Journal of
Biomedical and Pharmaceutical Sciences

Research Article Open Access

# QSAR Study and Molecular Docking of 2 Phenylaminoimidazo[4,5-H] Isoquinolin-9-Ones as Potent Inhibitors of P56$^{ick}$ Tyrosine Kinase (LCK) in Breast Cancer Therapy

**George Oche Ambrose[1]*, Olanrewaju John Afees[2,3], Sulaiman Adenike Faoziyat[1], Alakanse Suleiman Oluwaseun[1], Chia Terkuma[4,5], Kure Albert Alexander[1], Mtomga Iorwuese[6], Ayuba Abdullateef Kayode[1], Balogun Basheer Ajibola[1] and Taofeek Oluwatosin[1]**

[1]*Department of Biochemistry, University of Ilorin, Ilorin, Nigeria*

[2]*Department of Anatomy, University of Ilorin, Ilorin, Nigeria*

[3]*Ben Carson School of Medicine, Ileshan-Remo, Nigeria*

[4]*Department of Anatomy, Nile of Nigeria, Abuja, Nigeria*

[5]*Department of Anatomy, University of Nigeria, Enugu Campus, Nigeria*

[6]*Department of Anatomy, Ahmadu Bello University, Zaria, Nigeria*

## Abstract

The advances of Quantitative Structure-Activity Relationship (QSAR) studies has made the design and development of novel drugs simplified and more cost effective. QSAR in combination with molecular docking is useful in rational drug design. QSAR and Molecular docking methods were performed on 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones as inhibitors of lck. Docking studies were employed with the aid of PyRx to position the inhibitors into the lck active site to determine the optimum binding conformation and to elucidate the interactions with amino acid residues within the active site of the receptor. Based on AutoDuck vina scoring function, Compound 18 show a better binding affinity compared to the co-crystallized ligand (PBD ID: PM3). Twenty-one (21) compounds (Training dataset=14 compounds, Test dataset=7 compounds) were selected for this study. The statistical regression expressions were obtained using Multiple Linear Regression (MLR) and Partial Least Squares (PLS) with the MLR method showing more promising result than the PLS method. A QSAR model is generated by the training dataset with correlation coefficient $R^2$ of 0.79207, cross validation coefficient $Q^2$ (LOO) of 0.66644, $r^2$(correlation coefficient) for the external dataset is 0.89699 while $r^2$ of predicted dataset is 0.68432 by the Multiple Linear Regression Method.

**Keywords:** QSAR; 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones; lck; MLR; PLS

## Introduction

Lck is a member of the Src family non-receptor protein tyrosine kinase which is found to be mostly expressed in T cells and a few B cells. Recent study has shown that lck is also expressed in breast cancer tissues and cell lines [1,2]. Under-expression, Over-expression and inactive form of lck present with disrupted thymocyte development [3,4]. Therefore, lck is required for T cell receptor signaling in human jurkat T cells and for antigen receptor-dependent cytolytic effector function in the CTLL-2 T cells [5,6]. The activity of lck is regulated by phosphorylation of a highly conserved tyrosine residue, Tyr-505, which is located near the carboxyl terminus [7,8]. Previous study has shown that inducing the phosphorylation of p56$^{lck}$ stimulate lck-mediated NFkB activation which leads to the induction of urokinase type plasminogen activator (uPA) secretion that ultimately control cell motility, invasiveness and metastatic spread of breast cancer [9]. uPA, a member of serine protease plays a major role in malignant progression and tumor metastasis [9] which when up-regulated have been described in many human tumors [10]. Progress has been made recently on the development of highly potent and specific inhibitors of lck inhibitors which offers promise in finding effective drug candidates that may serve as a novel class of anticancer agents. Lck inhibitors are low molecular weight organic compounds which have been proposed to be prospective anti-proliferating agents. Based on the observation that clinical attrition rates are significantly reduced because the molecular weight falls below 500 Daltons, the recommended molecular weight is 500 Daltons [11]. Tyrosine kinase inhibitors (TKIs) are classified into three main groups; viz, type I, II and III. Type I TKIs are the most current TKIs which are referred to as ATP-competitive inhibitors. Type II and III are non-ATP competitors and act through induction of structural changes in the RTKs [12]. Of all the TKIs the most successful are Gleevec, Iressa and Tarceva [13]. The role of lck in breast cancer can be determined only when potent inhibitors of lck are developed and evaluated by clinicians. In view of this we performed a QSAR analysis to study the human p56$^{lck}$ tyrosine kinase inhibitory activity of a series of 2 phenylaminoimidazo[4,5-h] isoquinolin-9-ones. The aim of the present study is at rationalizing the substituent variations of these inhibitors to provide insight for future study. In QSAR study, some measures of physicochemical properties (descriptors) are correlated with biological activity in order to derive a mathematical model that illustrate the underlying Structural Activity Relationship (SAR). QSAR studies is greatly influential and important in modern chemistry and biochemistry. To understand SAR we need molecular descriptors that can effectively characterize molecular size, molecular shape and can influence the structure and its activities. Docking the 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones to lck followed by scoring to determine the affinity of binding and to

***Corresponding author:** George Oche Ambrose, University of Ilorin, Ilorin, Nigeria, Tel: +234840699446; E-mail: ocheab1@gmail.com

**Received:** July 12, 2018; **Accepted:** August 20, 2018; **Published:** August 27, 2018

**Citation:** Ambrose GO, Afees OJ, Faoziyat SA, Oluwaseun AS, Terkuma C, et al. (2018) QSAR Study and Molecular Docking of 2 Phenylaminoimidazo[4,5-H] Isoquinolin-9-Ones as Potent Inhibitors of P56ick Tyrosine Kinase (LCK) in Breast Cancer Therapy. J Biomed Pharm Sci 1: 108.

predict the strength of binding interaction was carried out since this is increasingly important in the drug discovery process. In this study, an attempt has been made to develop QSAR models adopting the multiple linear regression (MLR) and partial least squares (PLS) methodology. The concept of the training and test sets has been introduced for the prediction of lck inhibitory activity of structurally varied sets of compounds [14].

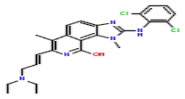
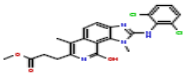
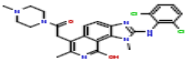
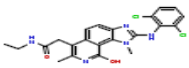
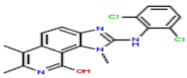
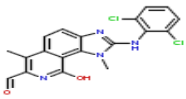
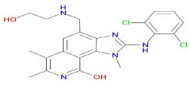
## Materials and Methods

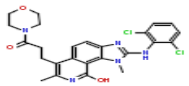
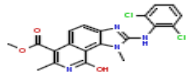
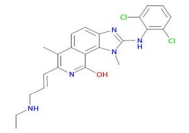
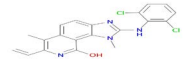
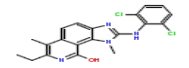
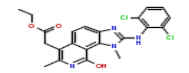
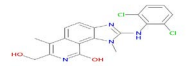
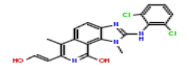
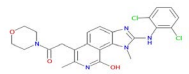
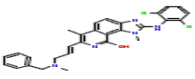
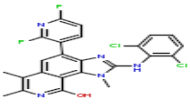
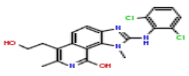
### Experimental data

For the present molecular modeling study, a set of twenty-one (21) 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones was retrieved from the CHEMBL database (http://www.ebi.ac.uk/chembl) with accession ID of CHEMBL 1136753 [15]. This dataset represent an in vitro autoimmune activity in terms of IC50 (μM) against lck. The biological activity data (IC50) were then converted to PIC50 values using the formula PIC50 = (-Log (IC50 X) (was used as the depended variable). The structures of 2 phenylaminoimidazo [4,5-h]isoquinolin-9-ones are listed in (Table 1) with their observed activities.

### Accession of chemical structures

The canonical smiles of the compounds retrieved from the CHEMBL

| S/N | COMPOUND ID | STRUCTURES | IC50 (μM) | PIC50 (μM) | NORMALIZED DATA (μM) |
|---|---|---|---|---|---|
| 1 | CHEMBL281957 |  | 0.0004 | 9.4 | 1 |
| 2 | CHEMBL284677 |  | 0.36 | 6.44 | 0.2 |
| 3 | CHEMBL29488 |  | 0.44 | 6.36 | 0.178 |
| 4 | CHEMBL26864 |  | 0.46 | 6.34 | 0.173 |
| 5 | CHEMBL27085 |  | 0.11 | 6.96 | 0.341 |
| 6 | CHEMBL284446 |  | 0.17 | 6.77 | 0.289 |
| 7 | CHEMBL27199 |  | 0.12 | 6.92 | 0.33 |

| | | | | | |
|---|---|---|---|---|---|
| 8 | CHEMBL26147 |  | 0.087 | 7.06 | 0.811 |
| 9 | CHEMBL26409 |  | 0.002 | 8.7 | 0.111 |
| 10 | CHEMBL27350 |  | 0.77 | 6.11 | 0.368 |
| 11 | CHEMBL26955 |  | 0.7 | 6.16 | 0.124 |
| 12 | CHEMBL27004 |  | 0.009 | 8.05 | 0.635 |
| 13 | CHEMBL281271 |  | 0.12 | 6.92 | 0.622 |
| 14 | CHEMBL27302 |  | 0.01 | 8 | 0.111 |
| 15 | CHEMBL27485 |  | 0.05 | 7.3 | 0.432 |
| 16 | CHEMBL281675 |  | 0.027 | 7.57 | 0.33 |
| 17 | CHEMBL26214 |  | 0.002 | 8.7 | 0.505 |
| 18 | CHEMBL26625 |  | 0.023 | 7.64 | 0.524 |
| 19 | CHEMBL27079 |  | 0.002 | 8.7 | 0.762 |

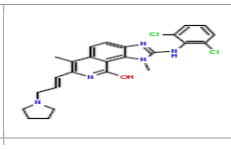
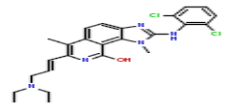
| | | | | | |
|---|---|---|---|---|---|
| 20 | CHEMBL432089 |  | 0.004 | 8.4 | 0.116 |
| 21 | CHEMBL287175 |  | 0.03 | 7.52 | 0.341 |

**Table 1**: Structures and biological activity of compounds.

database were converted to SDF files with 2D coordinates using dataWarrior software version 4.7.2. The 2D QSAR model generated in this study was derived from the training dataset of 14 molecules while the predictive potential of this model was evaluated by the testset of 7 molecules with uniformly distributed biological activities. (Table 2) shows the observed and predicted biological activities of the training and test datasets.

## Geometry optimization

After obtaining the SDF files of the compounds, their geometries were then optimized in order to make the conformations have least potential energy. Energy minimization were performed using Universal Force Field (UFF) with the optimization algorithm set at conjugate gradient. The total energy of a conformation can be calculated using the uff by the relation below:

$$E_{total} = E_B + E_A + E_{AB} + E_{OOP} + E_T + E_{VDW} + E_{ELE}$$

Where

$E_B$=Energy of bond stretching

$E_A$=Energy of angle bending

$E_{OOP}$=Out-of-plane bending energy

$E_{VDW}$=Van der waals energy

$E_{ELE}$=Electrostatic energy

$E_T$=Torsion energy term

$E_{AB}$=Energy of bond stretching and angle bending

## Descriptors generation

In order to develop a QSAR model, the activity of compounds must be quantitatively represented by molecular descriptors (9). The CDK descriptor version 1.0 was employed for the calculation of different descriptors under the following categories: Hybrid descriptors, Constitutional descriptors, Topological descriptors, Electronic descriptors and Geometric descriptors. The calculated descriptors were gathered in a data matrix. The preprocessing or pretreatment of the independent variables (i.e., descriptors) was done by removing invariable (constant column) and other descriptors based on a variance cut-off of 0.0001 and correlation coefficient cut-off of 0.99 using JFrameVWSP version 1.0. List of the physicochemical descriptors used in this study are summarized in (Table 3).

## Data normalization

Due to the existence of much variability in the range and distribution of each variable in the data set, the calculated values of the descriptors of each compound with their corresponding biological activity were subjected to a statistical technique known as min-max normalization using NormalizeTheData software version 1.0. In min-max normalization, the minimum and maximum value of each variable is adjusted to a uniform range between 0 and 1 according to the following equation:

$$x_{normalized} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

Where $x_{normalized}$ represents the min-max normalized value, $x_i$ represents the value of interest, $x_{min}$ represents the minimum value, and $x_{max}$ represents the maximum value.

## Selection of training and test set

The dataset of 21 molecules was divided into training and test set based on Kennard-Stone method [16] using the JFrameDivision software version 1.0. In this method, dissimilarity value gives an idea to handle training and test set size. This method is used for both MLR and PLS model with pIC50 activity values as dependent variable and the various 2D descriptors calculated for the molecules as independent variables.

## Model validation

Model validation is essential in QSAR modeling, it confirms the reliability of the developed QSAR model along with the acceptability of each step during model development [17]. This is done to test the internal stability and predictive ability of the QSAR models. The developed QSAR models in this study were validated by the following procedure:

**Internal validation:** Internal validation was carried out using leave-one-out (LOO) method. In the leave-one-out (LOO) method of cross validation, the process of removing a molecule, and creating and validating the model against the individual molecules is performed for all the $Q^2$ ($rCV^2$) values and reported. The $rCV^2$ (cross-validation regression coefficient) was calculated using equation (1), which describes the internal stability of a model.

$$rCV^2 = 1 - \frac{\sum\left(y_{obs} - y_{pred}\right)^2}{\sum\left(y_{obs} - \bar{y}\right)^2} \quad (1)$$

In the above equation, Y– means the average activity value of the training dataset, while Yobs and Ypred represent the observed and predicted activity values respectively. A high rCV (>0.5) suggests a reasonably robust model [18].

**Estimation of the predictive ability of a QSAR model:** After the internal validation process, the high predictive power of a QSAR model should be estimated from an external test set of compounds that are not used in building of the QSAR model. The external validation or predictive capacity of the obtained model was judged by predictive $R^2$ ($Rpred^2$) as shown in equation (2)

| Training Set | Selected Descriptors | | | Observed and Predicted values | | |
|---|---|---|---|---|---|---|
| Compounds | khs.dsCH | MDEO-11 | SC-5 | Observed | Predicted (MLR) | Predicted (PLS) |
| 2 | 0 | 0.714 | 0.296 | 0.2 | 0.21386781 | 0.215 |
| 4 | 0 | 0.714 | 0.228 | 0.173 | 0.207645118 | 0.21693 |
| 5 | 0 | 0 | 0 | 0.341 | 0.279259196 | 0.37513 |
| 6 | 0.5 | 1 | 0.296 | 0.289 | 0.384894561 | 0.33419 |
| 7 | 0 | 0.625 | 0.578 | 0.33 | 0.251201099 | 0.22589 |
| 8 | 0 | 0.833 | 0.378 | 0.811 | 0.722486046 | 0.72665 |
| 9 | 1 | 0 | 0.296 | 0.111 | 0.306346208 | 0.36672 |
| 11 | 0 | 0.714 | 0.228 | 0.124 | 0.207645118 | 0.21693 |
| 12 | 0 | 1 | 0.296 | 0.635 | 0.614594581 | 0.54965 |
| 13 | 1 | 0.833 | 0.296 | 0.622 | 0.722486046 | 0.72665 |
| 15 | 1 | 0 | 0.296 | 0.432 | 0.370769373 | 0.34669 |
| 17 | 0 | 0.714 | 0.228 | 0.505 | 0.207645118 | 0.21693 |
| 19 | 0 | 0 | 0.415 | 0.762 | 0.722486046 | 0.72665 |
| 20 | 0 | 0.714 | 0.578 | 0.116 | 0.23967368 | 0.20698 |
| **Testset** | **Selected Descriptors** | | | **Observed and Predicted values** | | |
| **Compounds** | **khs.dsCH** | **MDEO-11** | **SC-5** | **Observed** | **Predicted (MLR)** | **Predicted (PLS)** |
| 1 | 1 | 0 | 0.296 | 1 | 0.722486046 | 0.72665 |
| 3 | 0 | 0.714 | 0.578 | 0.178 | 0.23967368 | 0.20698 |
| 10 | 0 | 0 | 0.296 | 0.368 | 0.205958578 | 0.18738 |
| 14 | 0 | 0.714 | 0.578 | 0.111 | 0.176824642 | 0.15422 |
| 16 | 0 | 0 | 1 | 0.33 | 0.23967368 | 0.20698 |
| 18 | 1 | 0 | 0.296 | 0.524 | 0.317235919 | 0.36333 |
| 21 | 0 | 0 | 0.228 | 0.341 | 0.300123516 | 0.36865 |

**Table 2**: Normalized values of selected descriptors and the observed/predicted Y values (Normalized values).

| S/N | CHEMBL ID | BINDING AFFINITY (Kcal/mol) | RMSD/UB | RMSD/LB |
|---|---|---|---|---|
| 1 | CHEMBL281957 | -5.3 | 0 | 0 |
| 2 | CHEMBL284677 | -5.3 | 0 | 0 |
| 3 | CHEMBL29488 | -5.8 | 0 | 0 |
| 4 | CHEMBL26864 | -5.5 | 0 | 0 |
| 5 | CHEMBL27085 | -5.9 | 0 | 0 |
| 6 | CHEMBL284446 | -6.1 | 0 | 0 |
| 7 | CHEMBL27199 | -5.6 | 0 | 0 |
| 8 | CHEMBL26147 | -6.3 | 0 | 0 |
| 9 | CHEMBL26409 | -5.2 | 0 | 0 |
| 10 | CHEMBL27350 | -5.5 | 0 | 0 |
| 11 | CHEMBL26955 | -6.2 | 0 | 0 |
| 12 | CHEMBL27004 | -6.1 | 0 | 0 |
| 13 | CHEMBL281271 | -5.5 | 0 | 0 |
| 14 | CHEMBL27302 | -5.6 | 0 | 0 |
| 15 | CHEMBL27485 | -5.7 | 0 | 0 |
| 16 | CHEMBL281675 | -6.1 | 0 | 0 |
| 17 | CHEMBL26214 | -5.9 | 0 | 0 |
| 18 | CHEMBL26625 | -6.9 | 0 | 0 |
| 19 | CHEMBL27079 | -5.4 | 0 | 0 |
| 20 | CHEMBL432089 | -5.7 | 0 | 0 |
| 21 | CHEMBL287175 | -5.6 | 0 | 0 |

**Table 3**: 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones with their respective binding energies. Compound 18 has the highest docking score as compared with others.

$$r^2_{pred} = 1 - \frac{\sum\left(y_{pred\,(test)} - y_{(test)}\right)^2}{\sum\left(y_{(test)} - \overline{y}_{(training)}\right)^2} \quad (2)$$

Where $Y_{pred(test)}$ and $Y_{(test)}$ indicate the predicted and observed activity values, respectively, for test set compounds and Y(training) indicates the average bioactivity of compound in the training set. An acceptable predictive power of a QSAR model ($R_{pred}^2$) should be >0.6 for the test set molecules [19-21].

## QSAR model development

In this study, QSAR models were developed from the dataset using the methods MLR and PLS to screen potential leads against LCK within a training dataset set (14 compounds). The total molecular descriptors (108) was calculated for each compound using CDK algorithm.

Finally, a robust QSAR model equation was derived by MLR; irrelevant descriptors were removed through a forward stepwise method leading to a selection of three (3) 2D descriptors in the final QSAR regression equation (Table 2). The model creates a relationship in the form of a straight line (linear) equation that best approximates all the individual data points. Regression equation takes the form.

$$Y = b1x1 + b2x2 + b3x3 \text{ --------} \quad (3)$$

where Y is dependent variable, 'b's are regression coefficients for corresponding 'x's (independent variable), 'c' is a regression constant or intercept

The PLS model finds new variables or latent variables which are linear combinations of the original variables. The usefulness of PLS is obvious in cases where the data set contains highly inter-correlated descriptors (Multicollinearity) and in cases where the number of descriptors exceeds the number of observations [22]. The optimum number of PLS components (latent variables) for the study was determined based on leave one out cross validation approach. The same 108 descriptors, calculated using the CDK calculator, were selected for the PLS studies. Irrelevant descriptors were removed based on the Inter Correlation cut-off of 0.99 and Variance cut-off of 0.001 using the Genetic Algorithm v4.1 sofware.

### Ligands preparation for molecular docking

The MOL SDF format of these ligands were converted to PDBQT file using PyRx tool to generate atomic coordinates and energy was minimized by optimization using the optimization algorithm at force field set at uff (required) on PyRx.

### Accession and preparation of the target protein

The receptor LCK was prepared by retrieving the three-dimension crystal structure of lck in complex with a co-crystallized ligand (PDB:1CWD) from RCSB PDB (http://www.rcsb.org/pdb/home/home.do) [23]. The protein was subsequently cleaned by removing the bound complex molecule, the non-essential water molecules and all the heteroatoms using Pymol tool and Discovery studio visualizer. The co-crystallized ligand, 2-amino-3-(4 phosphonomethyl-phenyl)-propionic acid (PDB ID: PM3), was extracted (not removed) from the active site so as to reveal the grid coordinate around the binding pocket when viewed on pymol.

### Molecular docking using PyRx

After the preparation of the receptor (lck) and ligands, molecular docking analysis was performed by PyRx, AutoDockVina option based on scoring functions. For our analysis we used the PyRx, AutoDockVina exhaustive search docking function. After the minimisation process, the grid box resolution was centered at 7.1922 × 14.2175 × -12.6496 along the x, y and z axes respectively at grid dimension of 25 x 25 x 25 Å to define the binding site [24]. The co-crystallized ligand which serves as the standard was first docked within the binding site of lck and the resulting interaction was compared with that of 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones into the similar active sites using the same grid box dimension.

## Results and discussion

### QSAR study

**Multiple linear regression:** Based on the inter-correlation coefficients of the descriptors, highly correlated descriptors were removed from the study by a stepwise MLR method setting a correlation regression cut-off of 0.99. According to the rule of thumb in MLR (ratio of sample size to the number of descriptors should be greater than or equal to 5), a tetra-parametric model can be expected with the current training set of 14 compounds. This can be shown below.

pIC50=0.27926( +/- 0.09801)+0.41614( +/- 0.0876)khs.dsCH-0.12952( +/- 0.10151)MDEO-11+0.09151( +/- 0.16851) SC-5

n=14, $R^2$=0.74836, $R^2_a$=0.75537, F=40.8511, p=0.00001, $q^2$=0.6240, $r^2_{pred}$=0.89699 , t=6.39149, PRESS :0.19314

The above equation indicates that the model obtained with MLR showed good squared correlation coefficient ($R^2$) value and good internal predictive power ($rCV^2$) with an excellent external predictive power ($r^2_{pred}$).The scatter plot which is plotted between observed and predicted pIC50 values for training set and test set are shown in the (Figure1 a and Figure 1b) respectively.

**Partial least square regression:** The same training set, as used in MLR, was used to build the PLS model. The PLS regression was initially started with 108 descriptors. The descriptors with negligible regression coefficients were removed from the study until there was no improvement in $rCV^2$. The number of optimum components and descriptors for PLS model was found to be 3.The following model equation was obtained by PLS regression analysis:

pIC50=0.37513+0.35994(khs.dsCH)-0.21249 (MDEO-11)-0.02844 (SC-5) n=14, $R^2$=0.7194, $q^2$=0.54225, $r^2_{pred}$=0.70786

The scatter plot which is plotted between observed and predicted values for training set and test set are shown in the (Figure 2a and Figure 2b), respectively. The (Table 2) represents the observed and predicted values for both MLR and PLS models. The derived QSAR equation fitted with MLR presents a significant relationship between $pIC_{50}$ values (dependent variable) and the selected descriptors (independent

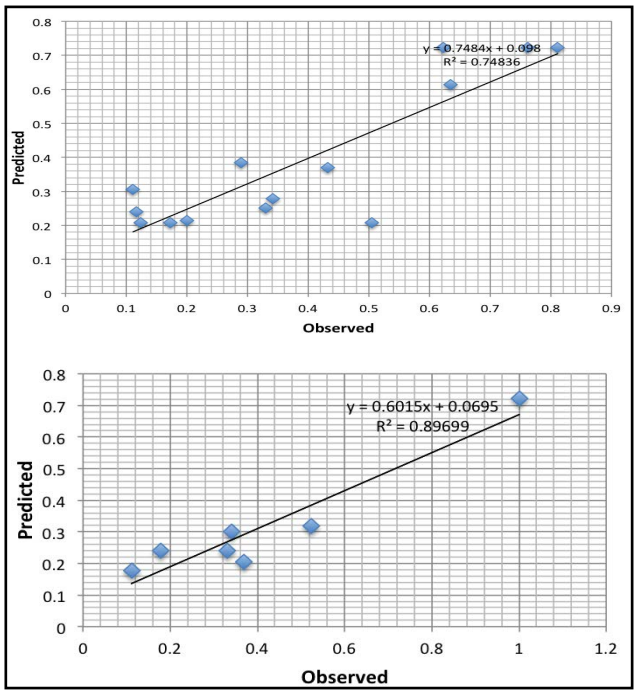



**Figure 1**: MLR analysis showing the correlation between observed and predicted pIC50 values for the (a) Training set and (b) Test set.

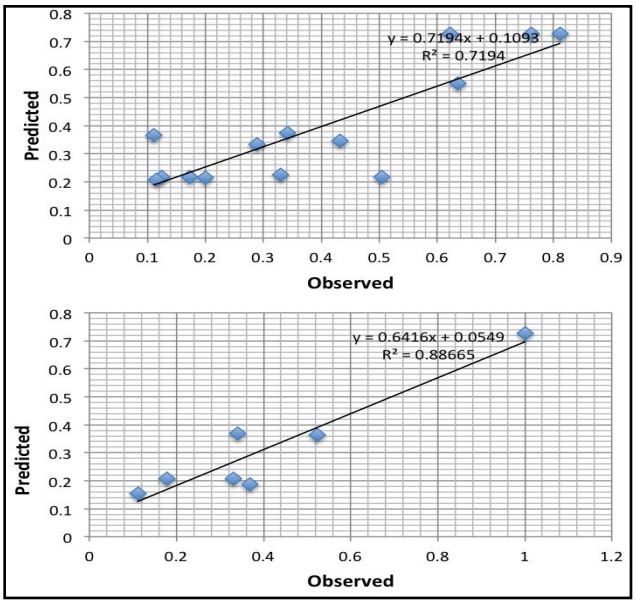



**Figure 2**: PLS analysis showing the correlation between observed and predicted pIC50 values for the (a) Training set and (b) Test set.

variables). The value of the regression coefficient (R2=0.74836) indicates the existence of ~75% correlation between the activity and the selected descriptors in the training dataset, while the value of the cross-validation regression coefficient (q2=0.6240) suggests ~62% prediction accuracy of this QSAR model. This QSAR model fitted with MLR can be use to predict future observations. $R_{pred}^2$=0.89699, shows the predictive power of the model.

## Molecular docking

In the present study, twenty-one (21) 2 phenylaminoimidazo[4,5-h] isoquinolin-9-ones were docked into the binding pocket of lck for their lck inhibitory (antagonistic) properties. All the compounds of 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones showed a better binding affinity when compared with the co-crystallized ligand (PDB ID: PM3) and compound 18 was discovered as the lead compound with the highest binding energy of -6.9 kcal/mol (Table 4). The drug-likeness of compound 18 was assessed by subjecting it to the Lipinski's rule of five, afterwards the lead compound violated none of the rules, this describes its bioavailability and binding potential (Table 5).

Compound 18, the lead compound has a binding energy of -6.9 kcal/mol, while the standard compound has binding energy of -3.7 kcal/mol (Table 4). The highest binding energy (-6.9kcal/mol) attributed to compound 18 in this regard is believed to be as a result of its chemical interactions at the receptor's active site (Figure 3) which includes:

Four (4) Hydrogen bonds involving VAL8 and GLU6 residues; (Figure 3a)

Two (2) Hydrophobic interactions involving VAL8 residue; (Figure 4a)

Two (2) Electrostatic bonding involving GLU2 and GLU6 residues

Four (2) Halogen bonds involving GLU2 residue

While that of the co-crystallized ligand (PDB Ligand ID: PM3) which serves as the standard presents with the following chemical interactions at the binding pocket (Figure 4)

Four (4) Hydrogen bonds involving GLU6 residue; (Figure 3b)

Two (2) Electrostatic interaction involving GLU2 and GLU6 residues; (Figure 4b)

The highest binding energy (-6.9 kcal/mol) attributed to compound 18 in this regard is believed to be as a result of the high number of chemical interactions (13) of compound 18 as against that of the co-crystallized ligand (6) (Figure 5).

| S/N | Ligand | Binding energy (kcal/mol) | RMSD/UBa | RMSD/LBb |
|---|---|---|---|---|
| 1 | CHEMBL26625 | -6.9 | 0 | 0 |
| 2 | PM3 | -3.7 | 0 | 0 |
| RMSD/UB: Root mean square deviation/upper bond; RMSD/LB: Root mean square deviation/lower bond | | | | |

**Table 4**: Docking scores and RMSD values of compound 18 and the co-crystalized ligand.

| Molecular Properties | Lipinski's rule of Five | Compound 18 drug-like properties |
|---|---|---|
| Molecular Mass | <500 | 483.000000 |
| Hydrogen bond Acceptor | <10 | 6 |
| Hydrogen bond Donor | <5 | 1 |
| LogP | <5 | -0.473900 |
| Molar Refractivity | Between 40-130 | 90.308289 |

**Table 5**: Lipinski's drug-like properties of compound 18: The rule describes drug candidate's pharmacokinetics in the human body which also including their absorption, distribution, metabolism, and excretion ("ADME") using an online server (http://www.scfbio-iitd.res.in/).

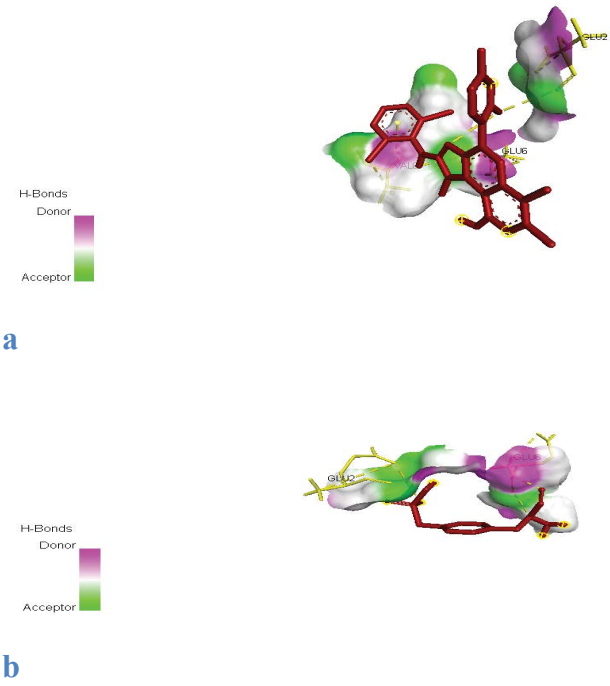



**Figure 3**: 3D interactions of ligands (Red sticks) within the binding pocket (a) compound 18 (b) PM3 (obtained by Discovery Studio visualizer software v 4.1).

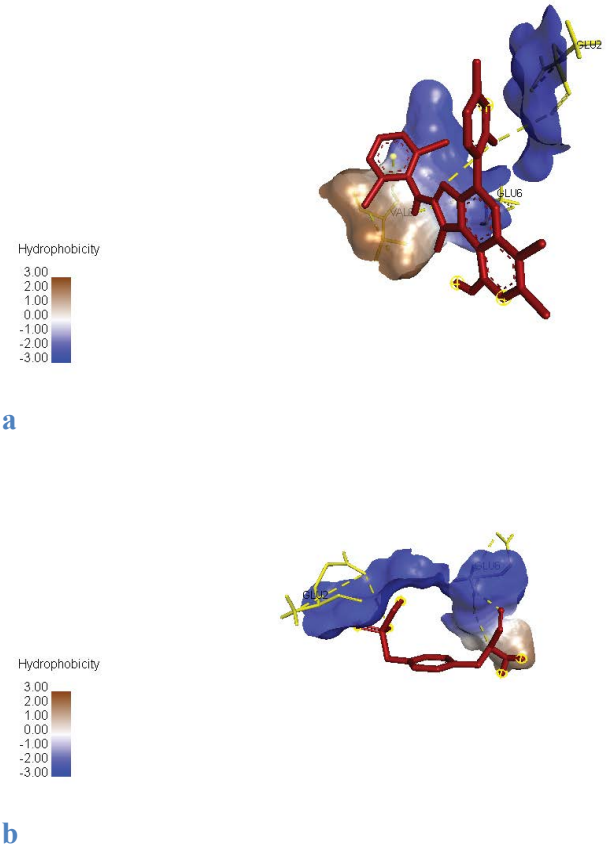



**Figure 4**: 2D interactions of ligands within the binding pocket (a) Compound 18 (b) PM3 (obtained by Discovery Studio visualizer software v 4.1).

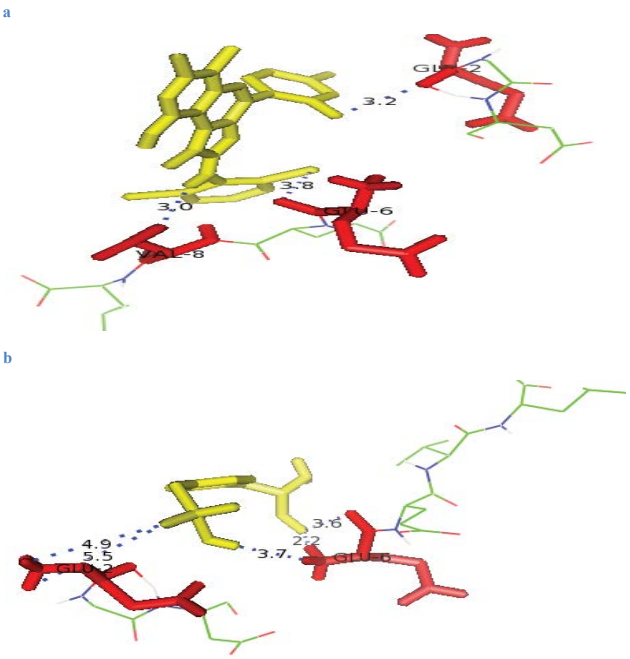



**Figure 5**: Measured bond distances between the ligands (yellow sticks) and amino acids residues (red sticks) (a) Compound 18 (b) PM3 (obtained by Pymol Molecular graphic software).

## Conclusion

In this paper, we have employed the QSAR and docking methodology to examine the structure-activity relationship of a series of 2 phenylaminoimidazo[4,5-h]isoquinolin-9-ones in order to evaluate lck inhibitors. The MLR and PLS were used to develop statistically significant models which is further validated by a cross validation method utilizing the LOO procedures. The models show good predictive potentials for lck inhibitors which can be use to predict new lck inhibitors. These QSAR models could provide a reliable tool for the design of lck inhibitors. Molecular docking studies of compound 18 with lck elucidate the relevance of hydrophobic interactions and hydrogen bonding to binding affinity. Furthermore, the Lipinski rule of five test show that compound 18 have the potential to serve as drugs against this target.

### References

1. Koster A, Landgraf S, Leipold A, Sachse R, Gebhart E, et al. (1991) Expression of oncogenes in human breast cancer specimens. Anti Cancer Res 11: 193-201.
2. Amundadottir LT, Leder P (1998) Signal transduction pathways activated and required for mammary carcinogenesis in response to specific oncogenes. Oncogene 16: 737-746.
3. Molina TJ, Kishihara K, Siderovski DP, VanEwijk W, Narendran A, et al. (1992) Profound block in thymocyte development in mice lacking p56lck. Nature 357: 161-164.
4. Allen JM, Forbush KA, Purlmutter RM (1992) Functional dissection of the lck proximal promoter. Mol Cell Biol 12: 2758-2768.
5. Straus DB, Weiss A (1992) Genetic evidence for the involvement of the lck tyrosine kinase in signal transduction through the T cell antigen receptor. Cell 70: 585-593.
6. Karnitz L, Sutor SL, Torigoe T, Reed JC, Bell MP, et al. (1992) Effects of p56lck deficiency on the growth and cytolytic effector function of an interleukin-2-dependent cytotoxic T-cell line. Mol Cell Biol 12: 4521-4530.
7. Amrein KE, Sefton BM (1988) Mutation of a site of tyrosine phosphorylation in the lymphocyte-specific tyrosine protein kinase, p56lck, reveals its oncogenic potential in fibroblasts. Proc Natl Acad Sci 85: 4247-4251.
8. Marth JD, Cooper JA, King CS, Ziegler SF, Tinker DA, et al. (1988) Neoplastic transformation induced by an activated lymphocyte-specific protein tyrosine kinase (pp56lck). Mol Cell Biol 8: 540-550.
9. Ganapati H, Mahabeleshwar, Kundu GC (2004) Tyrosine Kinase, p56lck Regulates Cell Motility and Nuclear Factor κBmediated Secretion of Urokinase Type Plasminogen Activator through Tyrosine Phosphorylation of IκBα Following Hypoxia/Reoxygenation. J Biol Chem 278: 52598-52612.
10. Andreasen PA, Egelund R, Petersen HH ( 2000) The plasminogen activation system in tumor growth, invasion, and metastasis. Cell Mol Life Sci 57: 25-40.
11. Garuti L, Roberti M, Bottegoni G (2010) Non-ATP competitive protein kinase inhibitors. Cuur Med Chem 17: 2804-2821.
12. Reuning U, Magdolen V, Wilhelm O, Fischer K, Lutz V, et al. (1998) Multifunctional potential of the plasminogen activation system in tumor invasion and metastasis. Int J Onco 13: 893-906.
13. Nandi S, Bagchi MC (2010) 3D-QSAR and molecular docking studies of 4-anilinoquinazoline derivatives: a rational approach to anticancer drug design. Mol Divers 14:27-38.
14. https://www.ebi.ac.uk/chembl/compound/inspect/CHEMBL1136753
15. Agus S, Tade MO, Hari V (2012) A Modified Kennard Stone Algorithm for Optimal Division of Data for Developing Artificial Neural Network Models. Chemical Product and Process Modeling 7.
16. Paul MK, Mukhopadhyay AK (2004) Tyrosine Kinase-Role and Significance in cancer. I J Med Sci 1: 101-115.
17. Roy K, Das RN, Ambure P, Aher RB (2016) Be aware of error measures. Further studies on validation of predictive QSAR models. Chemom Intell Lab Sys 152: 18-33.

18. Veber DF, Johnson SR. Cheng HY, Smith BR, Ward KW, et al. (2002) Molecular properties that influence the oral bioavailability of drug candidate. J Med Chem 45: 2615-2623.

19. Yadav DK, Khan F (2013) QSAR, docking and ADMET studies of camptothecin derivatives as inhibitors of DNA topoisomerase-I. J Chemom 27: 21-33.

20. Yadav DK, Kumar S, Rajput S, Singh H, Kim MH, et al. (2017) Molecular docking, QSARand ADMET studies of withanolide analogs against breast cancer. Drug Des Devel Ther 11: 1859-1870.

21. Yadav DK, Kalani K, Khan F, Srivastava SK (2013) QSAR and docking based semi-synthesis and in vitro evaluation of 18β-glycyrrhetinic acid derivatives against human lung cancer cell line A-549. Med Chem 9: 1073-1084.

22. Cramer RD (1993) Perspectives in Drug Discovery and Design. 1: 269.

23. Yadav DK, Mudgal V, Agrawal J, Tul ST, Maurya AK et al. (2013) Molecular docking and ADME studies of natural compounds of Agarwood oil for topical anti-inflammatory activity. Curr Comput Aided Drug Des 9: 360-370.

24. http://www.rcsb.org/pdb/home/home.do