

## Topical Perspectives

# Atom and receptor based 3D QSAR models for generating new conformations from pyrazolopyrimidine as IL-2 inducible tyrosine kinase inhibitors



Zaheer Ul-Haq<sup>a,\*</sup>, Juweria Shahrukh Effendi<sup>a</sup>, Sajda Ashraf<sup>a</sup>, Majdi M. Bkhaitan<sup>b</sup>

<sup>a</sup> Dr. Panjwani Canter for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan

<sup>b</sup> Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Umm Al-Qura University, Makkah, KSA, Saudi Arabia

## ARTICLE INFO

## Article history:

Received 8 April 2016

Received in revised form 24 March 2017

Accepted 24 March 2017

Available online 27 March 2017

## Keywords:

IL-2 Itk

Molecular docking

3D-QSAR CoMFA and CoMSIA

## ABSTRACT

In the current study, quantitative three-dimensional structure-activity-relationship (3D-QSAR) method was performed to design a model for new chemical entities by utilizing pyrazolopyrimidines. Their inhibiting activity on receptor IL-2 Itk correlates descriptors based on topology and hydrophobicity. The best model developed by ligand-based (atom-based) approach has correlation-coefficient of  $r^2$ : 0.987 and cross-validated squared correlation-coefficient of  $q^2$ : 0.541 with an external prediction capability of  $r^2$ : 0.944. Whereas the best selected model developed by structured-based (receptor-based) approach has correlation-coefficient of  $r^2$ : 0.987, cross-validated squared correlation-coefficient of  $q^2$ : 0.637 with an external predictive ability of  $r^2$ : 0.941. The statistical parameters prove that structure-based gave a better model to design new chemical scaffolds. The results achieved indicated that hydrophobicity at  $R_1$  location play a vital role in the inhibitory activity and introduction of appropriately bulky and strongly hydrophobic-groups at position 3 of the terminal phenyl-group which is highly significant to enhance the activity. Six new pyrazolopyrimidine derivatives were designed. Docking simulation study was carried out and their inhibitory activity was predicted by the best structure based model with predictive activity of ranging from 8.43 to 8.85 log unit. The interacting residues PHE435, ASP500, LYS391, GLU436, MET438, CYS442, ILE369, VAL377 of PDB 4HCT were studied with respect to type of bonding with the new compounds. This study was aimed to search out more potent inhibitors of IL-2 Itk.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

T-cell cancers include a group of disorders: acute lymphoblastic leukemia (T-ALL), Adult T-cell leukemia lymphoma and Sezary syndrome (a cutaneous T-cell lymphoma). Despite, that there are obvious variances between these carcinomas, management has some overlap especially at their terminal stages. Poor prognosis is due to the relapse of the disease. It is reported that 42% of T-ALL patients suffer with complications like relapse within 5 years [1]. Simultaneously the survival rate of Sezary syndrome is only 11% [2]. Tyrosine kinases are the central proteins of signal transmission

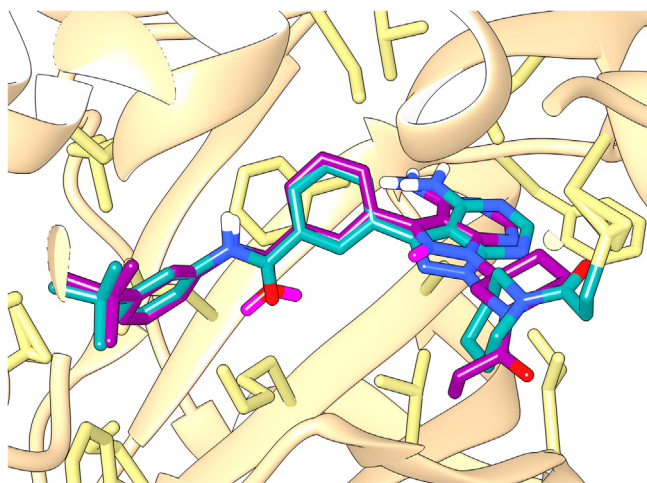
that originate from the cell surface [3]. IL-2 Itk is a non-receptor tyrosine-kinase and a member of Tec kinase family [4].

It's structure comprise of amino end which contains the Pleskstrin homology (PH) domain then the Zn<sup>++</sup> binding region called as the Btk homology (BH) and then a proline rich region (PRR) that follows the sequence of the SH3 ligand. As an antigen is recognized by the T-cell receptor (TCR) at major histocompatibility complex (MHC) a set of phosphorylation cascade inside the T-cell gets initiated. Downstream activation causes calcium flux which in turn results in activation of calcineurin and NFATc and simultaneously induces the expression of cytokines [5]. It is reported that Itk also signal through the G protein coupled receptor (GPCR) through CXCL14/SDF1 furthermore it induce actin polymerization and cellular migration [6]. Itk has proven to play a dominant role in T cell malignancies [7]. IL-2 Itk also has an aberrant role in inflammatory disorders like lung inflammation, skin dermatitis, and asthma and HIV infections [8,9].

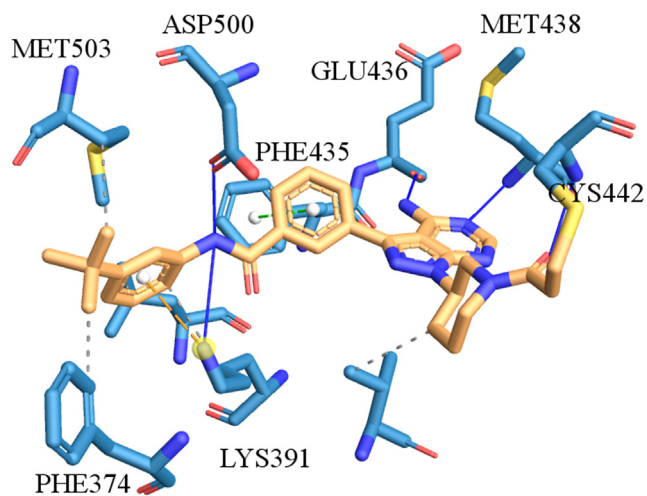
**Abbreviations:** 3D-QSAR, 3-dimensional quantitative structure-activity relationship; CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity index analysis; PLS, partial least square; IL-2 Itk, interleukin 2 inducible tyrosine kinase.

\* Corresponding author.

E-mail address: [zaheer.qasmi@iccs.edu](mailto:zaheer.qasmi@iccs.edu) (Z. Ul-Haq).



**Fig. 1.** Redocking pose of 4HCT with 18R compound as reference by GOLD software (3.0) with RMSD value of 1.5 Å.



**Fig. 2.** 3D representation of docking interactions between the 4HCT and the reference 18R compound showing the residues and the water molecules interacting in the active site within 10 Å.

Itk inhibitors reported in literature are 2 amino 5 [(thiomethyl)aryl] thiazoles [10], 2 amino 5 (thioaryl) thiazoles [11], 4/5 arylpyrazolyl indoles [12], benzimidazole derivatives [13–17] and pyrazolopyrimidines [18].

This detailed information of chemical features that can positively enhance the biological activity of inhibitor is considered to be important to design new potent compounds. The molecular modeling by 3D QSAR technique establishes a potent model. It explores new chemical scaffolds and studies their structure activity relationship (SAR). The structure activity prediction gives the information of the influencing molecular features contributing high biological activity. 3D QSAR modeling can be carried out by two approaches the ligand based (atom-based) and structure based (receptor-based) approaches [19]. Both approaches created CoMFA and CoMSIA models and determined correlations between biological activity and 3D structure using energy fields of Steric, Electrostatic, Hydrophobic, Electron donor and acceptor fields.

Current study modeling was performed by taking the reported pyrazolopyrimidines (Table 1). The aim was to design new compounds that can have better inhibitory activity on IL-2 Itk.

## 2. Methodology

### 2.1. Molecular dataset

The reported pyrazolopyrimidine derivatives and their experimental activities were included in this study [18]. A 44 compounds dataset was randomly distributed into two subsets the training (35 compounds) and test (7 compounds) based on the following rules: (a) to evaluate the significance of statistical model compound diversification was essential. (b) In-order to evade redundancy or biasness in activity range and structural features, the given information of the compounds must be precise. (c) The training dataset activity should range from the most potent to the least potent compound. The biological activities reported were in  $IC_{50}$ , which were then converted to  $pIC_{50}$  by applying formula of  $pIC_{50} = \text{Log}[1/IC_{50}]$ . The activity ranges from 4.19 to 8.7 log units (Table 1). Thus the compounds provide a comprehensive data set for 3D-QSAR study. Generally, the range of biological activity must cover at a minimum of 3 log units for a robust 3D QSAR model [20].

### 2.2. Ligand and receptor preparation

The 2D structures of the ligands were sketched by ChemDraw Ultra 8.0 software [21] which was further converted into 3D structures by using QUACPAC OpenEye Babel [22]. The ligands were charged AM1 by QUACPAC OpenEye molcharge program [23] followed by further correction by sketch module of Sybyl 7.3 [24] and the geometry optimization of the ligands was done by energy minimization by QUACPAC OpenEye Omega v2.5.1.4 [25]. In this study, the reported PDB used of IL-2 Itk was 4HCT (Resolution: 1.48 Å), which is resolved by X-ray crystallography [18] and retrieved from RCSB Protein Data Bank.

### 2.3. Molecular alignment

This is a step fundamental for comparison and co-relation of diverse structures in a dataset. Therefore, the alignment-based 3D QSAR models are considered to be sensitive to the different conformations of the molecules in the data set which are super-imposed over each other. To develop the optimal 3D QSAR models, two different methods of alignment were employed. The first one is an atom-based or ligand-based alignment, i.e., all ligands were then aligned to compound 41 (most potent compound of the series) by the Align Database command in SYBYL 7.3 software. The second one is a receptor-based alignment, which comprise dataset of optimal conformations of the ligands resulting from docking studies. The conformations were charged (AM1 charges) and imported for molecular alignment for 3D QSAR analysis [26].

### 2.4. Re-docking protocol

Prior to docking re-docking of co-crystallized ligand protocol was established. For this purpose, the protein 4HCT was retrieved from Protein Data Bank and the cognate ligand was removed. The water molecules found in the active site in interaction with the amino acids of the receptor protein were not removed while the rest were removed. This was followed by addition of polar hydrogens, software validation by re-docking and RMSD calculations. The GOLD software [27] resulted into better RMSD which was 1.5 Å. The efficacy of algorithm of docking and scoring function was assessed by evaluating the selected re-docked pose by visualizing and comparing its interactions with the receptor protein as shown in Fig. 1.

**Table 1**  
Structure of the Pyrazolopyrimidine compounds dataset with their reported biological activity.

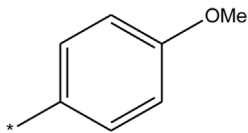
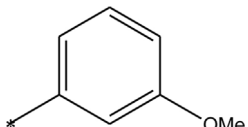
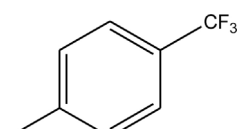
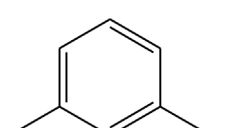
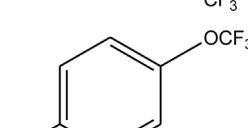
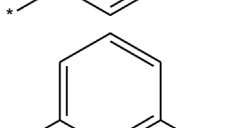
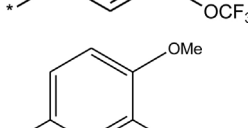
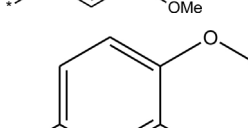
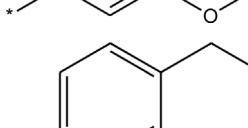
Compound Number	Data Set Compounds Substituents R <sub>1</sub>	IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> <sup>b</sup> μM
10		4.425	5.35
11		2.495	5.60
12		52.458	4.28
13		17.744	4.75
14		6.645	4.67
15		6.645	5.18
16		21.321	4.67
17		8.341	5.08
18		1.045	5.98

Table 1 (Continued)

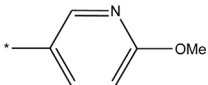
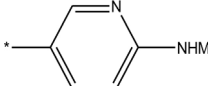
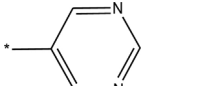
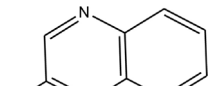
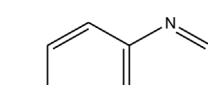
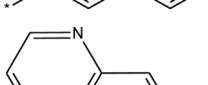
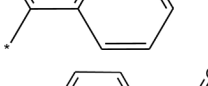
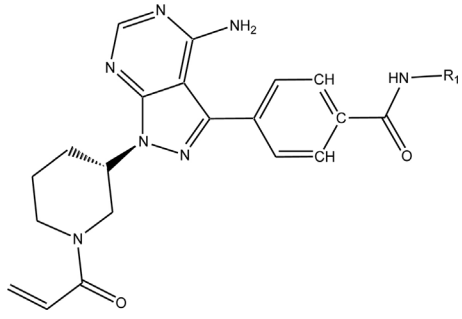
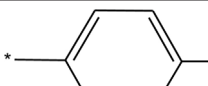
Compound Number	Data Set Compounds Substituents R <sub>1</sub>	IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> <sup>b</sup> μM
19		22.930	4.64
20		10.324	4.99
21		10.068	5.00
22		9.827	5.01
23		7.457	5.13
24		20.439	4.69
25		21.883	4.66
			
Compound Number	Substituents R <sub>1</sub>	IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> μM
30		0.13	6.89

Table 1 (Continued)

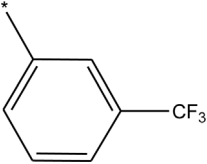
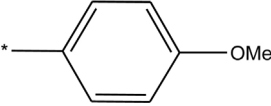
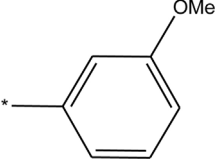
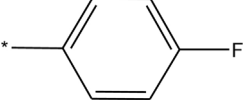
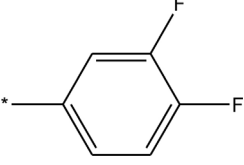
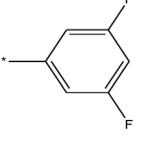
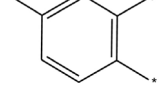
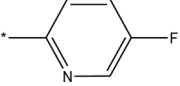
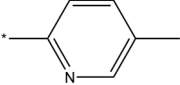
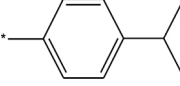
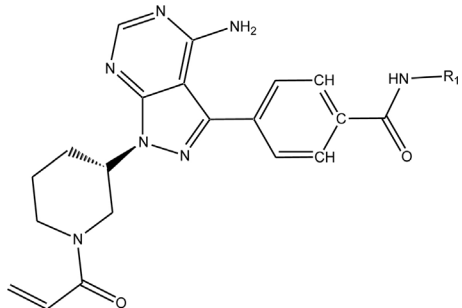
Compound Number	Substituents R <sub>1</sub>	IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> μM
31		0.19	6.72
32		0.18	6.74
33		0.20	6.70
34		0.72	6.14
35		1.41	5.85
36		0.21	6.68
37		0.52	6.28
38		1.02	5.99
39		0.72	6.14
40		0.01	8.00

Table 1 (Continued)



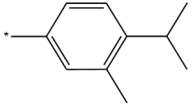
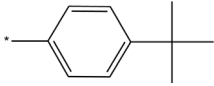
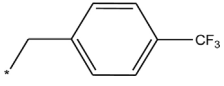
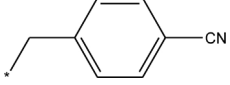
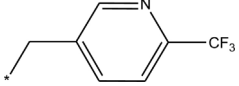
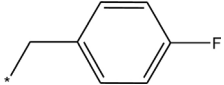
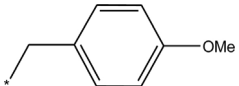
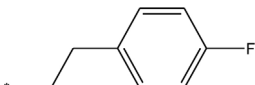
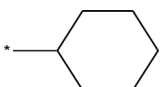
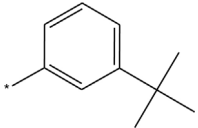
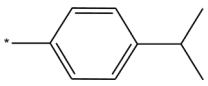
Compound Number	Substituents R <sub>1</sub>	IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> μM
41		0.002	8.70
42		0.01	8.00
43		0.20	6.70
44		4.08	5.39
45		1.52	5.82
46		2.46	5.61
47		6.47	5.19
48		1.75	5.76
49		10.88	4.96
51	-CH <sub>2</sub> CF <sub>3</sub>	36.22	4.44
52		0.073	7.15
53		0.78	6.107

Table 1 (Continued)

Compound Number		IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> μM	
58		65	4.19	
59		7.57	5.12	
61		0.1	7.00	
62		0.01	8.00	
Compound Number	R <sub>1</sub>	R <sub>2</sub>	IC <sub>50</sub> <sup>a</sup> μM	pIC <sub>50</sub> μM
65	Ph	–CH=CH <sub>2</sub>	8.80	5.06
67	4-MePh	–CH=CH <sub>2</sub>	1.68	5.77
68	4-MePh	Me	0.04	7.40

### 2.5. Molecular docking simulations

In this study, we used GOLD Dock suit for docking study which has a patented search engine and a Gold fitness scoring-function. For the docking simulation study the crystal structure of complex pdb of IL-2 Itk (4HCT complexed with 18R) was used. The scoring-function (GoldScore–fitness) incorporated comprises of four constituent energies relating to protein–ligand which includes the hydrogen-bond, Van der Waals (VDW) and relating to ligand are ligand–Internal VDW and ligand–torsional strain. An active site of 10 Å was selected around reference compound with centroid of X = 3.6049, Y = 8.0158, and Z = 19.6723. Other parameters were established in default. Each molecule of the dataset was docked and terminated till achieving ten poses lying within 1.5 Å (RMSD). The scores of ten docked poses of each pyrazolopyrimidine derivatives

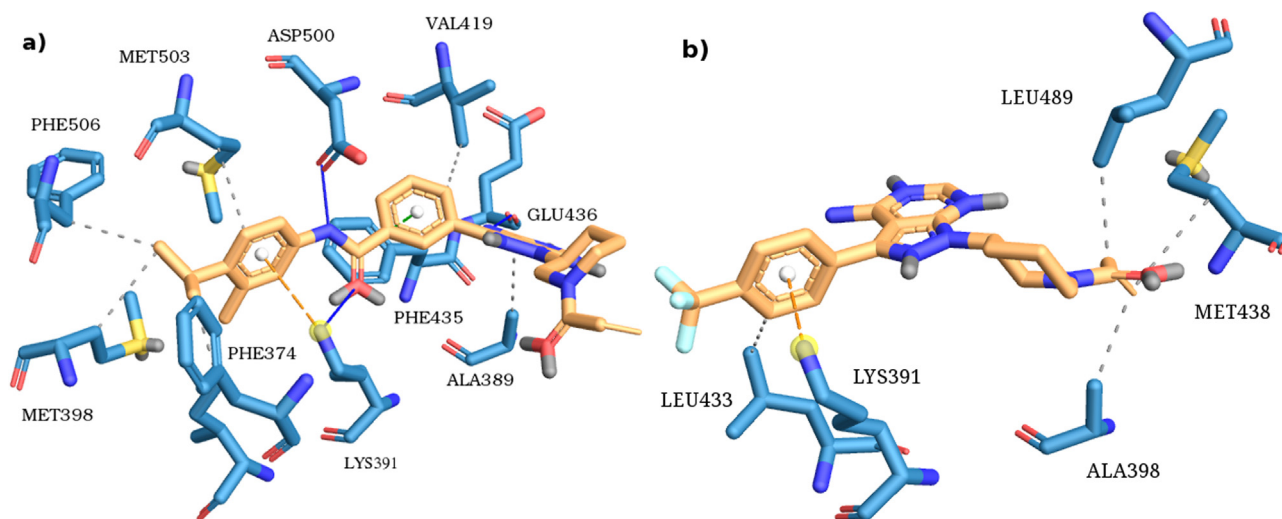
were rated in a molecular database sheet, and the topmost of the total score sheet was taken for the analysis of their protein–ligand interactions. The binding of the residues to the ligand (Fig. 2) was visualized by MOE 2015.1001 [28].

### 2.6. CoMFA and CoMSIA setup

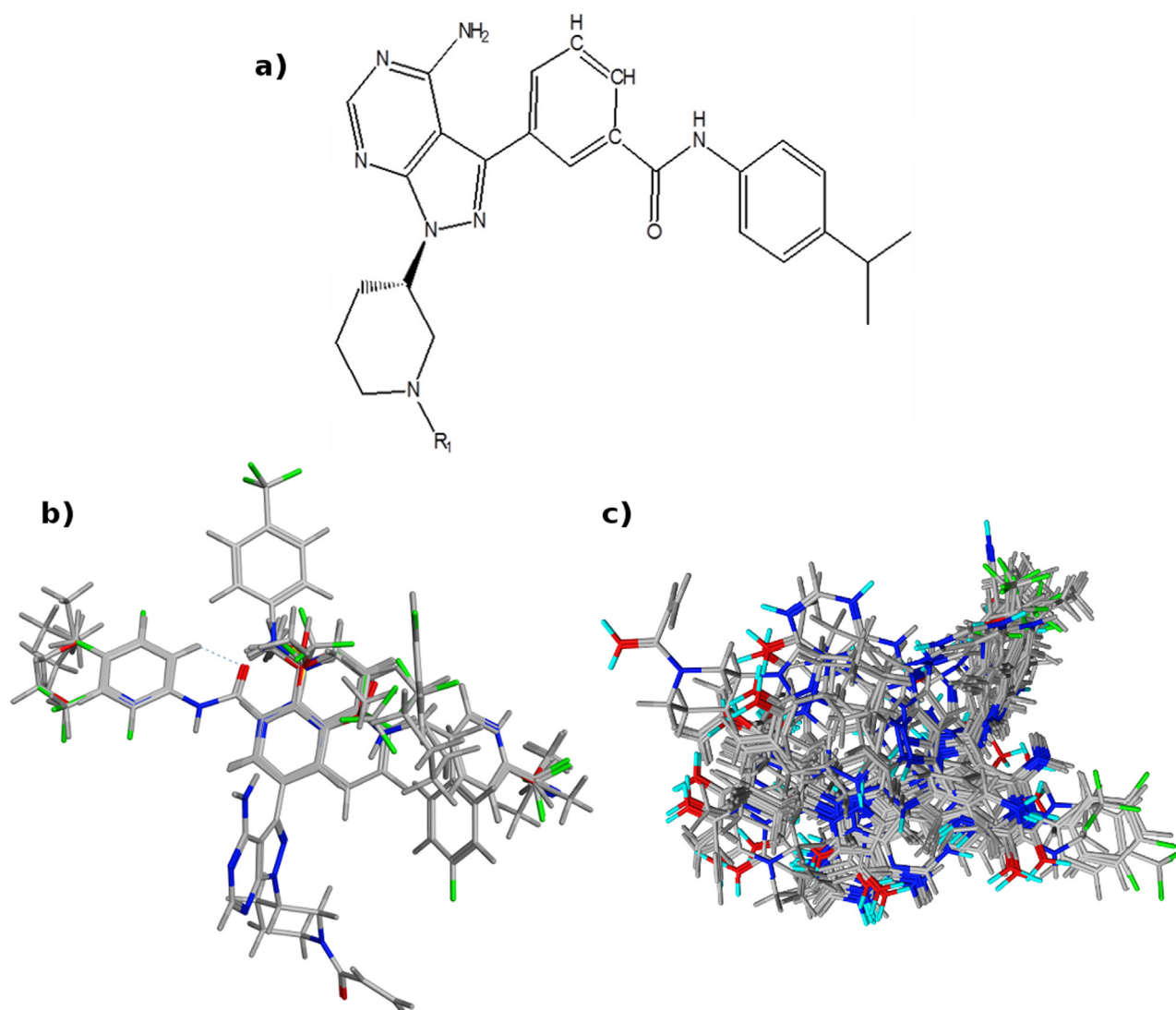
The 3-D lattice spacing in the x, y, and z directions was set at a distance of 2 Å. The energies were calculated for electrostatic and steric field by Coulomb and Lennard–Jones potentials [29]. For CoMFA technique, a sp<sup>3</sup> hybridized carbon (probe atom) with a + 1 charge was recognized to compute the magnitude of electrostatic and steric fields, which was truncated at 30 kcal/mol [30–32].

The CoMFA and CoMSIA methods were similar with reference to fields generated around the compound and relied on the postula-



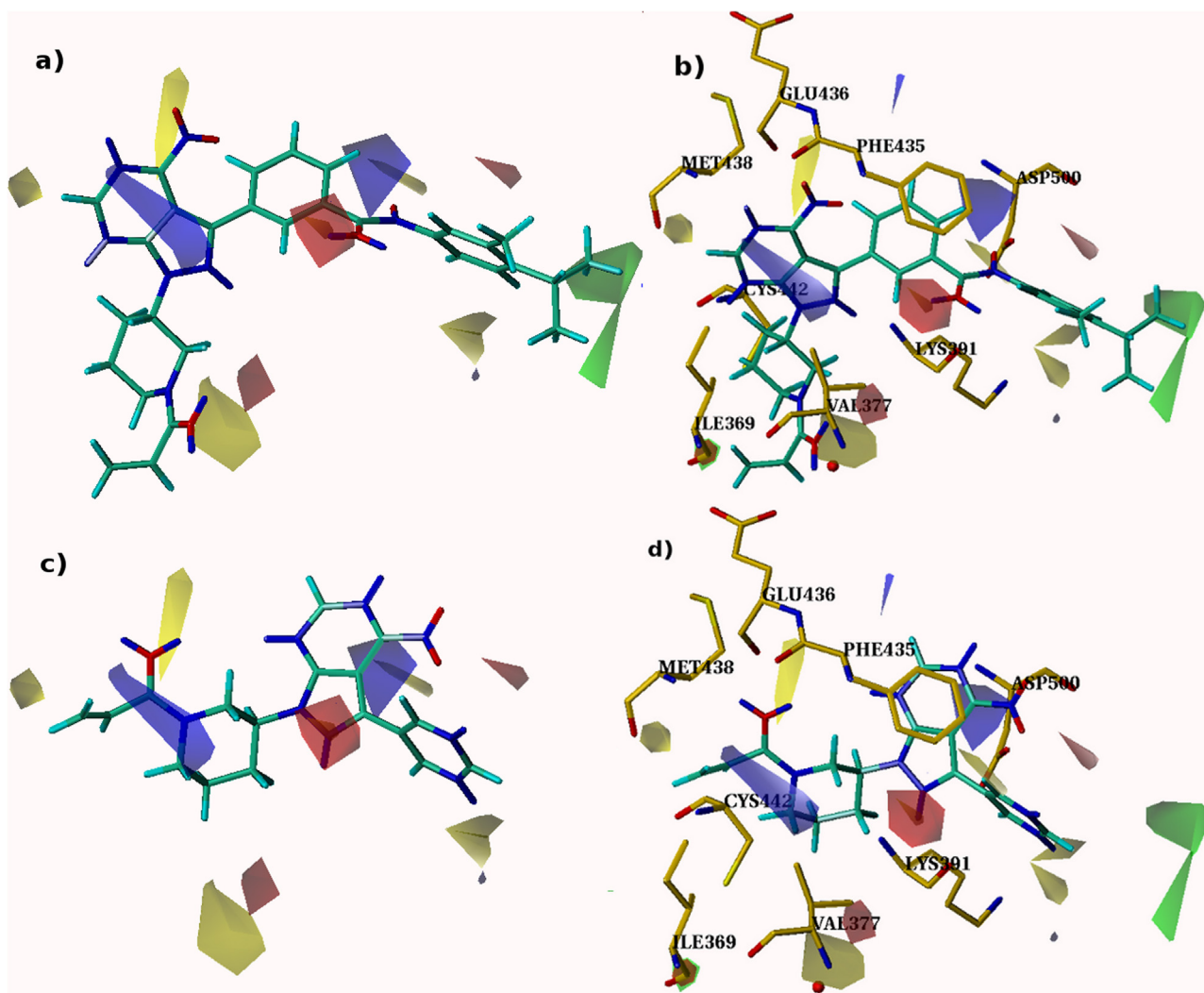


**Fig. 3.** a) 3D Representations of most active compound 41 showing three hydrogen bond interactions with ASP500, GLU436 and LYS391,  $\pi$ -stacking with the gatekeeper PHE435, pi cation interaction with LYS391 and hydrophobic interactions with residues (MET398, MET 503, ALA389, VAL419, PHE374, PHE506); b) 3D Representations of least active compound 12 interacted with hydrophobic interactions with (MET438, LEU489, ALA398) and pi cation interaction with LYS391.



**Fig. 4.** a) Core structure of Pyrazolopyrimidine derivatives used for alignment, b) Ligand-based alignment by using most active compound 41 as template; c) Receptor based alignment using most active compound 41 as template.





**Fig. 5.** CoMFA contour maps showing sterically favored areas which are represented by green isopleths while the yellow regions are served for sterically unfavorable regions; a) ligand-based CoMFA descriptors of most active (compound-41) b) demonstrates structure-based CoMFA contour maps of most active (compound-41) c) ligand-based CoMFA descriptors of least active (compound-12) d) demonstrates structure-based CoMFA contour maps of least active (compound-12). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tion that variation in the compound binding affinity is related with the modification of the molecular-property. Additional to steric (S) and electrostatic (E) fields, hydrophobic (H), hydrogen bond acceptor (HA) and hydrogen bond donor (HD) fields are also calculated [33]. Furthermore, Gaussian function was executed to calculate all grid points on the basis of similarity indices, outside and inside of different molecular surfaces. The Similarity indices were cumulated by the given equation:

$$A_{F,K(j)}^q = \sum_i W_{probe,k} W_{ik} e^{-\alpha r_{iq}^2}$$

Where, similarity index is represented by A at lattice point q, sum of total atoms “i” of the molecule “j” under examination. A probe atom:  $W_{probe,k}$  with a radius of 1 Å, carrying charge +1, hydrogen-bond donating +1 and hydrogen-bond accepting +1 and hydrophobicity +1.  $W_{ik}$  gives the value of the physicochemical property “k” of atom “i”.  $r_{iq}$  is the distance between the probe-atom at lattice point “q” and atom “i” of the test compound. Attenuation factor “ $\alpha$ ” whose optimum value is in the range of 0.2–0.4, with default value of 0.3 [34].

### 2.7. Partial least squares (PLS) analysis

The cross validation technique leave one out (LOO) was primarily used for partial least square (PLS) analysis to generate cross validated  $r^2$  value and the optimum number of component (N), with minimum standard error of prediction (SEP) which relates to better or higher cross validated ( $q^2$ ) squared coefficient. A higher count of components was included to avoid the over fitting of model. The value of  $q^2$  is a worthy indicator for actual and test set predictions, it should be >0.5 for an ideal model development [35]. For the establishment of final 3D QSAR model non-cross validation technique was executed with optimal number of component. For the establishment of final 3D QSAR model non-cross validation technique was executed with optimal number of component. The significant PLS models provided the substantial correlation coefficient ( $r^2$ ), F is the ratio between the variances of predicted and observed activities and standard errors of estimate (SEE).

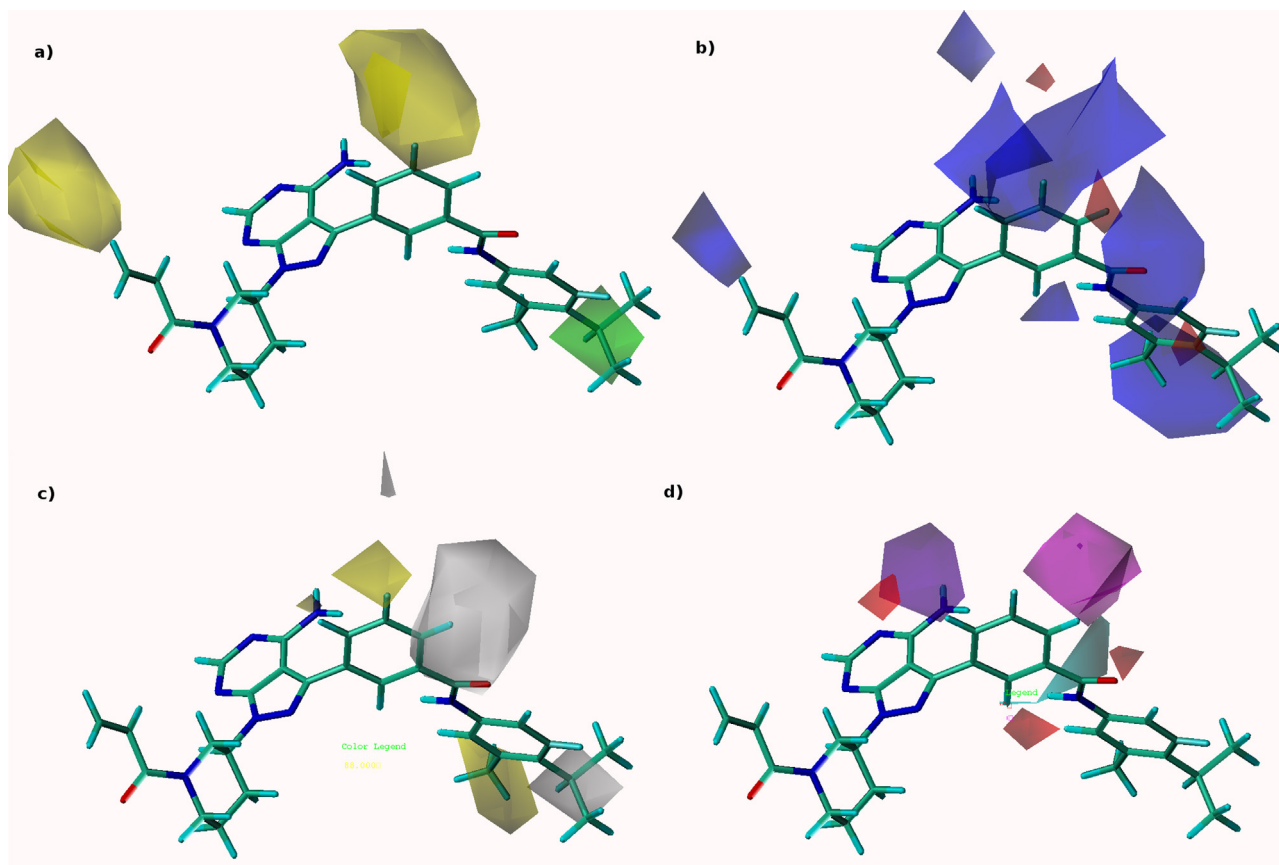
$$SEE = \sqrt{\frac{PRESS}{n - c - 1}}$$

Where, c and n is the number of components and compounds respectively, and PRESS is the addition of squared deviations

**Table 2**  
Comparison between the Statistical Parameters of CoMFA and CoMSIA Models between Ligand and Structure Based.

PLS Statistical Parameters	Ligand-Based Model		Structure-Based Model	
	CoMFA	CoMSIA	CoMFA	CoMSIA
<sup>a</sup> q <sup>2</sup> <sub>cv</sub> (LOO)	0.541	0.613	0.637	0.666
<sup>b</sup> ONC	6	5	6	4
<sup>c</sup> r <sup>2</sup> <sub>ncv</sub>	0.977	0.979	0.987	0.969
<sup>d</sup> F-value	208.8	221.425	338.300	136.582
<sup>e</sup> S.E.E	0.187	0.181	0.145	0.228
S.E.P	0.841	0.76	0.776	0.724
r <sup>2</sup> <sub>pred</sub>	0.944	0.84854	0.9414	0.86991
Fraction of field contribution				
Steric(S)	58.3%	16.4%	34.1%	16.2
Electrostatic (E)	41.7%	21%	65.9%	41.8
Hydrophobic (H)	–	45.1%	–	42.1
Hydrogen Bond Donor(HD)	–	7.6%	–	–
Hydrogen Bond Acceptor(HA)	–	9.9%	–	–

<sup>a</sup>cross-validated correlation coefficient, <sup>b</sup>optimum number of component, <sup>c</sup>non-cross-validated correlation, <sup>d</sup>Fischer statistic, <sup>e</sup>standard error of estimate.



**Fig. 6.** CoMSIA ligand-based descriptors; a) depicting steric field in which the green and yellow color represents favored and disfavored regions, respectively; b) electrostatic field showing the blue color which represents an increase of positive charge that will enhance affinity; red color represents an increase of negative charge that will enhance affinity; c) hydrophobic represented by yellow and disfavored areas represented by white colors; d) hydrogen bond donor by cyan and disfavored by purple polyhedral and hydrogen bond acceptor field favor region and disfavored areas are represented by magenta and red color, respectively with most active compound 41. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between actual and predicted activity values for individual molecule in the test dataset.

### 2.8. External validation

A q<sup>2</sup> is mostly useful but not satisfactory criteria for the validation of model. In number of cases model with good values of r<sub>cv</sub><sup>2</sup> and r<sup>2</sup> were found to be unsatisfactory. Although a model shows good prediction ability on the basis of test set statistics, but still it is not certain that every time model will predict well new data set of compounds. Thus, an external test dataset (r<sup>2</sup><sub>pred</sub>) [36,37] was sug-

gested for the assessment of predictive ability. Following equation is used to calculate predictive ability:

$$r^2_{\text{pred}} = 1 - (\text{PRESS}/\text{SD})$$

There, SD is the summed of squared variances between the calculated activities of test dataset and the average calculated activity of the training dataset.

**Table 3**  
Statistical results of all possible Ligand Based CoMSIA Models along with percentage contribution of their descriptors.

S.No.	Field Contribution	q <sup>2</sup> <sub>cv</sub>	r <sup>2</sup> <sub>ncv</sub>	ONC	S.E.E	S.E.P	F-ratio	r2pred	Field Contribution in Percentage (%)				
									S	E	H	HD	HA
1.	<sup>a</sup> S + <sup>b</sup> E	0.512	0.939	2	0.306	0.813	74.66	0.9148	49.4	50.6	–	–	–
2.	E + HD	0.456	0.832	2	0.508	0.858	24.015	0.9488	–	74.4	–	25.6	–
3	S + <sup>c</sup> H	0.614	0.958	3	0.255	0.734	110.014	0.8472	29.6	–	70.3	–	–
4	S + E + HA	0.512	0.961	3	0.245	0.825	119.36	0.8657	40.7	43.2	–	–	16.1
5	E + HD + HA	0.469	0.873	3	0.442	0.861	33.239	0.8901	–	66.2	–	14.2	19.6
6	E + HA	0.447	0.879	3	0.412	0.879	39.042	0.8217	–	72.7	–	–	27.3
7	HD + HA	0.435	0.628	3	0.757	0.888	8.176	0.895	–	–	–	53.5	46.5
8	S + E + H	0.6	0.976	4	0.193	0.759	195.688	0.8427	19.7	27.3	53	–	–
9	S + <sup>e</sup> HA	0.571	0.938	4	0.31	0.786	72.9	0.8523	71.8	–	–	–	28
10	E + H	0.569	0.977	4	0.188	0.788	205.502	0.8061	–	33.6	66.4	–	–
11	S + E + HD + HA	0.546	0.941	4	0.301	0.809	77.593	0.8879	35.8	39.3	–	10.7	14.1
12	S + E + HD	0.528	0.93	4	0.328	0.825	64.238	0.9434	41.8	45.2	–	13	–
13	S + H + HD	0.658	0.971	5	0.213	0.714	159.186	0.8502	24.3	–	63.2	12.5	–
14	S + HD + HA	0.618	0.929	5	0.33	0.754	63.714	0.8787	59.8	–	–	17.1	23.1
15	<b>S + E + H + HD + HA</b>	<b>0.613</b>	<b>0.979</b>	<b>5</b>	<b>0.181</b>	<b>0.76</b>	<b>221.425</b>	<b>0.8485</b>	<b>16.4</b>	<b>21</b>	<b>45.1</b>	<b>7.6</b>	<b>9.9</b>
16	S + E + H + HD	0.606	0.978	5	0.184	0.766	215.231	0.8578	17.7	24.2	50	8	–
17	E + H + HD	0.597	0.977	5	0.188	0.775	205.158	0.8418	–	29.7	60.4	10	–
18	H + HD + HA	0.703	0.974	6	0.199	0.677	182.8	0.8046	–	–	69.8	16.6	13.6
19	H + HD	0.692	0.974	6	0.202	0.689	177.714	0.8389	–	–	80.2	19.2	–
20	S + <sup>d</sup> HD	0.69	0.92	6	0.352	0.691	55.282	0.9398	74.8	–	–	25.2	–
21	H + HA	0.663	0.974	6	0.201	0.721	179.711	0.746	–	–	84	–	16
22	S + H + HD + HA	0.655	0.973	6	0.204	0.729	174.674	0.8249	21.6	–	53.3	11.5	13.6
23	E + H + HD + HA	0.628	0.977	6	0.187	0.757	207.446	0.8356	–	25.8	54.2	9.4	10.6
24	S + H + HA	0.623	0.976	6	0.193	0.762	195.56	0.7974	25.9	–	59	–	15.1
25	E + H + HA	0.605	0.98	6	0.174	0.78	240.808	0.8077	–	28.9	58.9	–	12.2
26	S + E + H + HA	0.599	0.979	6	0.18	0.786	225.89	–	18.2	23.1	47.7	–	11

<sup>a</sup>Steric field (S), <sup>b</sup> Electrostatic field (E), <sup>c</sup> Hydrophobic field (H), <sup>d</sup> Hydrogen bond donor (HD), <sup>e</sup> Hydrogen bond acceptor (HA), the best selected CoMSIA model is highlighted in bold character.

**Table 4**  
Statistical results of all possible Structure Based CoMSIA Models along with percentage contribution of their descriptors.

Model No.	Field Contribution	q <sup>2</sup> <sub>cv</sub>	r <sup>2</sup> <sub>ncv</sub>	ONC	S.E.E	S.E.P	F-ratio	r2pred	Field Contribution in Percentage (%)				
									S	E	H	HD	HA
1	<sup>a</sup> S + <sup>b</sup> E	0.662	0.965	4	0.242	0.722	118.825	0.943	27.8	72.2	–	–	–
2	S + <sup>c</sup> H	0.697	0.956	5	0.272	0.696	93.223	0.8691	30.5	–	69.5	–	–
3	S + <sup>d</sup> HD	0.603	0.923	6	0.358	0.812	51.67	0.8852	37.6	–	–	62.4	–
4	S + <sup>e</sup> HA	0.55	0.939	4	0.319	0.833	66.419	0.9336	36.5	–	–	–	63.5
5	E + H	0.681	0.969	5	0.227	0.721	137.589	0.8448	–	50.4	–	–	–
6	E + HD	0.599	0.936	5	0.328	0.808	63.696	0.8667	–	61.9	–	38.1	–
7	E + HA	0.500	0.955	4	0.277	0.886	91.171	0.8322	–	60.9	–	–	39.1
8	H + HD	0.623	0.944	6	0.308	0.798	73.041	0.7784	–	–	60.3	39.7	–
9	H + HA	0.636	0.960	5	0.260	0.784	104.138	0.8015	–	–	57.7	–	42.3
10	HD + HA	0.482	0.886	6	0.439	0.936	33.686	0.7718	–	–	–	48.8	51.2
11	<b>S + E + H</b>	<b>0.666</b>	<b>0.969</b>	<b>4</b>	<b>0.228</b>	<b>0.724</b>	<b>136.582</b>	<b>0.8699</b>	<b>16.2</b>	<b>41.8</b>	<b>42.1</b>	–	–
12	S + E + HD	0.601	0.942	6	0.314	0.821	69.835	0.8947	18.5	50.1	–	31.4	–
13	S + E + HA	0.552	0.955	4	0.274	0.839	93.002	0.9070	19.7	48.6	–	–	31.7
14	S + H + HD	0.635	0.942	5	0.313	0.771	70.390	0.8344	20.3	–	48.7	31.0	–
15	S + H + HA	0.618	0.956	6	0.271	0.803	95.143	0.8730	19.1	–	46.9	–	34
16	S + HD + HA	0.546	0.914	6	0.382	0.876	45.883	0.8710	23.1	–	–	37.3	39.6
17	E + H + HD	0.317	0.499	1	0.843	0.984	30.859	0.8429	–	34	31.1	34.9	–
18	E + H + HA	0.275	0.488	1	0.852	1.014	29.509	0.8513	–	37.5	34.3	–	28.2
19	E + HD + HA	0.286	0.482	1	0.857	1.006	28.863	0.8322	–	36	–	36.9	27.1
20	H + HD + HA	0.269	0.453	1	0.881	1.018	25.671	0.8299	–	–	34	38.1	27.9
21	S + E + H + HD	0.316	0.512	1	0.832	0.984	32.55	0.8481	11.4	30.1	27.6	30.9	–
22	S + E + HD + HA	0.288	0.498	1	0.844	1.005	30.699	0.8471	11.9	31.7	–	32.5	23.9
23	S + E + H + HA	0.280	0.505	1	0.837	1.011	31.672	0.8570	12.4	32.8	30.1	–	24.7
24	S + H + HD + HA	0.273	0.471	1	0.866	1.015	27.645	0.8341	12.3	–	29.8	33.4	24.5
25	E + H + HD + HA	0.289	0.482	1	0.857	1.004	28.889	0.8392	–	27.1	24.8	27.8	20.4
26	S + E + H + HD + HA	0.291	0.495	1	0.846	1.003	30.337	0.8439	9	24.6	22.5	25.2	18.5

<sup>a</sup>Steric field (S), <sup>b</sup> Electrostatic field (E), <sup>c</sup> Hydrophobic field (H), <sup>d</sup> Hydrogen bond donor (HD), <sup>e</sup> Hydrogen bond acceptor (HA), the best selected CoMSIA model is highlighted in bold character.

### 3. Results and discussion

#### 3.1. Molecular docking analysis

After evaluating the docking software by judging the resemblance of the docked pose with the co-crystallized ligand and the obtained RMSD. GOLD's default protocol and parameters was

used to search possible binding conformations of 4HCT IL-2 Itk inhibitors. Crystal structure of 4HCT with the reported active compound 52 (pIC<sub>50</sub>: 7.15) revealed the binding interaction which comprised the phenyl ring bound to the heterocyclic core and thus interacted through  $\pi$ -stacking with the gatekeeper PHE435. CYS442 was engaged through covalent bonding among the  $\beta$ -carbon (acrylamide) with cysteine-thio (C–S distance 2.0 Å). A

**Table 5**

These are the compounds with their reported  $pIC_{50}$  and predicted  $pIC_{50}$  by Ligand and Structure-based CoMFA Model along with their residuals.

Training Set of the Best Model					
CompoundNo.	$pIC_{50}$	Ligand Based Predictive	Ligand Based Residual	Structure Based Predictive	Structure Based Residual
10	5.35	5.327	-0.023	5.102	-0.248
12	4.28	4.28	0	4.551	0.271
13	4.75	4.75	0	4.903	0.153
14	4.67	4.67	0	4.703	0.033
15	5.18	5.186	0.006	5.012	-0.168
16	4.67	4.67	0	4.703	0.033
18	5.98	6.066	0.086	5.974	-0.006
20	4.99	5.235	0.245	5.008	0.018
21	5.00	5.09	0.09	4.962	-0.038
22	5.01	5.164	0.154	5.085	0.075
23	5.13	5.156	0.026	4.98	-0.15
25	4.66	4.633	-0.027	4.789	0.129
30	6.89	6.815	-0.075	6.995	0.105
31	6.72	6.654	-0.066	6.694	-0.026
32	6.74	6.696	-0.044	6.656	-0.084
33	6.70	6.945	0.245	6.528	-0.172
34	6.14	6.05	-0.09	6.215	0.075
35	5.85	6.093	0.243	5.682	-0.168
36	6.68	6.508	-0.172	6.547	-0.133
37	6.28	6.012	-0.268	6.177	-0.103
39	6.14	6.6	0.46	6.215	0.075
40	8.00	7.952	-0.048	8.23	0.23
41	8.70	8.652	-0.048	8.531	-0.169
44	5.39	5.355	-0.035	5.406	0.016
45	5.82	6.102	0.282	6.104	0.284
46	5.61	5.386	-0.224	5.517	-0.093
47	5.19	5.42	0.23	---	---
48	5.76	5.863	0.103	---	---
49	4.96	4.747	-0.213	---	---
51	4.44	4.276	-0.164	4.489	0.049
52	7.15	7.231	0.081	7.225	0.075
61	7.00	7.03	0.03	6.977	-0.023
62	8.00	7.872	-0.128	8.091	0.091
65	5.06	5.129	0.069	4.983	-0.077
Test Set of the Best Model					
11	5.60	5.346	-0.254	6.042	0.442
17	5.08	5.316	0.236	5.109	0.029
19	4.64	5.132	0.492	5.115	0.475
24	4.69	5.245	0.555	5.515	0.825
33	6.70	6.945	0.245	6.541	-0.159
38	5.99	5.725	-0.265	6.608	0.618
68	7.40	7.952	0.552	8.113	0.713

hydrogen bonding lied among the carbonyl group of acrylamide and NH of the CYS442 backbone (2.8 Å). Aminopyridine core interacted with residues MET438 and Glu436 residues with the hinge region forming two hydrogen bonds with distances 3.1 Å and 2.8 Å respectively. An interaction of  $\pi$ -stacking was between the phenyl ring of pyrazolopyrimidine with PHE435 (gatekeeper). This was presumed to determine the position of two hydrogen bonds including the carbonyl group of diarylamide forming hydrogen bond with LYS391 (3.2 Å) and nitrogen of diarylamide forming hydrogen bond with ASP500 (2.8 Å). The aminopyrazine heterocycle in pyridone interacted with residues LYS391, GLU406 by hydrogen bonding and overlapped with central phenyl ring. Most active compound 41 with  $pIC_{50}$  of 8.70 served three hydrogen bonds with residues ASP500 (1.47 Å), GLU436 (1.77 Å), LYS391 (3.19 Å). There was  $\pi$ -stacking with the gatekeeper PHE435,  $\pi$ -cation interaction with LYS391 and hydrophobic interactions were seen with the residues (MET398, MET503, ALA389, VAL419, ILE393, PHE374, PHE506) Fig. 3a. Least active compound 12 with  $pIC_{50}$  4.28  $\mu$ m showed no hydrogen bonding and only hydrophobic interactions were seen with the residues (MET438, LEU489, ALA398) and the  $\pi$ -cation interaction with LYS391 Fig. 3b. Other least active compound 19 with  $pIC_{50}$  4.64  $\mu$ m had hydrophobic interactions with (VAL377, ALA389, LYS391, LEU433, PHE435), one hydrogen bonding with LYS391 (3.19 Å) and  $\pi$ -stacking with the gatekeeper PHE435.

### 3.2. Statistics of ligand-based approach 3D QSAR models

In 3D QSAR study, CoMFA/CoMSIA techniques were used to generate models which greatly rely on the best alignment of the molecules in 3D space. The pyrazolopyrimidines dataset was aligned using compound 41 (Fig. 4a) as the template by Sybyl 7.3. The ligand-based active site alignment is shown in Fig. 4b. The final CoMFA and CoMSIA model selected was in accordance to the statistical parameters including the  $q^2_{cv} \geq 0.5$ ,  $r^2_{ncv} \approx 1$  which are as  $q^2_{cv}$ : 0.541,  $r^2_{ncv}$ : 0.977, F-value: 208.8 SEE: 0.187 as shown in Table 2. Both the CoMFA/CoMSIA models were validated by external test set containing 07 compounds. The predictive values with their residual values of the internal training and external test datasets are shown in Table 5.

### 3.3. Statistics of the structure based 3D QSAR models

The structure based CoMFA and CoMSIA models were developed on receptor based alignment. The model statistical results are shown in Table 2. It is evident from the results of CoMFA model, structure-based has given a satisfactory correlation coefficient  $q^2$  (cross-validation) of 0.637 than ligand-based 0.541 with 6 as an optimum number (of principal components) and also a satisfactory correlation coefficient  $r^2$  (non-cross validation) of 0.987 and 0.977 respectively. The resultant contributions of steric (Van



der Waals interactions) and electrostatic interactions in structure based 3D QSAR model was 34% and 66% and for ligand based was 58.3% and 41.7% respectively. The structure based CoMSIA model resulted with a correlation coefficient  $q^2$  of 0.631 and ligand based model was 0.601 and a non cross validation correlation coefficient  $r^2$  of 0.979 for structure based and 0.942 was for ligand based model. The structure based 3D QSAR model has proved to give better results than the reported ligand based model Table 2. The CoMFA/CoMSIA model predictions of the dataset were compared with the experimental  $pIC_{50}$  values and were demonstrated in graphical presentation in Fig. 8.

### 3.4. 3D QSAR models validation

The prediction of the models was validated by the test dataset. The resultant prediction of  $pIC_{50}$  values seem to correspond with the experimental  $pIC_{50}$  in satisfactory error range Table 5. The satisfactory predictive correlation coefficients value of CoMFA and CoMSIA models were 0.9414 and 0.86991 respectively. These best models can be used for the prediction of activities of novel pyrazolopyrimidines derivatives as Itk inhibitors (Table 3 and 4).

### 3.5. Contour maps analysis

Three dimensional coefficient contour maps were produced to visualize the CoMFA and CoMSIA models results. The results were graphically inferred by the field participation maps utilizing a field type of  $STDEV^*COEFFICIENTS$  which gives the enhanced view of model. The CoMFA contour maps were generated including steric and electrostatic fields are shown in Fig. 5. The CoMFA maps in Fig. 5a, b are showing the ligand based model of compound 41 and 12 and Fig. 5c, d are showing receptor based model of the compound 41 and 12 respectively. The molecular fields of CoMSIA including steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor are presented in Figs. 6 and 7 ligand-based and receptor-based respectively. The most active compound 41 was overlaid on the contour maps for visualization. Level contributions were shown by default 80% for favorable and 20% unfavorable regions.

### 3.6. Steric and electrostatic profile

The CoMFA and CoMSIA model contour maps with various fields' contribution are demonstrated with inhibitor 41 as reference compound with the highest inhibitory activity. In the steric field, green regions correspond to areas where steric tenancy with large polyhedral groups will increase in the inhibition activity and the yellow polyhedral areas have to be sterically avoided. In the electrostatic field, red contour contribution corresponds to the region where electronegative functional group is promising while conflicting is for blue contours. In the present study, we examined the contours maps regarding their reference individual group. In comparison the CoMSIA steric field is easier as compared to CoMFA in order to demonstrate exactly how steric effects affect the inhibitory activity of inhibitors Figs. 6 and 7. The most potent compound 41 of this series was placed on CoMFA contour maps inside the active site of receptor protein as shown in Fig. 5a (ligand-based model), Fig. 5b (receptor-based model). It is obvious from the figure that two yellow and one green contours are present in the developed CoMFA model. The green contour is found in close proximity to the isopropyl moiety of the R1 group; specify that bulkier group replacement at this position can enhance the inhibitory activity. It is further supported when we compare the most active compound 41 with least active compound 12 as shown in Fig. 5c (ligand based model), Fig. 5d (receptor based model), the decline in inhibitory activity was observed because the isopropyl group on benzene is replaced by cyanide group. Compounds 47, 49, 51 are true represen-

tative for this prediction in explanation; the occurrence of bulkier groups at the R1 position is suitable for improving inhibitory activity. First yellow contour is found near the oxygen of amide group while second yellow map is found close to the oxygen of vinyl ketone. This examination supports the sterically less bulky moiety at these positions could be better for good interaction, so as to accommodate within the binding site of Itk. In the current scenario, electrostatic field contribution is prominent over steric field and forming a powerful correlation with neighboring interacting amino acids. A large blue contour surrounds the isopropyl moiety at the R1 position and powerfully suggests the presence of an electron donating group at this point to improve the activity profile. Additionally, the electron withdrawing parts of the interacting residues of Itk were found in contact with electron donating part of inhibitors surrounded in the blue contours. So to enhance the activity profile powerful interaction with the electronegative side-chain of amino acids is preferred for established interaction of inhibitors with protein. This assessment was additional supported by another blue contour map which lies nearby  $NH_2$  moiety of amino acid residues. In CoMFA, two red contours positioned in the surrounding of N-substitute piperidine and another red contour found on the carbonyl oxygen which is attached with the  $NH_2$  group of isopropyl benzene exactly correspond with the hydrophobic residue. These contours determine the presence of an electron withdrawing group which established interactions with electron donating groups which is found close to these contours in the active binding site of Itk. The decrease in activity is observed in compounds 18, 43, 46, 36, 51 and 48 as these inhibitors were located away from red contours.

### 3.7. CoMSIA analysis

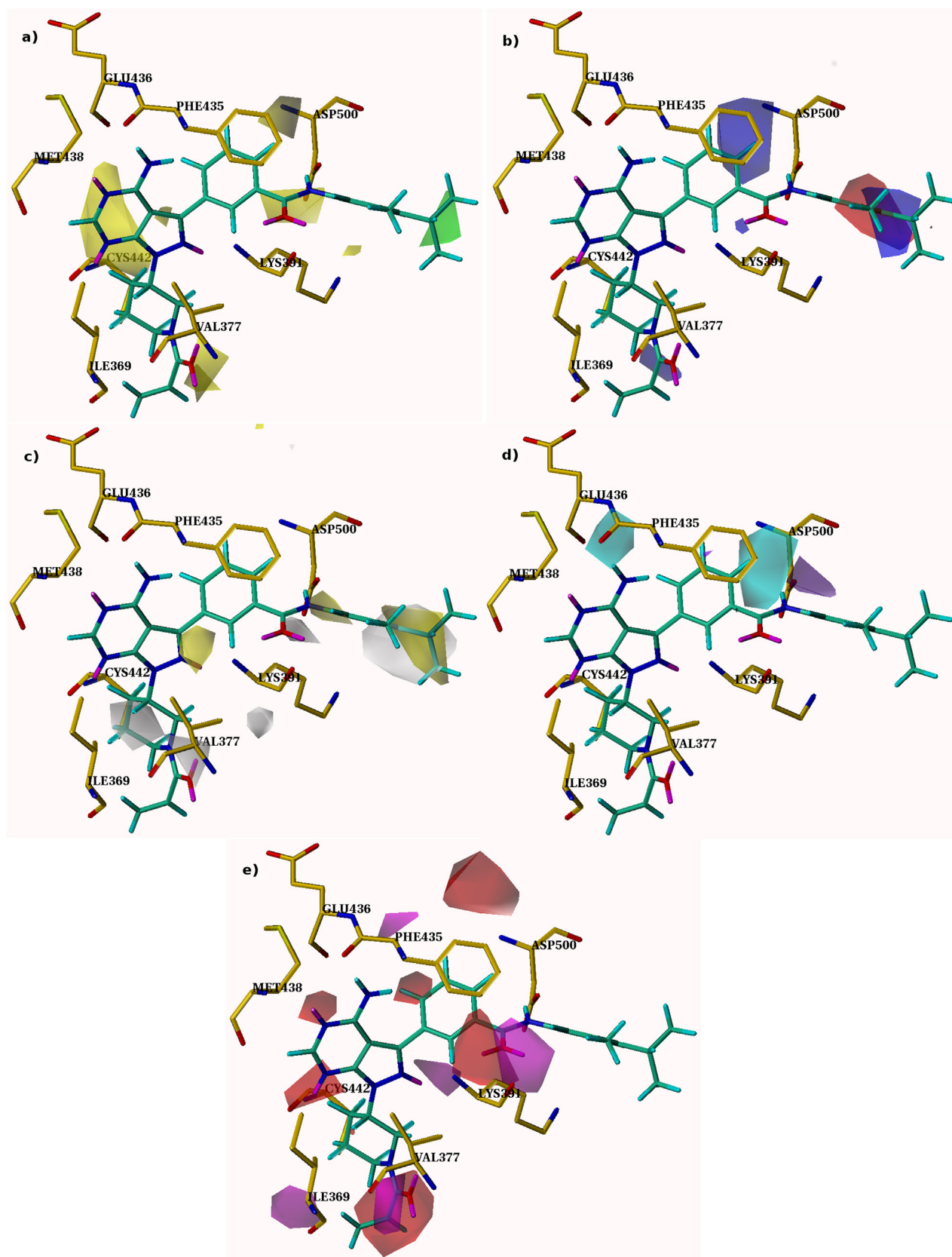
#### 3.7.1. Steric and electrostatic fields

The contour maps of steric and electrostatic fields are presented in Fig. 6 (ligand-based model) and Fig. 7 (receptor-based model). We found almost similar steric and electrostatic field contributions for designed Itk inhibitors so we mainly focus on molecular fields of hydrophobic, hydrogen bond donor and hydrogen bond acceptor.

### 3.8. Hydrophobic fields

The white and yellow colored contours show respective areas where hydrophilic and hydrophobic properties were found to be favorable. When reference compound overlaid on the developed contours, two yellow contours correspond the favorable perspective for hydrophobic group as showed in Fig. 7 one of yellow contour lies near the vinyl ketone while second contour is present near the isopropyl benzene close to R1 position. This observation showed that hydrophobicity at this part of the molecule play significant role in biological activity. This lipophilic pattern followed by compound 41 while the interacting residues present in yellow contour strongly suggest that more lipophilic character in place of less hydrophobic residues. If we observed that the low activity compound series as compared to reference, the reason behind that may be these compounds represented close to yellow contours and hydrophilic residues at this position are absent as compared to template. Additionally there are two white contour present near the piperidine and benzene ring specify that presence of hydrophobic moieties at this area are not suitable for activity. It is additionally supported after a comparison of the most potent compound 41 with less active compounds (12, 13, 15, 19, 31, 42, and 45).

The CoMSIA hydrogen bond donor and hydrogen bond acceptor contour plots are presented in Fig. 7 respectively. The cyan contours characterize the areas where hydrogen bond donating moieties enhance the activity, while the purple contours characterize the areas where hydrogen bond-donating moieties decline the activ-



**Fig. 7.** Structure Based CoMSIA Contour Maps a) Steric field is showing the favored and disfavored regions; b) In the Electrostatic field contour maps have blue color representing an increase of positive charge that will enhance affinity; red color represent an increase of negative charge that will enhance affinity; c) Hydrophobic favored region represented by favoured and disfavored areas; d) H-bond donor field represented by cyan and disfavored by purple polyhedra; e) H-bond acceptor field favor region and disfavored areas are represented by magenta and red color, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

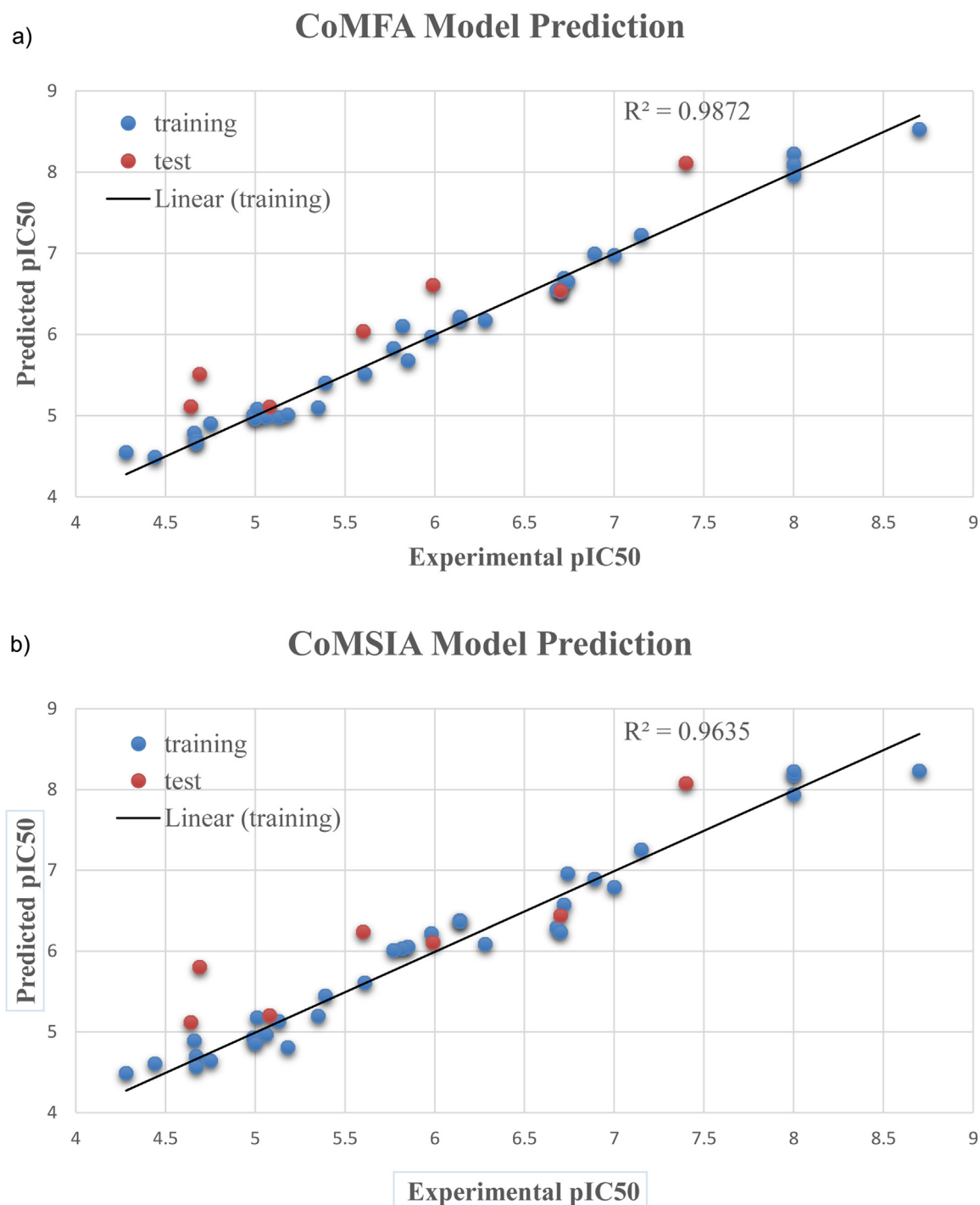


Fig. 8. Graphical presentation of Structure based CoMFA and CoMSIA.

ity. In the same way, the magenta contours point out the regions where hydrogen bond-accepting moieties enhance the inhibitory activity, while the red contours specify the areas where hydrogen bond-accepting moieties decline the activity. The cyan contours near the NH<sub>2</sub> group which is present on pyrimidine ring show that hydrogen bond-donating functionalities are favored. It is well established with the explanations that NH<sub>2</sub> group in this area forms hydrogen bonds with residues MET438 and Glu436 as donor. The purple contour close to N-linker favors the existence of this moiety which can take lone pair of electron from active site amino acids to make better interaction. A large magenta contour situated on the vinyl ketone, substituted on piperidine ring while second magenta

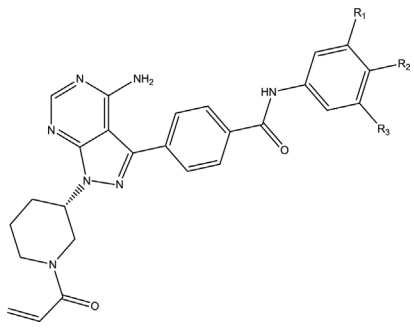
contour found close to oxygen of amide group recommends that hydrogen bond-accepting moieties are favored in these areas. Similarly one large red contour, found near the nitrogen of piperidine moiety suggest that its existence is not promising for H-bond acceptors group; this explanation is further proved by noticing lone pair donating group of compound can receive the electron pair from the protein for making better interactions.

### 3.9. Design of new inhibitors

The CoMFA and CoMSIA models provide detailed description about structural requirements for improved inhibitory activity of



**Table 6**  
Newly designed compounds with their predictive pIC<sub>50</sub> values.



Compound No.	R1	R2	R3	Predicted pIC <sub>50</sub> μM
1	NH <sub>2</sub>	–NH(CH <sub>3</sub> ) <sub>2</sub>	H	8.85
2	CH <sub>3</sub>	–CHCH <sub>3</sub> C <sub>2</sub> H <sub>5</sub>	CH <sub>3</sub>	8.46
3	CH <sub>3</sub>	–C(CH <sub>3</sub> ) <sub>2</sub>	H	8.45
4	NCH <sub>3</sub>	–N(CH <sub>3</sub> ) <sub>2</sub>	H	8.44
5	CH <sub>3</sub>	–CNH <sub>2</sub> (CH <sub>3</sub> ) <sub>2</sub>	H	8.43
6	C <sub>2</sub> H <sub>5</sub>	–N(CH <sub>3</sub> ) <sub>2</sub>	NH <sub>2</sub>	8.43

pyrazolopyrimidines derivatives. Particularly, the R<sub>1</sub> moiety plays an important role in its inhibition. Its substitution with hydrogen or less bulky group cause total loss of its activity. Replacement of linker oxygen atom with nitrogen and sulfur may compromise in the activity of the compounds. Insertion of steric and hydrophobic moieties at R<sub>3</sub>-position of benzene ring could enhance the activity. The practical values of these developed structure-activity relationships reveal by designing a series of new inhibitors (with predicted pIC<sub>50</sub> values ranging from 8.43 to 8.85 log units) with the best developed CoMFA and CoMSIA models (Table 6).

#### Author's contributions

Conceived and designed the experiments: ZUH, JSE. Performed the experiments: JSE, ZUH. Analyzed the data: JSE, SA. Manuscript written by: ZUH, JSE, SA. All authors read and approved the final manuscript.

#### Conflict of interest

The authors declared no competing interests.

#### Acknowledgement

Authors are highly acknowledged Higher Education Commission (HEC) of Pakistan for their financial grants (Grant ID: NRP/1509), and also grateful to Prof. Bernd M. Rode (University of Innsbruck) for providing their technical support to conduct this research work.

#### References

- [1] D.I. Marks, E.M. Paietta, A.V. Moorman, S.M. Richards, G. Buck, G. DeWald, et al., T-cell acute lymphoblastic leukemia in adults: clinical features, immunophenotype, cytogenetics, and outcome from the large randomized prospective trial (UKALL XII/ECOG 2993), *Blood* 114 (2009) 5136–5145.
- [2] Y.H. Kim, H.L. Liu, S. Mraz-Gernhard, A. Varghese, R.T. Hoppe, Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression, *Arch. Dermatol.* 139 (2003) 857–866.
- [3] A. August, S. Gibson, Y. Kawakami, T. Kawakami, G.B. Mills, B. Dupont, CD28 is associated with and induces the immediate tyrosine phosphorylation and activation of the Tec family kinase ITK/EMT in the human Jurkat leukemic T-cell line, *Proc. Natl. Acad. Sci.* 91 (1994) 9347–9351.
- [4] A.H. Andreotti, P.L. Schwartzberg, R.E. Joseph, L.J. Berg, T-cell signaling regulated by the Tec family kinase, Itk, *Cold Spring Harbor Perspect. Biol.* 2 (2010) a002287.
- [5] J.A. Grasis, D.M. Guimond, N.R. Cam, K. Herman, P. Magotti, J.D. Lambris, et al., In vivo significance of ITK-SLP-76 interaction in cytokine production, *Mol. Cell. Biol.* 30 (2010) 3596–3609.
- [6] T. Berge, V. Sundvold-Gjerstad, S. Granum, T.C. Andersen, G.B. Holthe, L. Claesson-Welsh, et al., T cell specific adapter protein (TSAd) interacts with Tec kinase ITK to promote CXCL12 induced migration of human and murine T cells, *PLoS One* 5 (2010) e9761.
- [7] C.C. Carson, S.J. Moschos, S.N. Edmiston, D.B. Darr, N. Nikolaishvili-Feinberg, P.A. Groben, et al., IL2 inducible T-cell kinase, a novel therapeutic target in Melanoma, *Clin. Cancer Res.* 21 (2015) 2167–2176.
- [8] S.B. Kanner, J.J. Perez-Villar, Altering T-cell activation by targeting the multidomain tyrosine kinase Itk, *Trends Immunol.* 24 (2003) 249–253.
- [9] J.A. Readinger, G.M. Schiralli, J.-K. Jiang, C.J. Thomas, A. August, A.J. Henderson, et al., Selective targeting of ITK blocks multiple steps of HIV replication, *Proc. Natl. Acad. Sci.* 105 (2008) 6684–6689.
- [10] J. Das, C. Liu, R.V. Moquin, J. Lin, J.A. Furch, S.H. Spergel, et al., Discovery and SAR of 2-amino-5-[(thiomethyl) aryl] thiazoles as potent and selective Itk inhibitors, *Bioorg. Med. Chem. Lett.* 16 (2006) 2411–2415.
- [11] J. Das, J.A. Furch, C. Liu, R.V. Moquin, J. Lin, S.H. Spergel, et al., Discovery and SAR of 2-amino-5-(thioaryl) thiazoles as potent and selective Itk inhibitors, *Bioorg. Med. Chem. Lett.* 16 (2006) 3706–3712.
- [12] A.D. Velankar, G. Quintini, A. Prabhu, A. Weber, G. Hunaeus, B. Voland, et al., Synthesis and biological evaluation of novel (4 or 5-aryl) pyrazolyl-indoles as inhibitors of interleukin-2 inducible T-cell kinase (ITK), *Bioorg. Med. Chem.* 18 (2010) 4547–4559.
- [13] R.J. Snow, A. Abeywardane, S. Campbell, J. Lord, M.A. Kashem, H.H. Khine, et al., Hit-to-lead studies on benzimidazole inhibitors of ITK: discovery of a novel class of kinase inhibitors, *Bioorg. Med. Chem. Lett.* 17 (2007) 3660–3665.
- [14] M.P. Winters, D.J. Robinson, H.H. Khine, S.S. Pullen, J.R. Woska, E.L. Raymond, et al., 5-Aminomethyl-1H-benzimidazoles as orally active inhibitors of inducible T-cell kinase (Itk), *Bioorg. Med. Chem. Lett.* 18 (2008) 5541–5544.
- [15] B.N. Cook, J. Bentzien, A. White, P.A. Nemoto, J. Wang, C.C. Man, et al., Discovery of potent inhibitors of interleukin-2 inducible T-cell kinase (ITK) through structure-based drug design, *Bioorg. Med. Chem. Lett.* 19 (2009) 773–777.
- [16] K.J. Moriarty, H. Takahashi, S.S. Pullen, H.H. Khine, R.H. Sallati, E.L. Raymond, et al., Discovery, SAR and X-ray structure of 1H-benzimidazole-5-carboxylic acid cyclohexyl-methyl-amides as inhibitors of inducible T-cell kinase (Itk), *Bioorg. Med. Chem. Lett.* 18 (2008) 5545–5549.
- [17] K.J. Moriarty, M. Winters, L. Qiao, D. Ryan, R. Desjarlis, D. Robinson, et al., Itk kinase inhibitors: Initial efforts to improve the metabolic stability and the cell activity of the benzimidazole lead, *Bioorg. Med. Chem. Lett.* 18 (2008) 5537–5540.
- [18] C.W. Zapf, B.S. Gerstenberger, L. Xing, D.C. Limburg, D.R. Anderson, N. Caspers, et al., Covalent inhibitors of interleukin-2 inducible T cell kinase (Itk) with nanomolar potency in a whole-blood assay, *J. Med. Chem.* 55 (2012) 10047–10063.
- [19] W. Sippl, Development of biologically active compounds by combining 3D QSAR and structure-based design methods, *J. Comput. Aided Mol. Des.* 16 (2002) 825–830.
- [20] Hevener, Kirk E. et al. Quantitative structure–activity relationship studies on nitrofuranyl anti-tubercular agents. *Bioorganic & medicinal chemistry* 16.17 (2008): 8042–8053.
- [21] UltraR CC. CambridgeSoft Corporation, C.D.C., MA. 2140.
- [22] OLBoyle NM, B.M., James CA, Morley C, Vandermeersch T, Hutchison GR. QUACPAC Open Babel v3.3 : An open chemical toolbox. *J Cheminf.* 2011;3:33.
- [23] Software., O.S.S.M.O.S. Openeye Scientific Software Molcharge v1.6.3.1. Openeye Scientific Software. 2010.
- [24] SYBYL v7.3 Tripos Associates St. Louis, MO. SYBYL v7.3 Tripos Associates St. Louis, MO 2007.
- [25] OpenEye Scientific Software OMEGA 2.5.1.4: Santa Fe, NM. Hawkins, P.C.D.; Skillman, A.G.; Warren, G.L.; Ellingson, B.A.; Stahl, M.T.
- [26] J. Verma, V.M. Khedkar, E.C. Coutinho, 3D-QSAR in drug design—a review, *Curr. Top. Med. Chem.* 10 (2010) 95–115.
- [27] GOLDv5.2.2 Cambridge Crystallographic Data Center, U. GOLDv3.2 Cambridge Crystallographic Data Center, UK. 2008.
- [28] Molecular Operating Environment (MOE), 2015.1001; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016.
- [29] P.R. Murumkar, R. Giridhar, M.R. Yadav, 3D-quantitative structure–activity relationship studies on benzothiadiazepine hydroxamates as inhibitors of tumor necrosis factor-α converting enzyme, *Chem. Biol. Drug Des.* 71 (2008) 363–373.
- [30] P. Lan, Z.-J. Huang, J.-R. Sun, W.-M. Chen, 3D-QSAR and molecular docking studies on fused pyrazoles as p38α mitogen-activated protein kinase inhibitors, *Int. J. Mol. Sci.* 11 (2010) 3357–3374.
- [31] R. Hu, F. Barbault, M. Delamar, R. Zhang, Receptor-and ligand-based 3D-QSAR study for a series of non-nucleoside HIV-1 reverse transcriptase inhibitors, *Bioorg. Med. Chem.* 17 (2009) 2400–2409.
- [32] R.N. Reddy, R.R. Mutyala, P. Aparoy, P. Reddanna, M.R. Reddy, An analysis of hydrophobic interactions of thymidylate synthase with methotrexate: free energy calculations involving mutant and native structures bound to methotrexate, *J. Mol. Model.* 16 (2010) 203–209.
- [33] G. Klebe, U. Abraham, Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries, *J. Comput. Aided Mol. Des.* 13 (1999) 1–10.

- [34] S.K. Sivan, V. Manga, Molecular docking and 3D-QSAR studies on triazolinone and pyridazinone, non-nucleoside inhibitor of HIV-1 reverse transcriptase, *J. Mol. Model.* 16 (2010) 1169–1178.
- [35] M. Clark, R.D. Cramer, The probability of chance correlation using partial least squares (PLS), *Quant. Struct.-Act. Relat.* 12 (1993) 137–145.
- [36] G. Alexander, T. Alexander, Beware of Q<sup>2</sup>, *J. Mol. Graph. Model.* 20 (2002) 269–276.
- [37] P.P. Roy, K. Roy, On some aspects of variable selection for partial least squares regression models, *QSAR & Comb. Sci.* 27 (2008) 302–313.