RESPONSE APPROXIMATION (LRA) APPROACHES FOR CALCULATING BINDING AFFINITIES OF QUINAZOLINE ANALOGUES AS ALK5 INHIBITORS

Md. Afroz Alam\*, Thalitha Jane Davina and Anurupa Devi

Department of Bioinformatics, Karunya University, Karunya Nagar, Coimbatore- 641 114, Tamil Nadu, India

### ABSTRACT

#### Keywords:

Quinazoline,  
Free energy of binding,  
Docking,  
Surface generalized Born continuum  
solvation model,  
Linear interaction energy

#### Correspondence to Author:

Dr. Md Afroz Alam

Assistant Professor (SG), Department of  
Bioinformatics, Karunya University,  
Karunya Nagar, Coimbatore, 641114,  
Tamil Nadu, India

Quinazoline and its analogues have important therapeutic value in the treatment of cancer to induce apoptosis in cancer cells in a proliferation-independent manner. The binding free energies of quinazoline based inhibitors of kinase were computed using linear interaction energy method with a surface generalized Born (SGB) continuum solvation model in the human ALK5 kinase domain. A training set of 20 quinazoline analogues was used to build a binding affinity model for estimating the free energy of binding for 12 inhibitors (test set) with diverse structural modifications. The root mean square error (RMSE) between the experimental and predicted activity values was 0.02  $\mu$ M which is comparable to the level of accuracy achieved by the most accurate methods, such as free energy perturbation (FEP) or thermodynamic integration (TI). The correlation coefficient between experimental and predicted activity based on SGB-LIE estimation for the test set compounds is also significant ( $R^2 = 0.9693$ ). Low levels of RMSE for the majority of inhibitors establish the structure-based LIE method as an efficient tool for generating more potent and specific inhibitors of kinase by testing rationally designed lead compounds based on quinazoline derivatives.

**INTRODUCTION:** Drug discovery and development is a cost and time intensive process involving many considerations in molecular design, synthesis, testing and evaluation of drug effects ranging from local interactions at the molecular/cellular level to global effects on the organism and population. *In silico* drug discovery process comprises mainly three stages.

First stage includes identification of a therapeutic drug target building a heterogeneous small molecule library tested against it. This is followed by development of a virtual screening protocol initialized by either docking the small molecule from the library or building these structures in the active site employing de novo design methods. Second stage is about binding affinity prediction/scoring and optimization of the set

molecules until the desired binding affinity is achieved. Molecular simulations are performed to obtain a more realistic appreciation of binding affinity and its dependence on solvent, salt and dynamics. The selected hits are checked for specificity by docking at binding sites of other known drug targets and third and final stage is the selected hits are subjected to detail *in silico* profiling studies<sup>1</sup>.

Multicellular organisms live in a complex milieu where signalling pathways contribute to critical links, for their existence. Tyrosine kinases are important mediators of this signal transduction process, leading to cell proliferation, differentiation, migration, and metabolism and programmed cell death<sup>2</sup>.

Tyrosine kinases are a family of enzymes, which catalyzes phosphorylation of select tyrosine residues in target proteins, using ATP. This covalent post-translational modification is a pivotal component of normal cellular communication and maintenance of homeostasis<sup>3</sup>.

Though, their activity is tightly regulated in normal cells, they may acquire transforming functions due to mutation(s), over expression and autocrine paracrine stimulation, leading to malignancy<sup>3</sup>. During cell division, the bipolar organization of mitotic spindles is essential for proper segregation of chromosomes. Inhibition of mitotic-spindle formation is an interesting target in cancer chemotherapy. However, so far used anti-mitotic agents target microtubule stability (e.g., tubulin) during cancer treatment (e.g., taxanes) and cause serious side effects, such as neurotoxicity. Furthermore, development of resistance against these anti-mitotic agents restricts their use.

Transforming growth factor- $\beta$ 1 (TGF- $\beta$ 1) is a key mediator in progressive fibrosis in the kidney, liver, heart, lung, bone marrow and skin, which enhances extracellular matrix production by increasing the transcription of matrix proteins, for example, fibronectin and collagen, and inhibiting enzymes responsible for matrix degradation<sup>4</sup>. TGF- $\beta$ 1 signals through a family of transmembrane serine/threonine kinase receptors. These receptors can be divided into two classes, the type I or activin like kinase (ALK) receptors and type II receptors.

Specifically, the binding of TGF- $\beta$ 1 to the type II receptor causes phosphorylation of the GS domain of the TGF- $\beta$  type I receptor, ALK5. The ALK5 receptor, in turn, phosphorylates the cytoplasmic proteins smad2 and smad3 at two carboxyl terminal serines. The phosphorylated smad proteins form heteromeric complexes with smad4, after which the complex translocates into the nucleus to affect gene transcription<sup>5</sup>. Therefore, identification of small-molecule inhibitors of the kinase activity of the TGF- $\beta$  type I receptor (also named activin-like kinase or ALK5) to block the pro-fibrotic effect of over expression of TGF- $\beta$ 1 represents an attractive target for the treatment of fibrotic diseases<sup>4</sup> and cancer<sup>5,6</sup>.

Quinazoline is potent and selective ALK5 inhibitors over p38MAP kinase from a rational drug design approach based on co-crystal structures in the human ALK5 kinase domain. Quinazoline and their derivatives constitute an important class of heterocyclic compounds. Many of them show insecticidal<sup>7</sup>, analgesic<sup>8</sup>, antifungal<sup>9</sup>, antibacterial<sup>10</sup>, anticancer<sup>11</sup>, anti-inflammatory<sup>11</sup> activities. Quinazoline nucleus is found in many bioactive natural products.

Different derivatives of quinazoline have been demonstrated to bind to ALK5. This binding mode was confirmed by the determination of a 1.80 Å X-ray structure of  $\alpha$  kinase complexes with quinazoline (PDB\_ID: 3GXL), showing that quinazoline also binds at the ALK5 site<sup>12</sup>. Over the years, a large number of analogues of quinazoline have been identified as kinase inhibitors. Since a wide variety of molecular scaffolds are available for optimization, this diversity presents a significant challenge to determining the essential features for activity.

A rational approach for the discovery of a pharmaceutically acceptable, economically viable activity model awaits development of a predictive quantitative structure-activity relationship. With the advent of parallel synthesis methods and technology, we might expect the number of quinazoline analogues to be tested to grow dramatically. Combinatorial methods could also be envisioned as a semi-rational approach to this discovery strategy. One method of orchestrating these strategies is to make use of linear interaction energy (LIE) models for the rapid prediction and virtual pre-screening of cytotoxic activity.

The linear interaction energy approximation is a way of combining molecular mechanics calculations with experimental data to build a model scoring function for the evaluation of ligand-protein binding free energies<sup>13</sup>. The LIE method is a semi-empirical model that has become widely used to predict protein-ligand binding affinities<sup>14</sup>. In LIE, the free ligand in water and the solvated protein-ligand complex are simulated and from these two calculations the ligand surrounding electrostatic and van der Waals (vdw) energies are collected. The binding free energy is then evaluated as proposed by Aqvist *et al.*<sup>14</sup>.

A continuum solvation model was developed based on the proposed LIE method by adding continuum electrostatic ligand-water interaction energies by using an equivalent form of equation <sup>15</sup>. However, the proposed generalized Born (GB)-LIE method overestimates the change in solvation energy and this is caused by consistent underestimation of the effective born radii in the protein-ligand complex <sup>15</sup>. To further assess the usefulness of continuum models for estimating binding free energies, more accurate GB models should be carried out.

The LIE method has been applied on a number of protein-ligand systems with promising results producing small errors on the order of 1 kcal/mol for free energy prediction <sup>16</sup>. This approach could then be applied to larger sets of inhibitors and contribute to fast and efficient ligand design. At present, a linear interaction energy method for rational design of quinazoline analogues for kinase polymerization inhibition has not been determined. The availability of structural information on kinase facilitates understanding the structure-activity relationships (SAR) for kinase polymerization inhibition. In this study, we have applied a structure-based linear interaction energy method implementing a surface generalized Born (SGB) continuum model for solvation to build a binding affinity model for estimating the binding free energy for a diverse set of quinazoline analogues with kinase.

In this regards, the development of virtual screening models based on SGB-LIE to facilitate the search for the potential drugs with low toxicity and better biological activity based on quinazoline skeleton and binding energy calculation based on the bonding and non bonding interaction and prediction correlation model with biological activity have been studied. The magnitude of predicted activity based upon linear interaction approach of inhibitors to kinase directly correlates with the experimental potency of these inhibitors.

Hence, fast and accurate estimation of binding free energies provides a means to screen the compound libraries for lead optimization and for generating more potent and specific inhibitors of kinase by testing rationally designed lead compounds based on quinazoline derivatization.

## MATERIALS AND METHODS:

**LIE Methodology:** The LIE method employs experimental data on binding free energy values for a set of ligands (referred as training set) to estimate the binding affinities for a set of novel compounds. The method is based on the linear response approximation (LRA), which dictates that binding free energy of a protein-ligand system is a function of polar and non-polar energy components that scale linearly with the electrostatic and van der Waals interactions between a ligand and its environment. The free energy of binding (FEB) for the complex is derived from considering only two states:

- (1) Free ligand in the solvent and;
- (2) Ligand bound to the solvated protein

The conformational changes and entropic effects pertaining to unbound receptor are taken into account implicitly and only interactions between the ligand and either the protein or solvent are computed during molecular mechanics calculations. Among the various formulations of the LIE methodology developed in the past, the SGB-LIE method <sup>16</sup> has been shown to be 1 order of magnitude faster than the methods based on explicit solvent with the same order of accuracy. In the LIE method,

$$\Delta G_{\text{bind}} = \alpha \langle \Delta U_{\text{ele}} \rangle + \beta \langle \Delta U_{\text{vdw}} \rangle + \gamma \langle \Delta S_{\text{SASA}} \rangle \dots \dots \dots (1)$$

where  $\langle \Delta U_{\text{ele}} \rangle$  and  $\langle \Delta U_{\text{vdw}} \rangle$  denotes the average change in the electrostatic and van der Waals interaction energy of the ligand in the free and bound states, respectively and  $\langle \Delta S_{\text{SASA}} \rangle$  is the change in the solvent-accessible surface area (SASA) of the ligand. The  $\alpha$ ,  $\beta$ , and  $\gamma$  terms are adjustable parameters that need to be determined by fitting the experimental data on the training set compounds. The SGB-LIE method also offers better accuracy in treating the long-range electrostatic interactions.

However, the SGB-LIE method used in this studied is based on the original formulation proposed by Jorgensen and implemented in *Liaison* using the OPLS-2005 force field. A novel feature of *Liaison* is that the simulation takes place in implicit (continuum) rather than explicit solvent- hence, the name *Liaison*, for

Linear Interaction Approximation in Implicit Solvation. The explicit-solvent version of the methodology was first suggested by Aqvist, based on approximating the charging integral in the free-energy-perturbation formula with a mean-value approach, in which the integral is represented as half the sum of the values at the endpoints, namely the free and bound states of the ligand.

The empirical relationship used by Liaison is shown below:

$$\Delta G_{bind} = \alpha (\langle U_{ele}^b \rangle - \langle U_{ele}^f \rangle) + \beta (\langle U_{vdw}^b \rangle - \langle U_{vdw}^f \rangle) + \gamma (\langle U_{cav}^b \rangle - \langle U_{cav}^f \rangle) \quad (2)$$

Here  $\langle \rangle$  represents the ensemble average,  $b$  represents the bound form of the ligand,  $f$  represents the free form of the ligand, and  $\alpha$ ,  $\beta$  and  $\gamma$  are the coefficients.  $U_{ele}$ ,  $U_{vdw}$  and  $U_{cav}$  are the electrostatic, van der Waals and cavity energy terms in the SGB continuum solvent model. The cavity energy term,  $U_{cav}$ , is proportional to the exposed surface area of the ligand. Thus, the difference:  $\langle U_{cav}^b \rangle - \langle U_{cav}^f \rangle$  measures the surface area lost by contact with the receptor. The energy terms involved can be computed using energy minimization, molecular dynamics, or Monte Carlo calculations. In the SGB model of solvation, there is no explicit van der Waals or electrostatic interaction between the solute and solvent.

The contribution for net free energy of solvation comes from two energy terms, namely, reaction field energy ( $U_{rxn}$ ) and cavity energy ( $U_{cav}$ ):  $U_{SGB} = U_{rxn} + U_{cav}$ . The cavity and reaction field energy terms implicitly take into account the van der Waals and the electrostatic interactions, respectively, between the ligand and solvent. The application of the SGB-LIE method for a given protein-ligand system essentially involves computing four energy components, i.e., the van der Waals and Columbic energy between the ligand and protein and the reaction field and cavity energy between the ligand and continuum solvent. The total electrostatic energy in the SGB-LIE method is the sum of Columbic and reaction field energy terms.

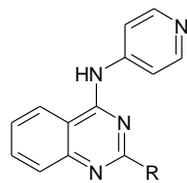
**Computational Description:** Preparation of receptor and ligands was done using the Schrödinger package from Schrödinger Inc<sup>17</sup>. All the calculations for the

SGB-LIE method were performed in the Liaison package from Schrödinger Inc<sup>18</sup>. The Liaison module performs LIE calculations in the OPLS force field with a residue-based cutoff of 15Å. The OPLS force field was also used for charge assignment and all energy calculations.

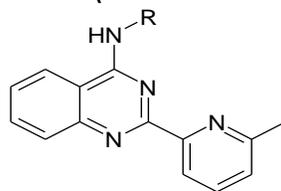
**Receptor preparation:** The X-ray structure of the complex between quinazoline and kinase protein (PDB\_ID: 3GXL) has been used as initial structure in the preparation of quinazoline binding site. After manual inspection and cleaning of structure we retained a complex composed of protein chains  $\alpha$  and quinazoline ligand. Hydrogen was added to the model automatically via the Maestro interface leaving no lone pair and using an explicit all-atom model<sup>17</sup>. All the water molecules were removed from the complex. The multi step Schrödinger's Protein preparation tool (Pprep) has been used for final preparation of protein. Pprep neutralizes side chains that are not close to the binding cavity and do not participate in salt bridges<sup>17</sup>.

This step is then followed by restrained minimization of co-crystallized complex, which reorients side chain hydroxyl groups and alleviates potential steric clashes. Progressively weaker restraints (tethering force constants 3, 1, 0.3, 0.1) were applied to non hydrogen atoms only. The complex structure was energy minimized using OPLS\_2005 force field and the conjugate gradient algorithm, keeping all atoms except hydrogen fixed. The minimization was stopped either after 1000 steps or after the energy gradient converged below 0.01 kcal/mol. The energy-minimized receptor structure was subsequently used for docking of quinazoline analogues and SGB-LIE calculations.

**Preparation of Compound Library:** Quinazoline is well known for its antitumor activity. However, the clinical application of it and its analogues in the treatment of cancer has been limited by severe toxic side effects during administration of the drugs<sup>12, 19</sup>. However, new findings related to their activities, mechanism of action and pharmacological properties have been unexplored. A total of 37 quinazoline analogues were used in the study and were taken from various sources belonging to different ring modifications<sup>12, 19</sup>. For better interpretation all these compounds were divided into following sub libraries.

**TABLE 1: COMPOUNDS 1-4 WERE PREPARED BY COUPLING OF 4-AMINOPYRIDINE WITH 4-CHLOROQUINAZOLINE INTERMEDIATES USING EITHER STANDARD BUCHWALD CONDITIONS****1-4**

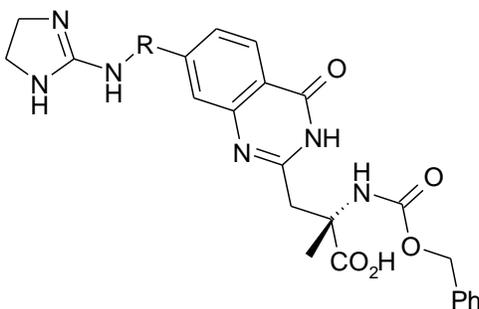
Analogue	R	ALK5 Expt. IC <sub>50</sub> (μM)	Analogue	R	ALK5 Expt. IC <sub>50</sub> (μM)
1		0.19	3		0.02
2		0.05	4		0.03

**TABLE 2: THE ANALOGUES 5-11 WERE PREPARED FROM 4-CHLORO-2-(6-METHYL-2-PYRIDINYL)****5-11**

Analogue	R	ALK5 Expt. IC <sub>50</sub> (μM)	Analogue	R	ALK5 Expt. IC <sub>50</sub> (μM)
5		0.20	9		0.59
6		0.26	10		0.64
7		0.04	11		3.71
8		0.03			

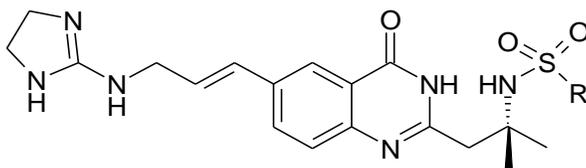
**TABLE 3: COUPLING OF THE 2-AMINO-4-METHYL-3-THIOPHENECARBOXAMIDE AND 3-AMINO-2-THIOPHENE CARBOXAMIDE ACID USING STANDARD PEPTIDE GAVE COMPOUNDS 12-15**

Analogue	Structure	ALK5 Expt. IC <sub>50</sub> (μM)	Analogue	Structure	ALK5 Expt. IC <sub>50</sub> (μM)
12		0.09	14		0.02
13		0.02	15		0.02

**TABLE 4: QUINAZOLINE DERIVATIVES WITH DIFFERENT LINKERS FROM 16-19. THESE ARE ALLYL WHICH IS UNIVALENT, UNSATURATED ORGANIC RADICAL C<sub>3</sub>H<sub>5</sub>.**

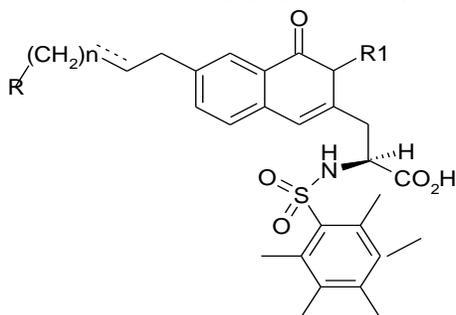
16-19

Analogues	R	Expt. IC <sub>50</sub>	Analogues	R	Expt. IC <sub>50</sub>
16		0.55	18		0.52
17		0.52	19		0.09

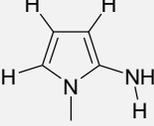
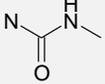
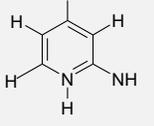
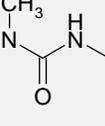
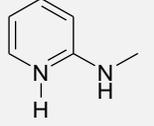
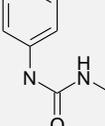
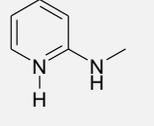
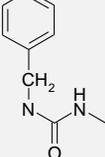
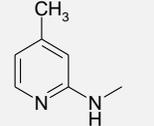
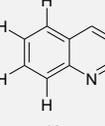
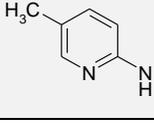
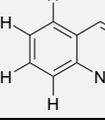
**TABLE 5: THE COMPOUNDS FROM 20-25 ARE METHOXY-SUBSTITUTED BENZENESULFONAMIDES WITH DIFFERENT ARYLSULFONYL GROUPS**

20-25

Analogues	R	Expt. IC <sub>50</sub>	Analogues	R	Expt. IC <sub>50</sub>
20		0.06	23		0.03
21		0.05	24		0.33
22		0.08	25		0.67

**TABLE 6: GUANIDINE MIMETIC GROUPS ARE 5 AND 6-MEMBERED CYCLIC AMIDINES**

26-27

Analogue	R	R1	Expt. IC <sub>50</sub>	Analogue	R	R1	Expt. IC <sub>50</sub>
26		H	0.07	32		H	0.22
27		H	0.11	33		H	3.9
28		H	0.64	34		H	1.7
29		CH <sub>3</sub>	4.6	35		H	0.82
30		H	0.88	36		H	14
31		H	1.8	37		H	2.2

All these quinazoline analogues were built from the scaffold by different ring modification and substitution of functional groups as mentioned in **Table 1-6**. We used ISIS Draw 2.3 software for sketching structures and converting them to their 3D representation by using ChemSketch 3D viewer of ACDLABS 8.0. LigPrep<sup>17</sup> was used for final preparation of ligands from libraries. LigPrep is a utility of Schrödinger software suit that combines tools for generating 3D structures from 1D (Smiles) and 2D (SDF) representation, searching for tautomers and steric isomers and performing a geometry minimization of ligands. The ligands were minimized by means of Molecular Mechanics Force Fields (MMFFs) with default setting. Each of these compounds had associated in vitro cytotoxicity values (IC<sub>50</sub> values reported in  $\mu$ M) against ALK5 and TGF- $\beta$  cell line.

**Docking of the ligands:** All the ligands were docked to the ALK5 receptor using Glide version 4.0<sup>17</sup>. After ensuring that protein and ligands are in correct form for docking, the receptor-grid files were generated using grid-receptor generation program, using van der

Waals scaling of the receptor at 0.4. The default size was used for the bounding and enclosing boxes was generated at the centroid of the ALK5 binding site by selecting the bound quinazoline ligand. The ligands were docked initially using the "standard precision" method and further refined using "xtra precision" Glide algorithm.

For the ligand docking stage, Vander Waals scaling of the ligand was set at 0.5. Of the 50,000 poses that were sampled, 4,000 were taken through minimization (conjugate gradients 1,000) and the 30 structures having the lowest energy conformations were further evaluated for the favourable Glide docking score. A single best conformation for each ligand was considered for further analysis.

**LIE Calculations:** The docked complex corresponding to each analogue was transported to the Liaison package for subsequent SGB-LIE calculations. Sampling technique such as molecular dynamics (MD) has been used for LIE conformation space sampling in the present work.

The system was initially heated to 300 K for 5 ps and then subjected to a MD simulation for 25 ps. A residue-based cut-off of 12 Å was set for the non-bonding interactions. The non-bonded pair list was updated every 10 fs. The time integration step of 1.0 fs and sampling LIE energies every 10 steps was used.

During the MD simulations, all the residues of the receptor beyond 12 Å from the bound ligands were frozen. Similarly, the average LIE energies for the ligands were obtained using the OPLS-2005 force field. The average LIE energy terms were used for building binding affinity model and free energy estimation for quinazoline analogues. The  $\alpha$ ,  $\beta$  and  $\gamma$  LIE fitting parameters were determined based on Gaussian elimination method using Mat lab 6.5 as described by fitting the experimental data on the training set compounds<sup>20</sup>.

In order to explore the reliability of the proposed model we used the cross validation method. Prediction error sum of squares (*PRESS*) is a standard index to measure the accuracy of a modelling method based on the cross validation technique. The  $r_{cv}^2$  was calculated based on the *PRESS* and *SSY* (sum of squares of deviations of the experimental values from their mean) using following formula.

$$r_{cv}^2 = 1 - \frac{PRESS}{SSY} = 1 - \frac{\sum_{i=1}^n (y_{exp} - y_{pred})^2}{\sum_{i=1}^n (y_{exp} - \bar{y})^2}$$

Where  $y_{exp}$ ,  $y_{pred}$  and  $\bar{y}$  are the predicted, observed and mean values of the cytotoxic activities of the quinazoline analogues. The cross validation analysis performed by using the leave one out (LOO) method in which one compound removed from the data set and its activity predicted using the model derived from the rest of the data points. The cross-validated correlation coefficient ( $q^2$ ) that resulted in optimum number of components and lowest standard error of prediction were considered for further analysis and calculated using following equations:

$$q^2 = 1 - \frac{\sum_y (y_{pred} - y_{observed})^2}{\sum_y (y_{observed} - y_{mean})^2}$$

$$PRESS = \sum_y (y_{predicted} - y_{observed})^2$$

Where  $y_{pred}$ ,  $y_{observed}$  and  $y_{mean}$  are the predicted, observed and mean values of the cytotoxic activities of the quinazoline analogues and *PRESS* is the sum of the predictive sum of squares.

$$r_{pred}^2 = \frac{SD - PRESS}{SD}$$

The predictive ability of the models is expressed by the  $r^2$  predictive value, which is analogous to cross-validated  $r^2$  ( $q^2$ ).

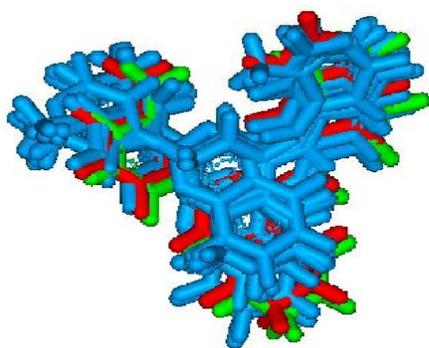
## RESULTS AND DISCUSSION:

**LIE Model:** The original crystal structure of ALK5 complex (PDB ID: 3GXL) was used to validate the Glide-XP docking protocol. This was done by moving the co-crystallized quinazoline ligand outside of active site and then docking it back into the active site. All the configurations after docking were taken into consideration to validate the result. The  $\Delta$ RMSD was calculated for eleven best configurations in comparison to the co-crystallized quinazoline and the value was found to be in between -0.94 to 0.52 Å and it is shown in **Table 7**. This revealed that the docked configurations have similar binding positions and orientations within the binding site and are similar to the crystal structure.

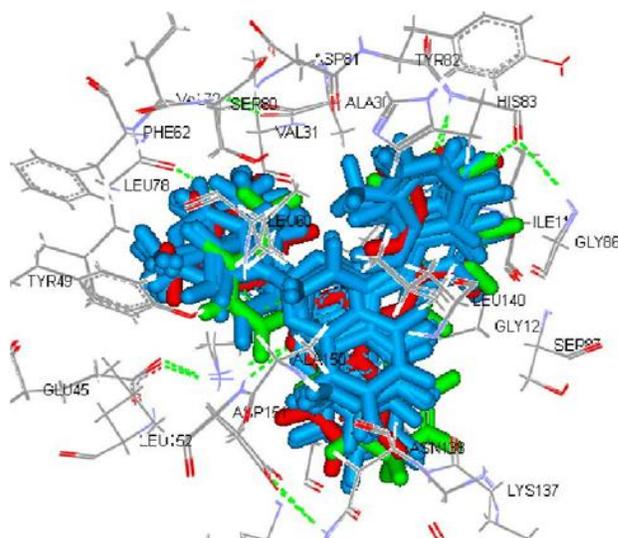
**TABLE 7: GLIDE SCORE AND RMSD OF 11 LOWEST CONFIGURATIONS OF CO-CRYSTAL QUINAZOLINE IN KINASE PROTEIN (3GXL)**

Configuration	Glide score	$\Delta G_{score}$	RMSD (Å <sup>o</sup> )	$\Delta$ RMSD (Å <sup>o</sup> )
4	-10.43	0	0.21	0
5	-9.47	-0.96	0.30	-0.09
37	-8.42	-1.05	0.56	-0.26
31	-8.10	-0.32	0.84	-0.28
15	-7.03	-1.13	0.27	0.37
27	-7.01	-0.02	0.78	-0.51
23	-6.27	-0.68	0.78	0
25	-5.84	-0.43	0.34	0.44
20	-5.01	-0.83	1.28	-0.94
16	-4.83	-0.18	0.76	0.52

The best docked structure, which is the configuration with the lowest Glide score is compared with the crystal structure and is shown in **Figure 1**. These docking results illustrate that the best-docked quinazoline complex agrees well with its crystal structure and that Glide (XP)-docking protocol successfully reproduces the crystal ALK5- quinazoline complex is shown in **Figure 2**.



**FIG. 1: SUPERPOSITION OF 10 BEST DOCKED CONFIGURATIONS OF QUINAZOLINE ON CRYSTAL STRUCTURE (RED-STICK).  $\Delta$ RMSE (HEAVY ATOMS) = -0.94 TO 0.52Å**

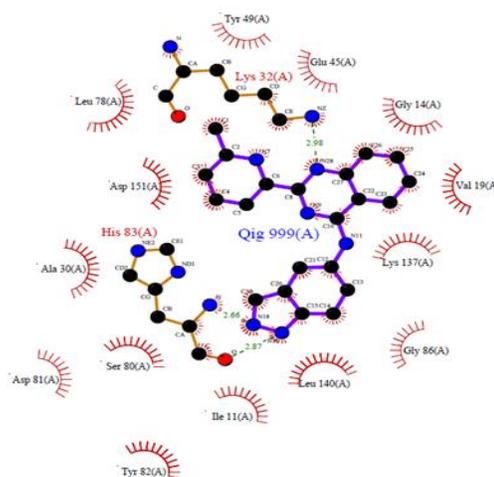


**FIG. 2: SHOWS THE VALIDATION OF DOCKING. THIS REVEALS THAT DOCKED CONFIGURATIONS HAVE SIMILAR BINDING POSITIONS AND ORIENTATIONS WITHIN THE BINDING SITE AND ARE SIMILAR TO THE CRYSTAL STRUCTURE**

With the wide range of difference between the  $IC_{50}$  values and the large diversity in the structures, the combined set of 20 ligands is ideal to be considered as a training set, as the set does not suffer from bias, due to the similarity of the structures. Also, the training set containing 20 analogues contains enough data points not to suffer from over parameterization by the LIE model. Training set compounds were docked into the ALK5 binding site of kinase protein and the SGB-LIE calculations were performed using the Liaison module. The simulations were performed both for the ligand-free and ligand-bound state.

The largest contribution for the binding energy comes from the Vander Waals (vdw) interactions. The energy values in **Table 8** were used to fit equation 2 using the Gaussian elimination method. The values obtained for the three fitting parameters,  $\alpha$ ,  $\beta$  and  $\gamma$  are -0.141, -0.093 and -1.071, respectively. The large value of the

**Validation of Docking:** We have applied the SGB-LIE method to a training set of 20 quinazoline analogues to build a binding affinity model that was then used to compute the free energy of binding and predicted  $pIC_{50}$  for a test set of 12 analogues. Further the SGB-LIE model developed was validated using 5 new quinazoline analogues for which the experimental kinase polymerization inhibition was known. The training set for building the binding affinity model was comprised of subsets of quinazoline analogues. For all the subsets included in the training set the experimental  $IC_{50}$  values against the ALK5 and TGF- $\beta$  cell lines are available.



cavity energy term signifies the fact that binding is largely driven by the ligand's ability to bury itself in the binding cavity, which is understandable given that most of the ligands are highly hydrophobic in nature. Even though the R value is low, vdw interactions contribute significantly toward the free energy of binding due to the large magnitude of the vdw interaction term.

In **Table 8**, the binding free energy ( $\Delta G$ ) values obtained from estimated using SGB-LIE fitting parameters are presented. The root mean square error (RMSE) between the experimental values and the values obtained by the fit was 0.02  $\mu$ M, which is an indicator of the robustness of the fit. The quality of the fit can also be judged by the value of the squared correlation coefficient ( $R^2$ ), which was 0.9998 for the training set. **Figure 3** graphically shows the quality of fit. The statistical significance of the SGB-LIE model is

evaluated by the correlation coefficient R, standard error s, F-test value, significance level of the model P, leave-one-out cross validation coefficient  $q^2$  and predictive error sum of squares PRESS.

$$\Delta G_{bind} = (-0.141) \langle \Delta U_{ele} \rangle + (-0.093) \langle \Delta U_{vdw} \rangle + (-1.071) \langle \Delta S_{ASA} \rangle \quad (3)$$

( $n = 19$ ,  $R^2 = 0.9998$ ,  $r^2_{pred} = 0.9978$ ,  $s = 0.01$ ,  $F = 71.31$ ,  $P = 0.001$ ,  $q^2 = 0.9979$ ,  $PRESS = 12.35$ )

The SGB-LIE model developed in this study is statistically ( $q^2 = 0.9979$ ,  $R^2 = 0.9998$ ,  $F = 71.31$ ) best fitted and consequently used for prediction of cytotoxic activities (pIC<sub>50</sub>) of training and test sets of molecules as reported in **Table 8** and **Table 9**.

**TABLE 8: AVERAGE ELECTROSTATIC (ELE), VANDER WAALS (VDW) AND CAVITY (CAV) ENERGY TERMS AS BINDING AFFINITY MODEL CALCULATIONS FOR THE TRAINING SET USING SGB-LIE METHOD**

Analogue	Glide Score	<sup>1</sup> <U <sub>ele</sub> >	<sup>1</sup> <U <sub>vdw</sub> >	<sup>1</sup> <U <sub>cav</sub> >	$\Delta G_{bind}$ LIE	<sup>3</sup> Expt. IC <sub>50</sub>	<sup>4</sup> Pred. IC <sub>50</sub>
1	-7.24	10.3	-41.6	3.1	-0.87	0.19	0.22
2	-10.2	11.9	-44.5	5.7	-0.28	0.05	0.07
3	-9.52	12.3	-53.0	6.4	-0.18	0.02	0.05
5	-9.47	11.5	-41.2	5.8	-0.75	0.20	0.19
6	-8.52	9.9	-34.2	5.2	-1.00	0.26	0.26
7	-6.87	7.2	-49.9	5.4	-0.15	0.04	0.04
8	-8.05	12.3	-53.8	6.4	-0.11	0.03	0.03
9	-5.51	12.4	-47.6	7.9	-2.27	0.59	0.58
10	-7.71	8.7	-54.0	8.2	-2.53	0.65	0.65
11	-7.62	8.7	-45.6	18.6	-14.43	3.72	3.71
16	-4.83	12.6	-43.5	7.4	-2.10	0.55	0.54
17	-8.31	11.1	-30.3	6.0	-2.04	0.52	0.52
18	-6.76	18.5	-62.9	9.8	-2.07	0.52	0.53
19	-5.09	10.2	-65.0	7.3	-0.33	0.09	0.09
20	-5.01	8.9	-53.9	6.1	-0.26	0.06	0.07
21	-8.37	10.3	-41.9	5.2	-0.22	0.05	0.06
22	-7.85	11.1	-52.0	6.3	-0.36	0.08	0.09
23	-6.27	13.9	-38.0	5.3	-0.18	0.03	0.05
24	-6.61	5.9	-44.0	5.8	-1.28	0.33	0.33
25	-5.84	8.2	-45.6	7.5	-2.64	0.67	0.68

<sup>1</sup><U<sub>ele</sub>>, <sup>1</sup><U<sub>vdw</sub>> and <sup>1</sup><U<sub>cav</sub>> energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment. <sup>2</sup> $\Delta G_{bind,LIE}$  refer to the absolute free energy values obtained using SGB-LIE method. <sup>3</sup>Expt. IC<sub>50</sub> = Experimental cytotoxic activity of ligands. <sup>4</sup>Pred. IC<sub>50, pred</sub> refers to predicted cytotoxic activity of ligands and is estimated using the relationship:  $Pred. IC_{50,pred} = -(\Delta G_{bind,LIE} / 2.303 RT)$ , where 298 K is used in the work for temperature T.

**TABLE 9 AVERAGE ELECTROSTATIC (ELE), VAN DER WAALS (VDW) AND CAVITY (CAV) ENERGY TERMS AS WELL AS BINDING AFFINITY MODEL CALCULATIONS FOR THE TEST SET USING SGB-LIE METHOD.**

Analogue	Glide Score	<sup>1</sup> <U <sub>ele</sub> >	<sup>1</sup> <U <sub>vdw</sub> >	<sup>1</sup> <U <sub>cav</sub> >	$\Delta G_{bind}$ LIE	<sup>3</sup> Expt. IC <sub>50</sub>	<sup>4</sup> Pred. IC <sub>50</sub>
26	-6.95	11.1	-45.3	2.8	-0.35	0.08	0.09
27	-7.01	16.9	-51.3	3.4	-1.25	0.11	0.32
28	-9.48	12.4	-45.1	4.7	-2.58	0.64	0.66
29	-9.76	14.9	-41.9	18.3	-17.79	4.60	4.57
30	-8.05	10.4	-41.1	5.5	-3.53	0.88	0.91
31	-8.10	17.0	-46.8	7.6	-6.15	1.80	1.58
32	-6.83	11.5	-53.8	6.9	-3.96	0.22	1.02
33	-4.37	11.7	-51.6	16.5	-14.46	3.90	3.72
34	-8.54	11.7	-53.9	8.8	-6.04	1.70	1.55
35	-5.64	6.7	-45.0	5.7	-2.86	0.82	0.73
36	-7.89	4.3	-47.9	7.2	-3.82	0.82	0.98
37	-8.42	8.0	-48.6	10.9	-8.26	2.20	2.12

<sup>1</sup><U<sub>ele</sub>>, <sup>1</sup><U<sub>vdw</sub>> and <sup>1</sup><U<sub>cav</sub>> energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment. <sup>2</sup> $\Delta G_{bind,LIE}$  refer to the absolute free energy values obtained using SGB-LIE method. <sup>3</sup>Expt. IC<sub>50</sub> = Experimental cytotoxic activity of ligands. <sup>4</sup>Pred. IC<sub>50, pred</sub> refers to predicted cytotoxic activity of ligands and is estimated using the relationship:  $Pred. IC_{50,pred} = -(\Delta G_{bind,LIE} / 2.303 RT)$ , where 298 K is used in the work for temperature T.

The predicted activity calculated from free energy of binding is satisfactory with small deviation compared with experimental activity of training and test sets of molecules. The calculated free energy of binding represents the experimental activity well. Theoretically, FEB can be partitioned into several components: vdw, electrostatic and solvent accessible surface area (SASA)<sup>14</sup>.

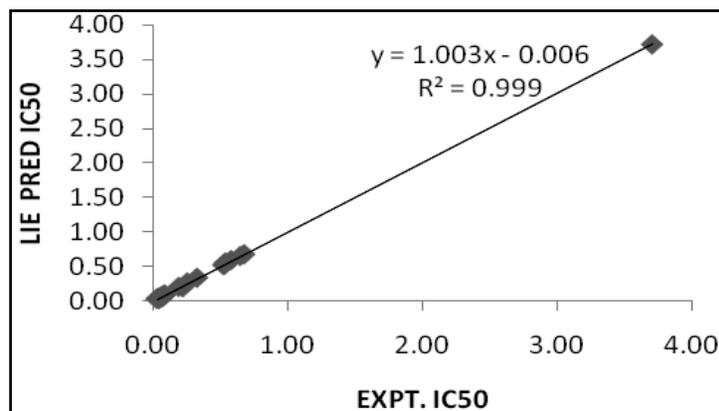
In this study the SASA energy term has been replaced by the cavity energy term<sup>16</sup>. Satisfied with the robustness of the binding affinity model developed using the training set, we applied the LIE model to the quinazoline analogues comprising the test set. The test set includes 12 compounds mentioned below **Table 9**. The analogues comprising the test set were also obtained from different sources<sup>12, 19</sup>. Since, the experimental values of IC<sub>50</sub> for these inhibitors are already available; this set of molecules provides an excellent data set for testing the prediction power of the SGB-LIE method for new ligands<sup>12, 19</sup>.

**Table 9** presents the free energy values estimated for the 12 test compounds for which experimental IC<sub>50</sub> values were available to enable the accuracy check. The free energy values were estimated based on optimized SGB-LIE parameters  $\alpha$ ,  $\beta$ , and  $\gamma$  from the training set. The overall RMSE between the experimental and predicted free energy of binding values was 0.02  $\mu$ M which is comparable to the level of accuracy achieved by the most accurate methods such as free energy perturbation.

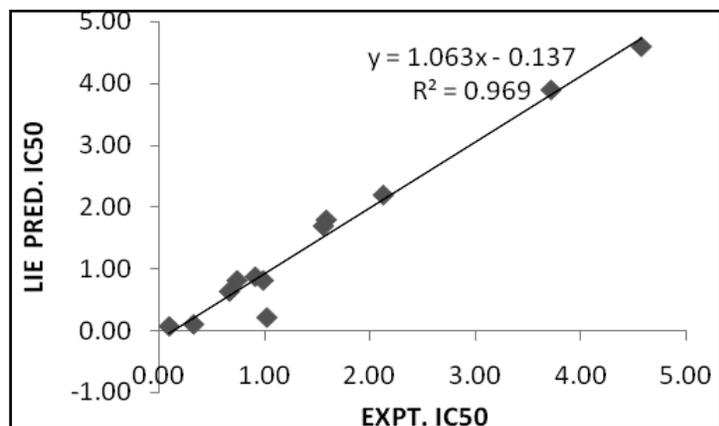
The squared correlation coefficient between experimental and SGB-LIE estimates for the free energy for the test set compounds and regression coefficient between Expt. IC<sub>50</sub> and Pred. IC<sub>50</sub> also significant ( $R^2 = 0.9693$ ). The estimated free energy values for the test set ligands are plotted against the experimental data in **Figure 4**. There is a close match between the Expt. IC<sub>50</sub> and Pred. IC<sub>50</sub> free energy values of the ligands in the test set (**Table 9**).

The quality of the fit can also be judged by the value of the squared correlation coefficient ( $R^2$ ), which was 0.9998 for the training set. Similar type of work has been studied by Alam & Naik, 2009 for applying linear

interaction energy method for binding affinity calculations of podophyllotoxin analogues with tubulin using continuum solvent model and prediction of cytotoxic activity.



**FIG. 3: FREE ENERGY VALUES ESTIMATED BY THE SGB-LIE METHOD FOR 20 QUINAZOLINE ANALOGUES COMPRISING THE TRAINING SET PLOTTED AGAINST CORRESPONDING EXPERIMENTAL DATA. THE RMS ERROR IS 0.02 $\mu$ M BETWEEN THE TWO DATA SETS FOR 20 LIGANDS STUDIED HERE**



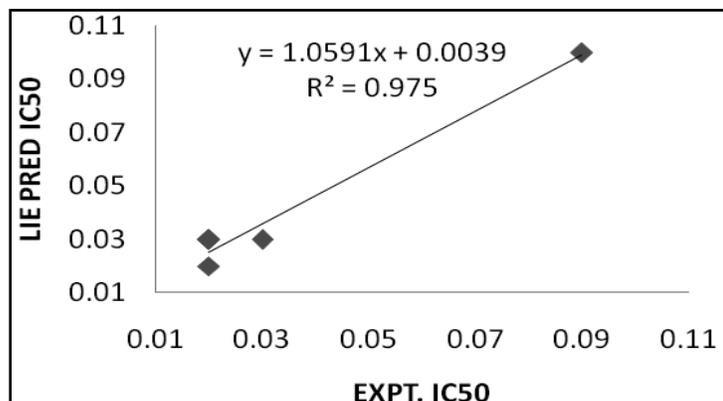
**FIG. 4: FREE ENERGY VALUES ESTIMATED BY THE SGB-LIE METHOD FOR 12 QUINAZOLINE ANALOGUES COMPRISING THE TEST SET PLOTTED AGAINST CORRESPONDING EXPERIMENTAL DATA**

To evaluate the accuracy of the SGB-LIE estimation for kinase inhibition potencies, we have taken a separate data set called as validation set consisting of 5 analogues of quinazoline (**Table 10**). Their experimental activity and chemical structures were obtained from literature<sup>12</sup>. The quality of the fit can also be judged by the value of the squared correlation coefficient ( $R^2$ ), which was 0.975 for the validation set shown in **Figure 5**.

**TABLE 10: LIE FITTING, FREE ENERGY VALUES ( $\Delta G_{\text{BIND}}$ , KCAL/MOL) AND PREDICTED POTENCIES ( $\text{IC}_{50}$ ) OBTAINED FROM THE SGB-LIE METHOD AND EXPERIMENTAL DATA FOR THE VALIDATION SET**

Analogue	Glide Score	$^1\langle U_{\text{ele}} \rangle$	$^1\langle U_{\text{vdw}} \rangle$	$^1\langle U_{\text{cav}} \rangle$	$^2\Delta G_{\text{bind LIE}}$	$^3\text{Expt. IC}_{50}$	$^4\text{Pred. IC}_{50}$
4	-10.43	7.8	-43	2.8	-0.1	0.03	0.03
12	-7.87	8.6	-44.6	3.1	-0.38	0.09	0.1
13	-8.13	7.8	-51.8	3.6	-0.08	0.02	0.02
14	-6.89	4	-43.9	3.4	-0.12	0.02	0.03
15	-7.03	11.9	-43.3	2.3	-0.12	0.02	0.03

$^1\langle U_{\text{ele}} \rangle$ ,  $^1\langle U_{\text{vdw}} \rangle$  and  $^1\langle U_{\text{cav}} \rangle$  energy terms represents the ensemble average of the energy terms calculated as the difference between bound and free state of ligands and its environment.  $^2\Delta G_{\text{bind,LIE}}$  refer to the absolute free energy values obtained using SGB-LIE method.  $^3\text{Expt. IC}_{50}$  = Experimental cytotoxic activity of ligands.  $^4\text{Pred. IC}_{50, \text{pred}}$  refers to predicted cytotoxic activity of ligands and is estimated using the relationship:  $\text{Pred. IC}_{50\text{pred}} = -(\Delta G_{\text{bind,LIE}} / 2.303 \text{ RT})$ , where 298 K is used in the work for temperature T.

**FIG. 5: FREE ENERGY VALUES ESTIMATED BY THE SGB-LIE METHOD FOR 5 QUINAZOLINE ANALOGUES COMPRISING THE VALIDATION SET PLOTTED AGAINST CORRESPONDING EXPERIMENTAL DATA**

**CONCLUSION:** We have demonstrated that the SGB-LIE method can be applied to estimate the binding free energy with a high level of accuracy for a diverse set of quinazoline analogues with kinase domain. The predicted activity value based upon free energy changes calculated by LIE approach of these analogues to kinase inhibition activity have directly correlated with the experimental potency of these inhibitors. Despite the limitation imposed by the insufficient sampling inherent in the energy minimization protocol, the method has reproduced experimental data with reasonably small error for the majority of quinazoline analogues.

The SGB-LIE predictions could produce exactly the same trend of kinase inhibition for the quinazoline analogues as experimentally activity have been synthesized. Quinazoline is well known for its antitumor activity.

However, the clinical application of it and its analogues in the treatment of cancer has been limited by severe toxic side effects during administration of the drugs.

With a view to achieving greater therapeutic efficiency many quinazoline analogues have been isolated and via molecular manipulation, a large number of synthetic derivatives have been synthesized. However, new findings related to their activities, mechanism of action and pharmacological properties have been unexplored. Most of the toxic effects of the quinazoline and its derivatives are due to their scant selectivity between cancer and normal cells.

Moreover, the SGB-LIE method is able to predict the binding free energy and cytotoxic activity of rationally designed quinazoline congeners with relative success. The difference in the exact magnitudes of predicted vs. experimental inhibition activity mode for compounds in the training set, test set and validation set may be due to the limitations imposed by inadequate sampling and force field parameterization. In practical the  $\text{IC}_{50}$  value of a drug molecule is dependent upon a number of factors including solubility, membrane permeability, p-glycoprotein activity against the compound, etc.

However, the SGB-LIE model developed is able to predict the binding energy of the validation set quite accurately in comparison to the binding kinetics in vitro. This may be the fact that kinase is the most potential target for quinazoline. Further, the strong relationship between the experimental and predicted activity could be established by in vitro studies of all these quinazoline analogues with isolated ALK5.

The close estimation of inhibition potencies of a wide range of structural derivatives for quinazoline establishes the SGB-LIE methodology as an efficient tool for screening novel compounds with very different structures. The mechanism of action of any drug is very important in drug development.

Generally, the drug compound binds with a specific target, a receptor, to mediate its effects. Therefore, suitable drug–receptor interactions are required for high activity. Understanding the nature of these interactions is very significant and theoretical calculations, in particular the SGB-LIE method, seem to be a proper tool for gaining such understanding.

The results obtain will give information on how the chemical structure of the drug should be modified to achieve suitable interactions and for the rapid prediction and virtual pre-screening of anti-tumor activity. This will lead to new proposals regarding possible improvements to the therapeutic indices of quinazoline.

Compared to the empirical methods, such as scoring function approaches, the LIE method is more accurate due to the semi empirical approach adopted in which experimental data are used to build the binding affinity model. The SGB-LIE method seems promising when compared to the FEP or TI methods in achieving comparable accuracy with must faster speed even for structurally very different ligands.

Thus, it has been the objective of numerous studies to prepare better and safer anticancer drugs. As a result of this a large number of Quinazoline analogues have been developed over the year. This provides the data set for development of computational models for virtual screening and prediction of biological activity of the new analogues development in the near future. This computational model also reduces the number of candidate molecules that needs to be synthesized and tested. This reduces both cost and time in the process of drug development.

**ACKNOWLEDGEMENTS:** We are thankful to Dr. Patrik Gomez and Dr. Janet Vaniella, for providing a very nice working atmosphere and resources.

#### REFERENCES:

1. Shaikh SA, Jain T, Sandhu G, Latha N, Jayaram B: From drug targets to leads-sketching a physiochemical pathway for lead molecule design insilico. *Curr Pharm Design* 2007; 13:3454-3470.
2. Schlessinger J: Cell Signaling by receptor tyrosine kinases. *Cell* 2000; 103:211-225.
3. Hunter T: Signaling and Beyond. *Cell* 2000; 100:113-127.
4. Roberts A, Mishra L: Role of TGF- $\beta$  in stem cells and cancer. *Oncogene* 2005; 24:5667.
5. Derynck R, Zhang Y, Feng XH: Transcriptional Activators of TGF- $\beta$  Responses: Smads. *Cell* 1998; 95:737.
6. Paul MK, Mukhopadhyay AK: Tyrosine kinase – Role and significance in Cancer. *Int J Med Sci* 2004; 1:101-115.
7. Tripti S, Shalabh S, Virendra, KS, Ashok K: Synthesis, insecticidal and antimicrobial activities of some heterocyclic derivatives of quinazolinone. *Ind J Chem* 2006; 45:2558-2565.
8. Veerachamy A, Veluchamy M, Nagendran P, Poongavanam V, Rajappan R: Synthesis, Analgesic and Anti-inflammatory Activities of Some Novel 2,3-Disubstituted Quinazolin-4 (3H)-ones. *Bio Pharm Bull* 2003; 26:557-559.
9. Guiping O, Peiquan Z, Gangfang X, Baoan S, Song Y, Linhong Jin, Wei X, Deyu H, Pin, L, Zhuo C: Synthesis and Antifungal Bioactivities of 3-Alkylquinazolin- 4-one Derivatives. *Molecules* 2006; 11:383-392
10. Ashis KN, Subarna G, Ranadhir C: Antibacterial Activity of Some 3- (Arylideneamino)- 2- phenylquinazolin- 4(3H)-ones: 169 Synthesis and Preliminary QSAR Studies. *Molecules* 2007; 12:2413-2426
11. Mani, CP, Yakaiah T, Raghu RRT, Narsaiah B, Chakra RN, Sridhar V, Venkateshwara RJ: Synthesis of novel 4,6-disubstituted quinazoline derivatives, their anti-inflammatory and anti-cancer activity (cytotoxic) against U937 leukemia cell lines. *Euro J Med Chem* 2000; 147-152.
12. Gellibert F, Fouchet MH, Nguyen V, Wang R, Krysa, Gouville A, Huet SND: Design of novel quinazoline derivatives and related analogues as potent and selective ALK5 inhibitors. *Bio Med Chem* 2009; 19:2277–2281.
13. Alam MA, Naik PK: Applying linear interaction energy method for binding affinity calculations of podophyllotoxin analogues with tubulin using continuum solvent model and prediction of cytotoxic activity. *Jour Mol Graph Mod* 2009; 27:930–943.
14. Aqvist J, Medina C, Samuelsson JE: A new method for predicting binding affinity in computer-aided drug design. *Prot Eng.* 1994; 7:385–391.
15. Carlsson, Ander M, Nervall M: Continuum solvation models in the linear interaction energy method. *J Phys Chem* 2006; 110:12034–12041.
16. Zhou R, Frienser RA, Ghosh A, Rizzo RC, Jorgensen WL, Levy RM: “New linear interaction method for binding affinity calculations using a continuum solvent model. *J Phys Chem* 2001; 105:10388–10397.
17. Schrodinger LLC., <http://www.schrodinger.com>, (accessed: 24. 04.2008).
18. Liaison, version 4.0, Schrodinger, LLC, New York, NY, 2005.
19. Lawson EC, Kinney MH, Costanzo MH, Hoekstra KH, Kauffman JA, Luci DK, Santulli R, Tounge BA, Yabut CS, Patricia, Andrade G, Maryanoff BE: Structure-Function Study of Quinazolinone-Based Vitronectin Receptor ( $\alpha\beta 3$ ) Antagonists: Computer-Assisted Analysis of Ligand-Receptor Interactions, *Lett Drug Des Disc* 2004; 114-18.
20. G.B. Thomas, R.L. Finny: In *Calculus and Analytic Geometry*, 9th ed., Addison-Wesley, 2001.

\*\*\*\*\*