MEDICINAL
CHEMISTRY
RESEARCH

ORIGINAL RESEARCH

A combined 3D QSAR and pharmacophore-based virtual screening for the identification of potent p38 MAP kinase inhibitors: an in silico approach

**Gugan Kothandan · Thirumurthy Madhavan ·
Changdev G. Gadhe · Seung Joo Cho**

Abstract p38 kinase plays a vital role in inflammation mediated by tumor necrosis factor-α and interleukin-1β pathways. Inhibition of p38 kinase provides an effective way to treat inflammatory diseases. 3D-QSAR study was performed to obtain reliable comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) models for a series of p38 inhibitors with three different alignment methods (Receptor based, atom by atom matching, and pharmacophore based). Among the different alignment methods, better statistics were obtained with receptor-based alignment (CoMFA: $q^2 = 0.777$, $r^2 = 0.958$; CoMSIA: $q^2 = 0.782$, $r^2 = 0.927$). Superposing CoMFA/CoMSIA contour maps on the p38 active site gave a valuable insight to understand physical factors which are important for binding. In addition, this pharmacophore model was used as a 3D query for virtual screening against NCI database. The hit compounds were further filtered by docking and scoring, and their biological activities were predicted by CoMFA and CoMSIA models.

Keywords p38 inhibitors · 3D QSAR · CoMFA · CoMSIA · Pharmacophore · GALAHAD

G. Kothandan · T. Madhavan · C. G. Gadhe · S. J. Cho (✉)
Department of Bio-New Drug Development, College of
Medicine, Chosun University, 375 Seosuk-dong, Dong-gu,
Gwangju 501-759, Republic of Korea
e-mail: chosj@chosun.ac.kr

S. J. Cho
Department of Cellular Molecular Medicine and Research
Center for Resistant Cells, College of Medicine, Chosun
University, Gwangju 501-759, Republic of Korea

Introduction

Rheumatoid arthritis (RA) is characterized by the chronic inflammation of joints that leads to destruction of cartilage and bone deformation (Chen *et al.*, 1993; Pincus, 1995). Patients suffering from arthritis and other autoimmune diseases have elevated levels of pro-inflammatory cytokines such as TNFα and IL-1β in the synovial fluid (Dinarello, 1991; Feldmann *et al.*, 1996). They are produced by synovial macrophages and exhibit similar biological activities like cell proliferation, collagen synthesis, adhesion molecule expression, etc. (Dominguez *et al.*, 2005). Clinical therapies with anticytokines were employed to obstruct the hyperactivated pro-inflammatory cytokines and prevent the joint damage caused by TNFα and IL-1β. The cost of anticytokine treatment, mode of administration, and post-treatment infections hampered the effective use of these anticytokine therapies (Andreakos *et al.*, 2002). Intracellular signaling pathways involving p38 mitogen-activated protein kinase (MAPK) regulate the production of several pro-inflammatory cytokines and are considered as a focal point in the development of new therapeutic agents to treat inflammatory disease such as RA (Kulkarni *et al.*, 2006). p38 MAP kinases are proline-directed serine-threonine protein kinases that are activated by various growth factors and pro-inflammatory cytokines by the dual phosphorylation of TXY motifs in the activation loop (Pearson *et al.*, 2001). Four variants of p38 kinase have been recognized, namely, p38α (also known as p38), p38β, p38γ (ERK6/SAPK3), and p38δ (SAPK4) with 40–60 % structural similarity (Lee *et al.*, 1994; Jiang *et al.*, 1996; Li *et al.*, 1996; Jiang *et al.*, 1997). Several reports indicate that p38α has profound role in RA and is involved in the expression of TNFα and IL-1β at both transcription and translation levels while the roles of p38β, p38γ, and p38δ have to be identified critically (Newton and Holden, 2003).

In the last couple of decades, the three dimensional structure-activity relationship (3D QSAR) techniques such as comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) (Cramer et al., 1988; Klebe et al. 1994; Bohm *et al.*, 1999) have been routinely used in modern drug design. Current studies are now focused on the development of novel inhibitors. Several series of p38α kinase inhibitors are reported and despite these, more structurally diverse inhibitors should be discovered. The primary aim of this study is lead optimization and lead identification of p38 kinase inhibitors. QSAR techniques are most often used as a tool for lead optimization within the congeneric domain of molecules. There is strong literature evidence suggesting that lead identification can be achieved by pharmacophore-based methods (Martin *et al.*, 1993; Jones *et al.*, 1995; Patel *et al.*, 2002; Willett and Winterman, 1986). In this study, we combined both these approaches to identify lead compounds using pharmacophore-based approach and optimize the lead compounds using QSAR approach. We have developed 3D-QSAR models for a series of dihydroquinazolinone inhibitors (John *et al.*, 2003) using receptor-based and ligand-based schemes (atom by atom matching and pharmacophore based). The generated CoMFA and CoMSIA models were further validated by the test set. Receptor-based model was selected for further explanation of contour maps. Furthermore, the pharmacophore which was generated for 3D-QSAR studies was validated with the crystal structure-bound conformation of potent p38 kinase inhibitor Dihydroquinazolinone (PDB code: 1M7Q) and then used in the virtual screening of NCI database. Compounds with high screening scores obtained after different filtering were further supported by CoMFA and CoMSIA with good predicted pIC50 values. The results obtained could be useful for further prospects.

Materials and methods

Dataset

Forty five novel p38α kinase inhibitors were taken from literature with their biological activities (John *et al.*, 2003) in terms of IC_{50} values and forty one molecules were selected for the development of model. Four molecules were excluded because of their low inhibition concentration (IC_{50}). These compounds were excluded as they might act as outliers and it would affect the quality of the model. The IC_{50} values i.e., the concentration (nM) of inhibitors that produces 50 % inhibition of p38 kinase were converted into pIC_{50} as reported in Tables 1, 2, and 3. The dataset is randomly divided into a training set of 32 molecules and test set of nine molecules. The IC_{50} (nM) values

Table 1 Structure and biological values (pIC50) of dihydroquinazolinone inhibitors (series 1)
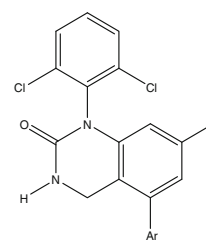
Compound 9c-9r, Compound 10a-10c

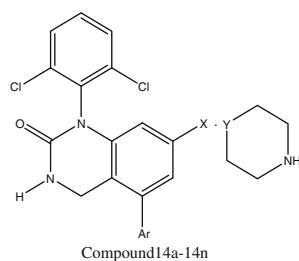| Compound no.	Ar	R	X	Y	pIC_{50}
9c	Phenyl	Co_2Me	–	–	7.523
9d	2-Fl-Phenyl	Co_2Me	–	–	8.155
9e[a]	3-Fl-Phenyl	Co_2Me	–	–	7.102
9f	4-Fl-Phenyl	Co_2Me	–	–	7.658
9g[a]	2-Cl-Phenyl	Co_2Me	–	–	9.000
9h	3-Cl-Phenyl	Co_2Me	–	–	7.174
9i	4-Cl-Phenyl	Co_2Me	–	–	7.328
9k[a]	3-CF_3-Phenyl	Co_2Me	–	–	6.770
9l	4-CF_3-Phenyl	Co_2Me	–	–	6.886
9m	2-CH_3-Phenyl	Co_2Me	–	–	7.959
9n[a]	3-CH_3-Phenyl	Co_2Me	–	–	6.495
9o	4-CH_3-Phenyl	Co_2Me	–	–	7.357
9p	2,4-di-Fl-Phenyl	Co_2Me	–	–	8.155
9q	2,6-di-Fl-Phenyl	Co_2Me	–	–	7.357
9r[a]	2-CH_3-4-Fl-Phenyl	Co_2Me	–	–	7.959
10a	2-Cl-4-Fl-Phenyl	OMe	–	–	8.523
10b	2,4-di-Fl-Phenyl	OMe	–	–	7.959
10c	2-Cl-Phenyl	OMe	–	–	8.222

[a] Test set compounds

were taken in molar (M) range and converted to pIC_{50} according to the formula given in following Eq.

$$pIC_{50} = -\log IC_{50}$$

Molecular docking

Docking study was performed to validate the hits obtained from virtual screening using SYBYL 8.1[1] (SYBYL 8.1., Tripos, St. Louis, MO) molecular modeling package installed on Linux system. Protein structure was prepared using biopolymer module of SYBYL. Hydrogen atoms were added to structure, atom types and charges were assigned using AMBER7 FF99 force field, and side chain amides were modified. Docking study was performed using

Table 2 Structure and biological values (pIC$_{50}$) of dihydroquinazolinone inhibitors (series 2)

Compound14a-14n

Compound no.	Ar	R	X	Y	pIC$_{50}$
14a	2-Cl-Phenyl	–	–	CH	9.046
14b	2,4-di-Fl-Phenyl	–	–	CH	9.699
14c	2-Cl-4-Fl-Phenyl	–	–	CH	10.000
14d	2-Cl-Phenyl	–	–	N	8.854
14e	2,4-di-Fl-Phenyl	–	–	N	8.585
14f	2-Cl-Phenyl	–	O	CH	9.699
14g[a]	2-Cl-4-Fl-Phenyl	–	O	CH	10.000
14h	2-Cl-Phenyl	–	NH	CH	9.301
14i	2,4-di-Fl-Phenyl	–	NH	CH	9.222
14j	2-Cl-Phenyl	–	CH$_2$	N	10.000
14k[a]	2-Cl-4-Fl-Phenyl	–	CH$_2$	N	9.886
14l	2-Cl-Phenyl	–	CO	N	8.824
14m	2,4-di-Fl-Phenyl	–	CO	N	8.620
14n	2-Cl-4-Fl-Phenyl	–	CO	N	8.959

[a] Test set compounds

Surflex-Dock module of SYBYL, which uses empirical scoring function to score ligand and protomol (Ruppert et al., 1997) guided docking. All the molecules were docked to the inhibitor binding site of p38 kinase crystal structure in complex with DQO (PDB code: 1M7Q) (John et al., 2003).

Pharmacophore generation

The pharmacophore hypothesis was generated using GALAHAD (genetic algorithm with linear assignment of hypermolecular alignment of datasets) module of SYBYL. GALAHAD operates in two main stages: the ligands are aligned to each other in internal coordinate space, and then the conformations produced are aligned in Cartesian space. The features considered in developing the pharmacophore model includes hydrogen bond donor atoms, hydrogen bond acceptor atoms, hydrophobic and charged centers.

Table 3 Structure and biological values (pIC$_{50}$) of dihydroquinazolinone inhibitors (series 3)

Compound15a-15i
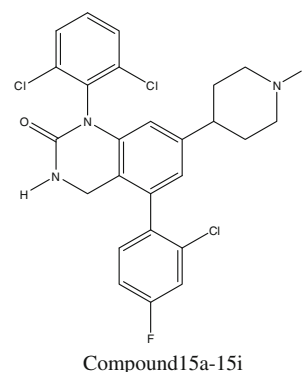
Compound no.	Ar	R	X	Y	pIC$_{50}$
15a	–	Methyl	–	–	9.301
15b	–	Ethyl	–	–	8.921
15c	–	i-Propyl	–	–	9.222
15d[a]	–	Cyclopropyl	–	–	8.959
15e	–	Methyl cyclopropyl	–	–	9.301
15f	–	Ethyl 1-cyclopropyl	–	–	9.523
15g[a]	–	Cyclobutyl	–	–	9.398
15h	–	Methyl cyclobutyl	–	–	9.222
15i	–	t-Butyl	–	–	9.699

[a] Test set compounds

In this study, eight compounds (i.e., **9g**, **14b**, **14c**, **14e**, **14k**, **15a**, **15f**, and **15i**) were selected to carry out pharmacophore hypothesis and the genetic algorithm was used to create conformers for all molecules. The compounds which were selected to generate pharmacophore hypothesis are highly active. Five models are generated with default parameters. The pharmacophore indicates that donor and acceptor atoms are quite crucial to target the hinge region. The selected pharmacophore model is shown in Fig. 1.

Atom by atom matching alignment

In this scheme, one of the most active compounds (**14c**) was used as the template. Systematic search was used in the conformational analysis and all rotatable bonds were searched in 10° increments from 0° to 360°. Conformational energy was computed with electrostatic term and the lowest energy conformer was selected. The template was modified for other ligands in the dataset. The common moiety was constrained for each molecule and the substituents were minimized at Tripos force field. The minimized structures were aligned over the template using atom by atom matching method and are displayed in Fig. 2. The alignment obtained was subsequently used for CoMFA and CoMSIA.
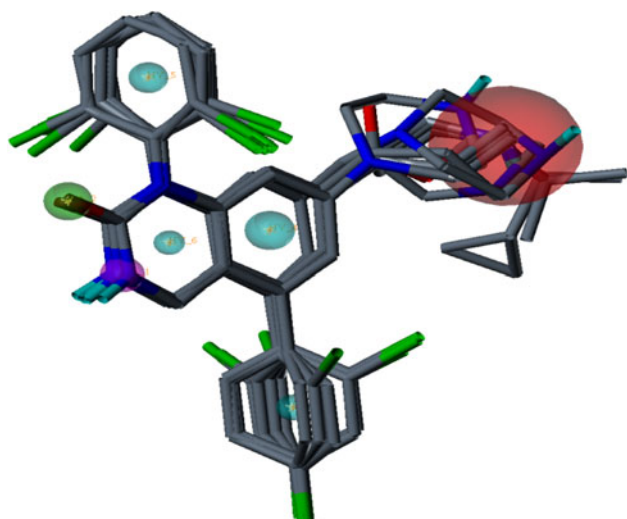
Fig. 1 Representation of the selected pharmacophore model. *Cyan* indicates hydrophobes, *green* indicates hydrogen bond acceptors, *magenta* indicates hydrogen bond donors, and *red* indicates positive nitrogens (Color figure online)
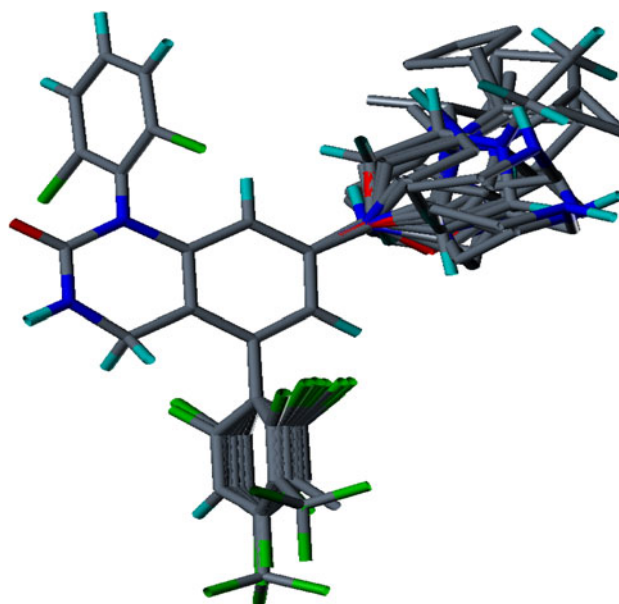

Fig. 2 Alignment of molecules in the dataset by atom and atom matching method

Pharmacophore-based alignment

Pharmacophore-based alignment was done using GALA-HAD. In this scheme, the molecules in the dataset (41 compounds) were aligned to the selected pharmacophore model. The alignment was done based on the common pharmacophores between the molecules in the dataset. The alignment obtained from this scheme along with their pharmacophoric features is shown in Fig. 3 and subsequently used for CoMFA and CoMSIA.

Receptor-based alignment

Among the inhibitors in the dataset, one of the compound (**14e**) was complexed with the receptor is known (21). Therefore, the ligand (**14e**) extracted from the protein structure (PDB code: 1M7Q) was used as the template molecule. The common moiety was constrained for each molecule and other molecules in the dataset were modified. The compounds in the dataset were minimized within the receptor site using Tripos force field, but the whole protein active site was fixed during minimization. All minimized structures inside the receptor were superimposed to get the molecular alignment for CoMFA and CoMSIA, and subsequently used for analysis. The alignment of the molecules inside the receptor is shown in Fig. 4.

Database search

The selected pharmacophore model was validated and converted into a UNITY query for pharmacophore-guided
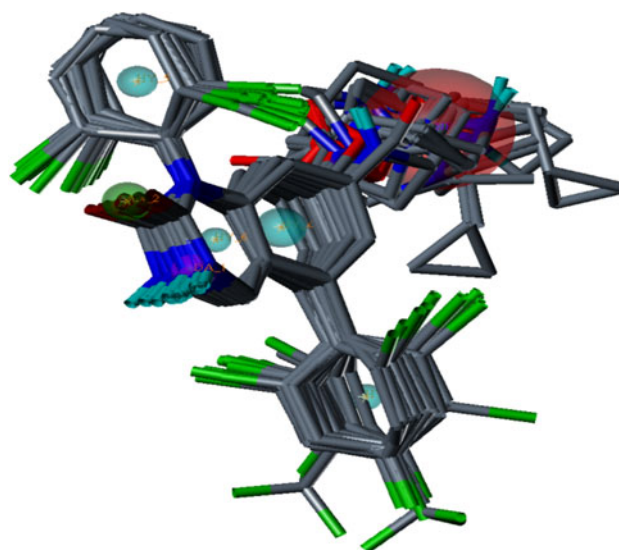

Fig. 3 Alignment of molecules by pharmacophore-based alignment methods. 41 molecules in the dataset are aligned based on the common pharmacophores using GALAHAD

virtual screening studies. The query was screened against NCI2000 database. The "flexible database search" option was implemented to perform virtual screening. Primary filters such as Lipinski's rule of five, Van der Waals bumps, restricting the number of rotatable bonds to ≤7, and QFIT (pharmacophoric match between query and the hit compound) were applied to reduce the dataset. Further screening of the hits was carried out using the docking algorithm, Surflex-Dock in SYBYL. The generated UNITY query with distance constraints is shown in Fig. 5.
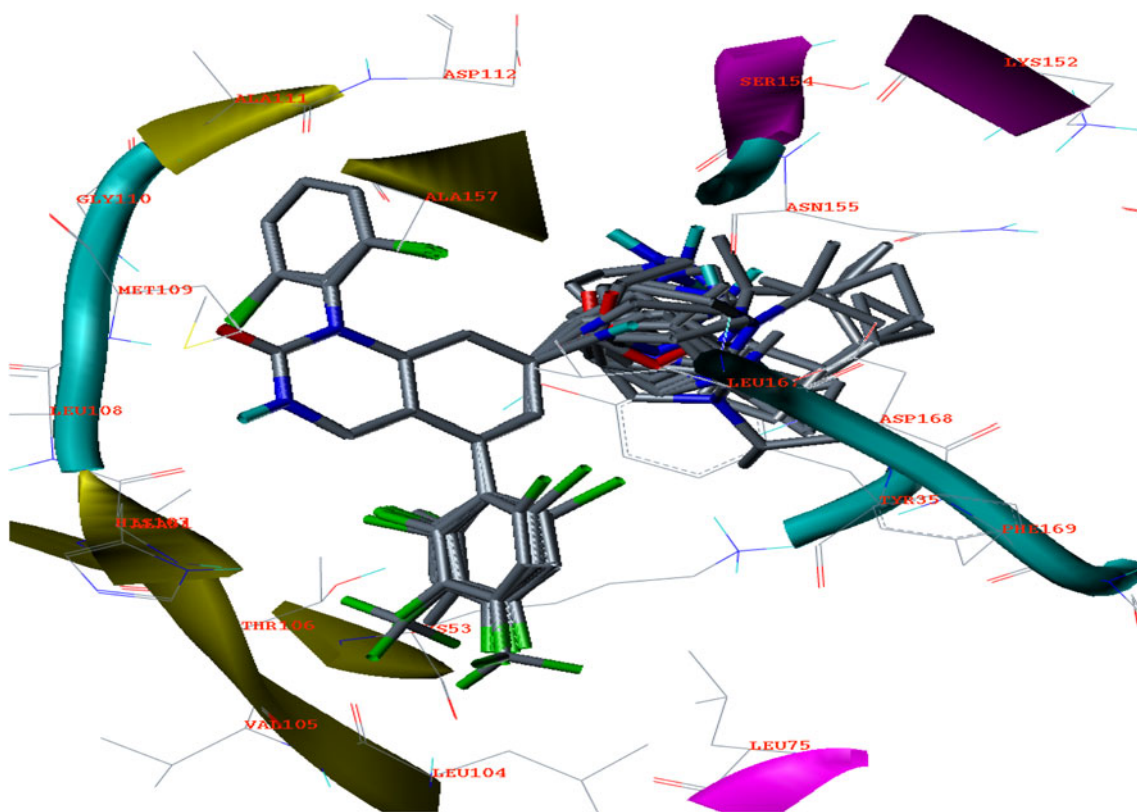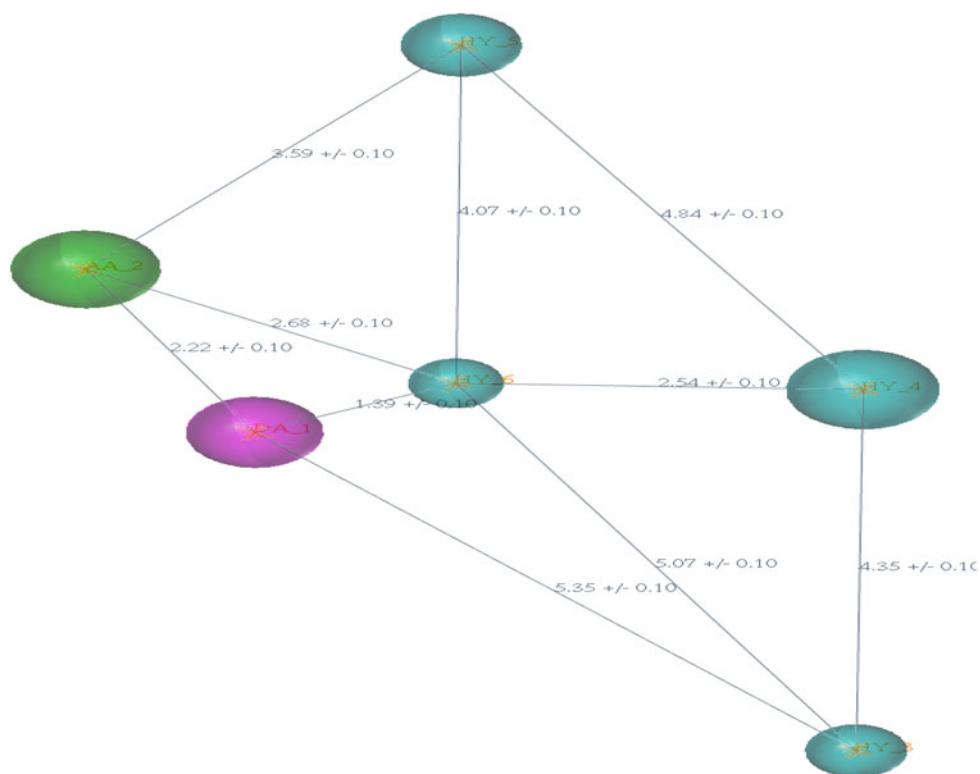
Fig. 4 Alignment of molecules inside the receptor using receptor-guided alignment scheme

Fig. 5 The generated UNITY query for virtual screening against NCI database along with the distance constraints

CoMFA model

The steric and electrostatic field effects were calculated using the Tripos force field with a distance-dependent dielectric constant at all intersections in a regular space (2 Å) grid. A sp^3 carbon atom was used as steric probe and a +1 charge as electrostatic probe. The generated steric and electrostatic fields were scaled by CoMFA-Standard scaling method in SYBYL with default energy cutoff value 30 kcal/mol. With standard options for scaling of variables, the regression analysis was carried out using the full cross-validated PLS method (Cramer et al., 1988a, b; Wold et al., 1984) of leave-one-out (LOO) (Wold, 1978). The minimum sigma (column filtering) was set to 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points whose energy variation was below this threshold. The final model, a non-cross-validated conventional analysis, was developed with the optimum number of components to yield a non-cross-validated r^2 value.

CoMSIA model

The reported CoMSIA method is based on molecular similarity indices (Klebe et al. 1994) with the same lattice box as was used in CoMFA. Molecular similarity is expressed in terms of five different properties, namely steric, electrostatic, hydrophobic, and hydrogen bond donors and acceptors which were calculated using a C^+ probe atom with a radius of 1 Å placed at a regular grid spacing of 2 Å. CoMSIA similarity indices (A_F) for molecule j with atoms i at a grid point q were calculated using Eq. 1.

$$A_{F,K}^q(j) = -\sum \omega_{prob,k}\omega_{ik}e^{-\alpha r^2}iq, \qquad (1)$$

where k represents the following physicochemical properties: steric, electrostatic, hydrophobic, and hydrogen bond donor and acceptor. A Gaussian-type distance dependence was used between grid point q and each atom i of the molecule. The default value (0.3) was used as the attenuation factor (α). The steric indices were related to the third power of the atomic radii, electrostatic descriptors were derived from atomic partial charges, hydrophobic fields were derived from atom-based parameters (Viswanadhan et al., 1989), and hydrogen bond donor and acceptor indices were obtained by a rule-based method based on experimental results (Klebe, 1994).

Partial least square (PLS) analysis and validation of QSAR models

To derive 3D-QSAR models, the CoMFA and CoMSIA descriptors were used as independent variables and the pIC_{50} values as the dependent variable. The PLS method (Wold and Ruhe, 1984; Geladi, 1988) was used to linearly

Table 4 The regression summary of CoMFA and CoMSIA models by different alignment methods

	CoMFA			CoMSIA		
	Alignment 1 (ligand based)	Alignment 2 (pharmacophore based)	Alignment 3 (receptor based)	Alignment 1 (ligand based)	Alignment 2 (pharmacophore based)	Alignment 3 (receptor based)
q^{2a}	0.775	0.734	0.777	0.806	0.750	0.782
r^{2b}	0.951	0.894	0.958	0.927	0.862	0.927
SEE[c]	0.212	0.299	0.196	0.257	0.342	0.258
F value	130.278	122.917	153.910	86.288	90.613	86.132
r^2 pred[d]	0.778	0.631	0.818	0.927	0.688	0.710
No. of compounds	41	41	41	41	41	41
NOC[e]	4	2	4	4	2	4
Contribution						
Steric	0.714	0.448	0.598	0.146	0.089	0.139
Electrostatic	0.286	0.552	0.402	0.347	0.366	0.408
Hydrophobic	–	–	–	0.126	0.127	0.120
Hydrogen bond donor	–	–	–	0.255	0.183	0.259
Hydrogen bond acceptor	–	–	–	0.125	0.235	0.074

[a] Cross-validated correlation after leave-one-out procedure

[b] Correlation coefficient

[c] Standard error of estimate

[d] Predicted correlation coefficient for test set compounds

[e] Optimal number of components

correlate these CoMFA and CoMSIA descriptors to the activity. The CoMFA cutoff values were set to 30 kcal/mol for both steric and electrostatic fields, and all fields were scaled by the default options in SYBYL. The cross-validation analysis was performed using the LOO method in which one compound was removed from the dataset and its activity was predicted using the model derived from rest of the dataset. The cross-validated correlation coefficient (q^2) that resulted in the optimum number of components and the lowest standard error of prediction was considered for further analysis and calculated using Eq. 2.

$$q^2 = 1 - \frac{\sum (Y_{predicted} - Y_{observed})^2}{\sum (Y_{observed} - Y_{mean})^2} \quad (2)$$

where, $Y_{predicted}$, $Y_{observed}$, and Y_{mean} are the predicted, observed, and mean values of the target property (pIC_{50}), respectively. The non-cross-validated PLS analyses were performed with a column filter value of 2.0, to reduce analysis time with small effect on the q^2 values. The predictive power of the developed 3D-QSAR models has been assessed by test set (nine molecules) predictions. The predictive abilities of the models were expressed by the predictive r^2 value, which is analogous to the cross-validated $r^2(q^2)$ and is calculated using Eq. 3.

$$r^2_{pred} = \frac{SD - PRESS}{SD}, \quad (3)$$

$$PRESS = \sum_y (Y_{predicted} - Y_{observed})^2 \quad (4)$$

where SD is the sum of the squared deviations between the biological activities of the test set and mean activities of the training molecules, and PRESS is the sum of squared deviation between the predicted and observed activity of the test set molecules and is calculated using Eq. 4.

Results and discussion

Molecular docking

To predict the appropriate binding conformation for p38 kinase inhibitors and the reported hit compounds from virtual screening, Surflex-Dock was used to generate an ensemble of docking conformations. Before the docking was done, the reliability of Surflex-Dock was evaluated by re-docking the co-crystallized ligand [14e (DQO)] into the binding site. It was observed that Surflex-Dock has successfully reproduced the binding pose. Moreover, the resulted hit compounds from screening after QFIT filtering were further screened using molecular docking into the binding site. The docked compounds were filtered

Table 5 Observed, predicted, and residual values for the test set and training set by CoMFA and CoMSIA models

Compound	pIC_{50}	CoMFA		CoMSIA	
		Predict	Residual	Predict	Residual
9c	7.523	7.540	−0.017	7.571	−0.048
9d	8.155	7.728	0.427	7.705	0.450
9e[a]	7.102	7.499	−0.397	7.463	−0.361
9f	7.658	7.699	−0.041	7.623	0.035
9g[a]	9.000	7.928	1.072	7.710	1.290
9h	7.174	7.267	−0.093	7.444	−0.270
9i	7.328	7.474	−0.146	7.506	−0.178
9k[a]	6.770	7.361	−0.591	7.300	−0.530
9l	6.886	6.901	−0.015	7.173	−0.287
9m	7.959	7.855	0.104	7.626	0.333
9n[a]	6.495	7.384	−0.889	7.559	−1.064
9o	7.357	7.384	−0.027	7.504	−0.147
9p	8.155	7.926	0.229	7.739	0.416
9q	7.357	7.650	−0.293	7.709	−0.352
9r[a]	7.959	8.011	−0.052	7.677	0.282
10a	8.523	8.419	0.104	8.266	0.257
10b	7.959	8.278	−0.319	8.257	−0.298
10c	8.222	8.256	−0.034	8.213	0.009
14a	9.046	9.475	−0.429	9.643	−0.597
14b	9.699	9.440	0.259	9.676	0.023
14c	10.000	9.639	0.361	9.688	0.312
14d	8.854	8.695	0.159	8.630	0.225
14e	8.585	8.639	−0.054	8.671	−0.086
14f	9.699	9.913	−0.214	9.677	0.022
14g[a]	10.000	10.082	−0.082	9.729	0.271
14h	9.301	9.407	−0.106	9.351	−0.050
14i	9.222	9.297	−0.075	9.388	−0.166
14j	10.000	10.046	−0.046	9.946	0.054
14k[a]	9.886	10.144	−0.258	9.990	−0.104
14l	8.824	8.716	0.108	8.637	0.187
14m	8.620	8.643	−0.023	8.678	−0.058
14n	8.959	8.886	0.073	8.688	0.271
15a	9.301	9.124	0.177	9.193	0.109
15b	8.921	8.926	−0.005	9.123	−0.202
15c	9.222	9.370	−0.148	9.129	0.093
15d[a]	8.959	9.074	−0.115	9.166	−0.207
15e	9.301	9.228	0.073	9.321	−0.020
15f	9.523	9.484	0.039	9.570	−0.047
15g[a]	9.398	9.827	−0.429	9.482	−0.084
15h	9.222	9.277	−0.055	9.431	−0.209
15i	9.699	9.670	0.029	9.478	0.221

[a] Test set compounds

based on scoring function and interaction with crucial active site residues (Leu107, Met109, Gly110) in the binding site.

CoMFA and CoMSIA statistical results

The CoMFA and CoMSIA studies were carried out using three different schemes. Results of the PLS analyses of three different alignments are summarized in Table 4. The ligand-based model gave better result for CoMFA with cross-validated $r^2(q^2) = 0.775$ and non-cross-validated $r^2 = 0.951$, while for CoMSIA model, combination of steric, electrostatic, donor, acceptor and hydrophobic terms yielded a cross-validated $r^2(q^2) = 0.806$ and non-cross-validated $r^2 = 0.927$. These models were validated by a test set of nine molecules with predictive $r^2 = 0.778$ for CoMFA model and 0.927 for CoMSIA model. The pharmacophore-based model also gave good results for both CoMFA and CoMSIA. CoMFA yields a cross-validated $r^2(q^2) = 0.734$ and non-cross-validated $r^2 = 0.894$, while CoMSIA model yielded a cross-validated $r^2(q^2) = 0.750$ and non-cross-validated $r^2 = 0.862$. The predictive ability of these models was calculated by an external test set and it gave 0.631 for CoMFA and 0.688 for CoMSIA. In comparison, receptor-guided

alignment gave better statistical results. CoMFA yields a cross-validated $r^2(q^2) = 0.777$ and non-cross-validated $r^2 = 0.958$, while CoMSIA model yielded a cross-validated $r^2(q^2) = 0.782$ and non-cross-validated $r^2 = 0.927$. These models were also validated on a test set of nine molecules with predictive $r^2 = 0.818$ for CoMFA and 0.710 for CoMSIA. It can be seen that all the models are better in statistics, so we decided to explain receptor-based model to better understand the relationship between the inhibitors and the macromolecule. The predicted pIC$_{50}$ value for test and training set from CoMFA and CoMSIA models are given in Table 5. Graphical representations of actual and predicted inhibitory activities of CoMFA and CoMSIA are shown in Figs. 6 and 7.

CoMFA and CoMSIA contour plots

Since the model was built on the basis of receptor-guided scheme, we could overlay the 3D contour maps produced by CoMFA and CoMSIA onto the receptor binding pocket.
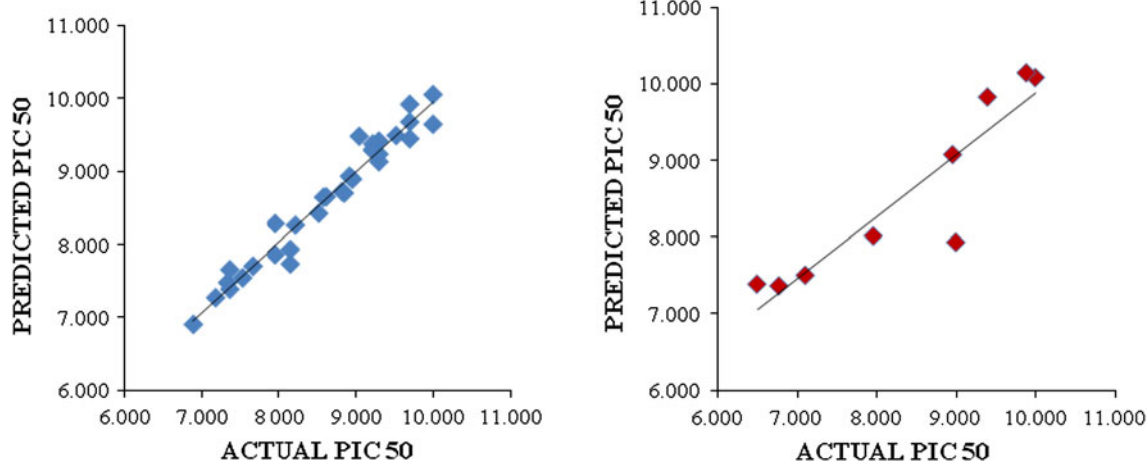

Fig. 6 Trend of actual and predicted activities of training and test set by receptor-based model (CoMFA)
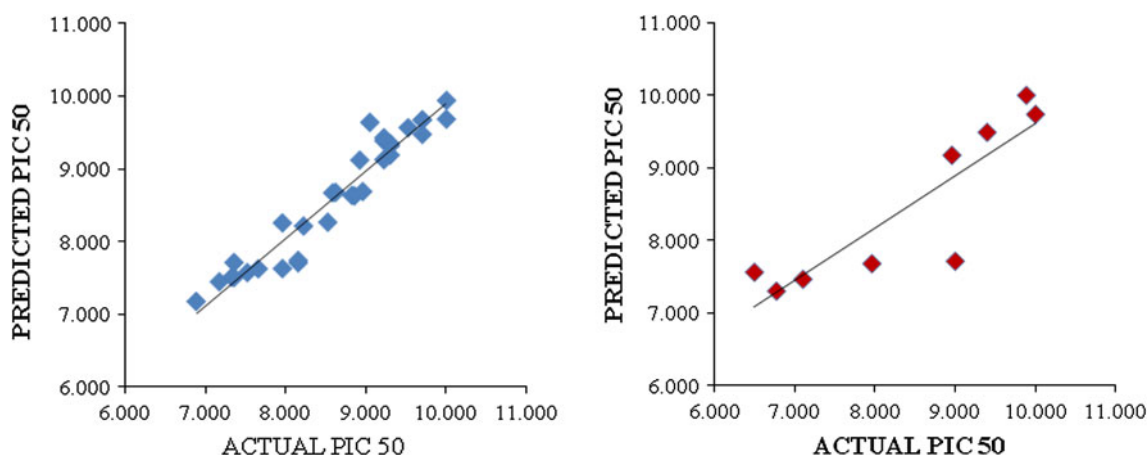

Fig. 7 Trend of actual and predicted activities of training and test set by receptor-based model (CoMSIA)

The results obtained from CoMFA indicate that steric and electrostatic properties play a major role in inhibition activity (Table 4). However, electrostatic properties play a major role in CoMSIA, followed by donor and steric fields. This is because of the fact that the substituents are highly bulky, electropositive, and electronegative in nature. So we explain more on steric and electrostatic field contributions. The steric interactions are represented by green and yellow colored contours while electrostatic interactions are represented by red and blue colored contours. In steric field, green contour represents region where bulky substituent enhances activity, whereas yellow contour indicates region where bulky substituents decrease the activity. In case of electrostatic interactions, the blue contour represents region where electropositive groups enhance the activity, while red-colored region indicates that electronegative groups increases the activity. One of the most active compounds in the series (compound **14j**) is shown with CoMFA contour maps of steric and electrostatic fields in Figs. 8 and 9. The green contour region near the piperidine or piperazine moiety of C7 substituent indicates that bulky substituent at this position increases activity. This is consistent with series 2 and 3 where the bulky substituent shows potent activity, whereas in series 1 the absence of bulky groups shows less activity. The presence of yellow contour near the second and third position of phenyl ring at C5 indicates that steric bulk disfavors activity. This is quite visible that compounds (**9k, 9l, 9n, 9o**) having bulky substitutions at those positions are lower in activity. A blue contour present around the C7 substituent indicates that electropositive and

hydrogen donors are favorable at this position. This blue contour also reflects the fact that electronegative groups are unfavorable at this position. This is quite clear that series 1 having carbonyl substitutions lacks activity. Red contours around the phenyl group of C5 indicate that electronegative groups are favorable. Compounds having chlorine, fluorine at those regions are highly active and it is quite evident throughout the dataset.

The CoMSIA steric and electrostatic contour maps were more or less similar to that of CoMFA. Figures 10 and 11 show compound **14j** superimposed on CoMSIA plots. The presence of green contour near the piperidine or piperazine moiety indicates that bulky substitution at this position enhances the activity. A small yellow contour present at the fourth position of phenyl ring at C5 indicates that bulky position around this region disfavors activity. Compounds such as **9l** and **9o** having bulkier substitutions at these positions lack in activity, whereas compounds having smaller substitutions at this position show better inhibitory activity. This is quite evident throughout the dataset. These results were supported by the presence of red contour around the same region which indicates that electronegative groups such as chlorine and fluorine are favorable.

Evaluation of pharmacophore with crystal complex

The crystal structure of p38α in complex with DQO (PDB code: 1M7Q) was further utilized to evaluate the selected pharmacophore model. It was reported that the complex forms three hydrogen bond interactions with hinge region.

Fig. 8 CoMFA contour maps for steric field with highly active compound (**14j**), where *green contour* indicates regions where bulky groups increase activity and *yellow contours* indicate that bulky groups decrease activity (Color figure online)
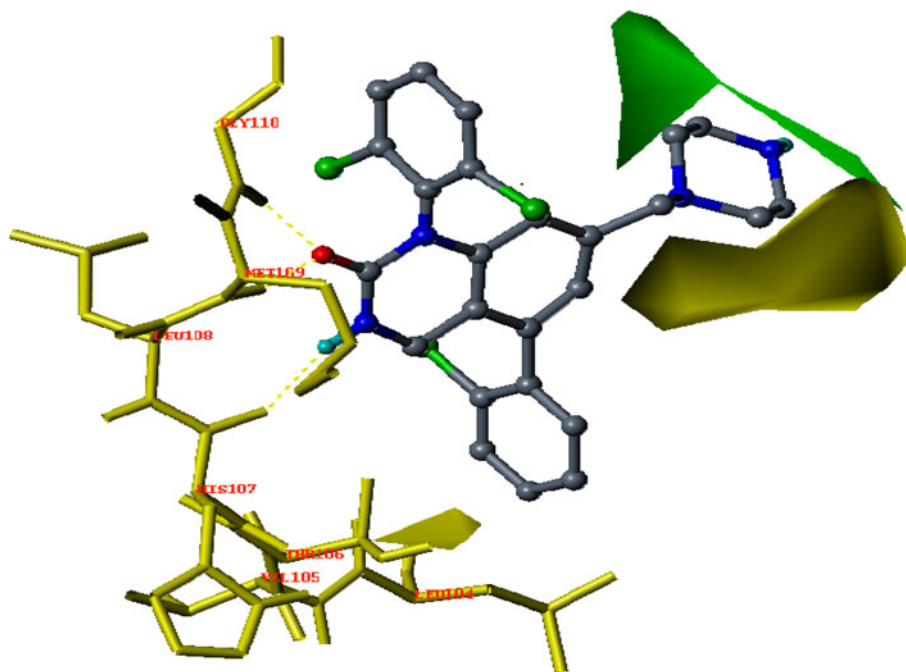
Fig. 9 CoMFA contour maps
for electrostatic field with
highly active compound (**14j**),
where *blue contour* indicates
regions where bulky
electropositive groups increase
activity and *red contours*
indicate regions where
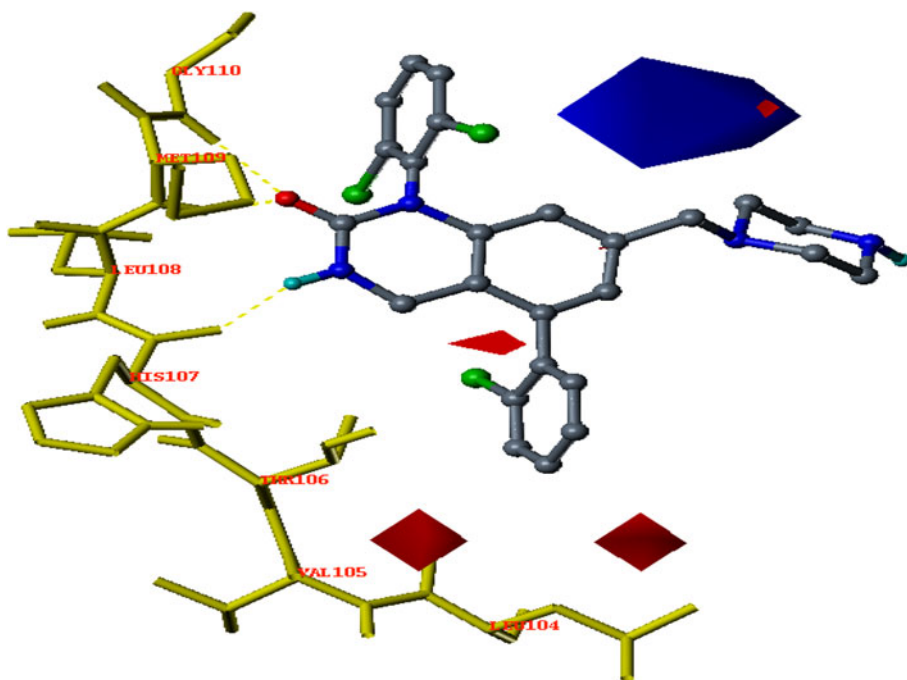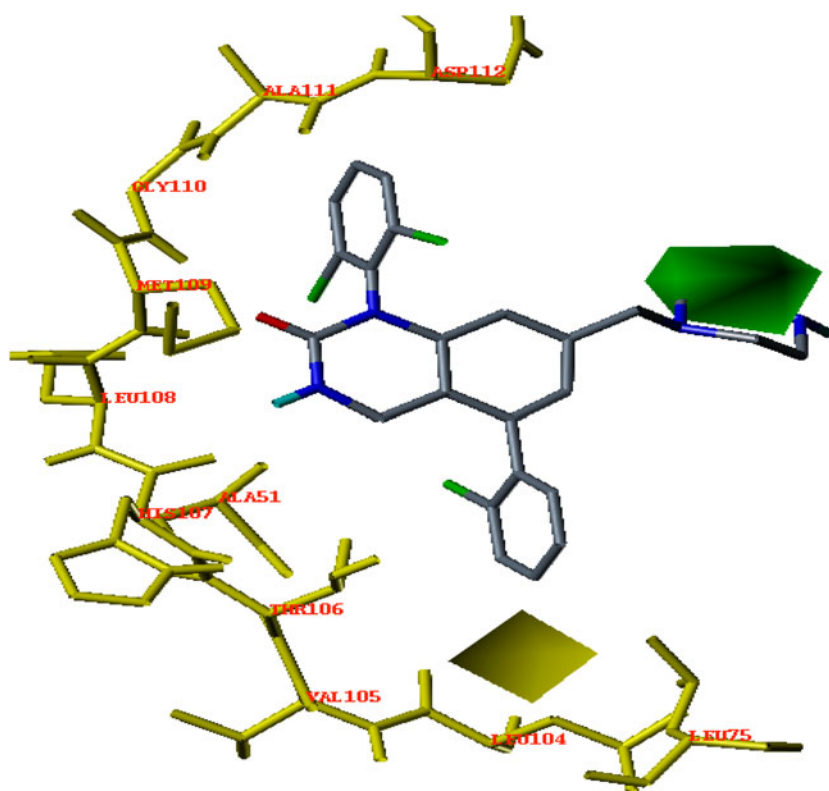electronegative groups increase
activity (Color figure online)

Fig. 10 CoMSIA contour maps
for steric field with highly active
compound (**14j**), where *green
contour* indicates regions where
bulky groups increase activity
and *yellow contours* indicate
that bulky groups decrease
activity (Color figure online)

The complex structure forms hydrogen bonding with Leu107, Met109, and Gly110 (21). The selected pharmacophore model was mapped on to the bound conformation of DQO. The bound conformation fits quite well on to the pharmacophore model. The RMSD between the pharmacophore and DQO is 0.59 Å. From the mapping we can see

Fig. 11 CoMSIA contour maps for electrostatic field with highly active compound (**14j**), where *blue contour* indicates regions where bulky electropositive groups increase activity and *red contours* indicate regions where electronegative groups increase activity (Color figure online)
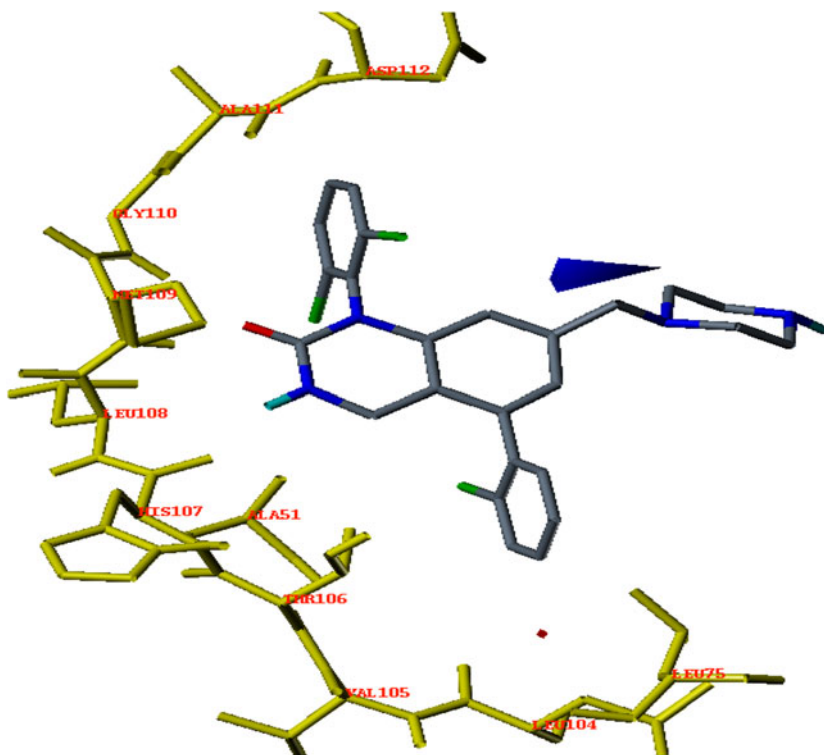
Fig. 12 The generated pharmacophore (*sticks*) was mapped with the co-crystallized ligand DQO (dihydroquinazolinone in *capped sticks*) inside the active site
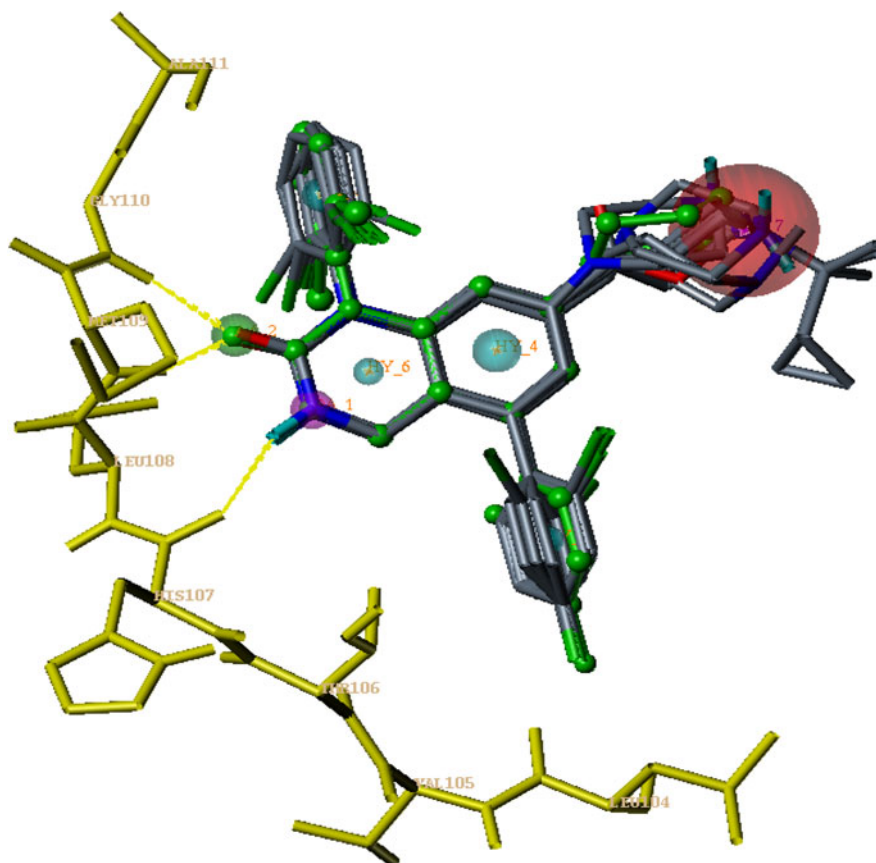
that the pharmacophore features were mapped exactly on the complex, where the co-crystal ligand interacts with the key residues of active site. The mapping of pharmacophore with DQO is shown in Fig. 12. These results suggested that the generated pharmacophore was reliable enough to retrieve compounds from a chemical database that fits all the features of the query.

Common pharmacophore hypothesis and virtual screening

A set of highly active molecules were used to identify the common pharmacophore, GALAHAD produces five models with default parameters. The results are given in Table 6. The resulted models are more or less similar in statistics. All the models show reasonable specificity, which is a logarithmic indicator of the expected discrimination of each query. The number of hits column indicates that all models hit all ligands in the dataset and each of them has seven features in it. Pareto rank indicates that no model is more or less superior to each another. We have selected one model and the model has been validated for subsequent virtual screening analysis. The virtual screening workflow is shown in Fig. 13. The screening of the pharmacophore query yielded 1,182 hits that met the specific requirements. The compounds are further screened on the basis of QFIT, where QFIT is the pharmacophore match between query and the hit. 268 compounds which were further filtered by QFIT were then subjected to molecular docking to the binding site of p38 to select the compounds on the basis of their ability to form favorable interactions with the active site. Finally, seven compounds are selected on the basis of dock score and favorable interaction with key residues. The pIC_{50} values of final hits were predicted using CoMFA and CoMSIA models generated on the basis of pharmacophore alignment. The results of hit compounds with their dock score and predicted biological activity values are shown in Table 7. Thus, these potential hits are expected to induce improved binding affinity with p38α
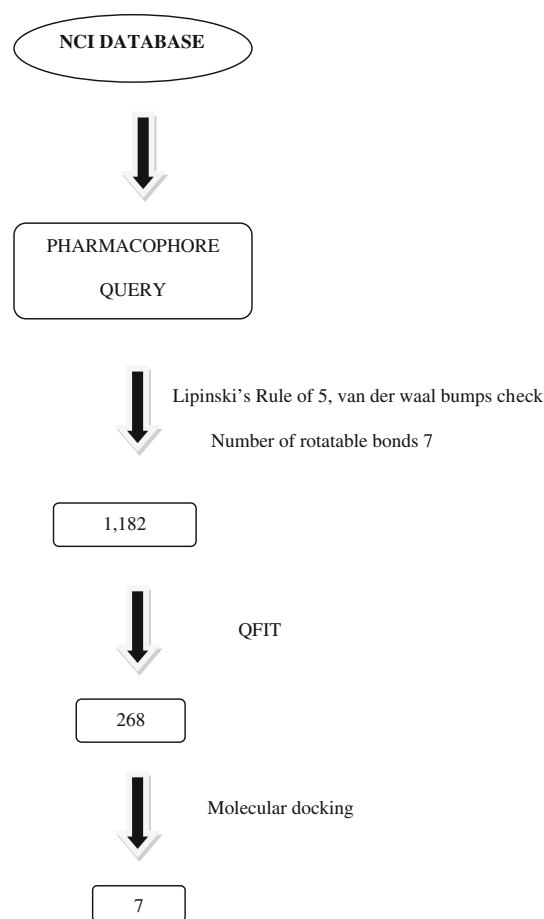

Fig. 13 The results obtained from virtual screening. The *numbers* given in the figure represent the number of molecules after employing the filters

kinases. The hit compound NCI 211823 which has good docking score is shown in Fig. 14. The binding mode of this hit compound is similar to that of co-crystallized compound (**14e**). All the final hit compounds have hydrogen bonding interactions with the hinge residues which is crucial to target the ATP binding site.

Table 6 Number of models obtained along with the pharmacophoric features and their statistical values using GALAHAD

S. no.	Specification	No of hits	Features	Pareto rank	Energy	Sterics	H-bond	Mol_Qry
Model_01	3.747	8	7	0	6.41	3,356.9	199.9	80.37
Model_02	3.749	8	7	0	7.48	3,515.0	199.5	79.83
Model_03	3.746	8	7	0	7.22	3,415.9	199.7	80.06
Model_04	3.748	8	7	0	9.74	3,500.5	199.3	86.90
Model_05	3.748	8	7	0	15.93	3,778.4	195.8	81.27

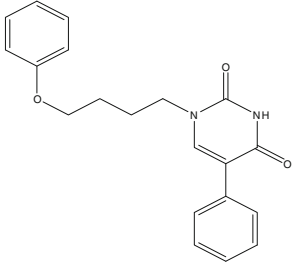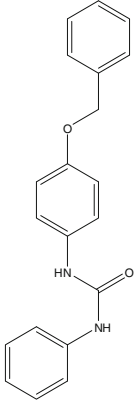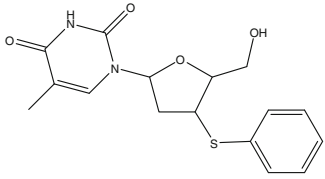**Table 7** Chemical structures of hit compounds, their dock scores, predicted CoMFA and CoMSIA
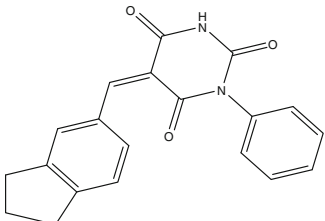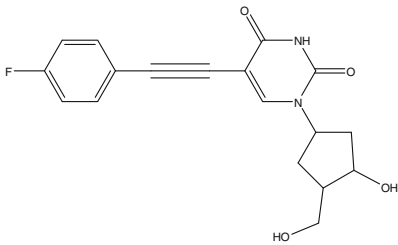
Hit compound	Structure	Dock score	Predicted CoMFA	Predicted CoMSIA
NCI 211823		9.05	9.024	8.436
NCI 321056		8.61	8.119	8.506
NCI 662429		8.03	8.521	8.150
NCI 371720		7.44	7.853	8.376
NCI 648477		7.25	8.318	8.153

Table 7 continued
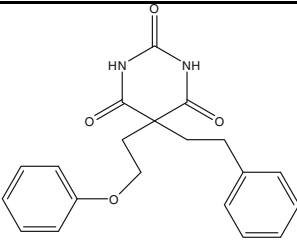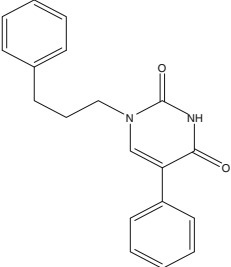
Hit compound	Structure	Dock score	Predicted CoMFA	Predicted CoMSIA
NCI 2103		7.34	8.062	8.057
NCI 107558		7.13	8.870	8.413

Fig. 14 Binding mode and hydrogen bonding pattern of the lead molecule (NCI 211823) obtained from virtual screening along with co-crystallized compound **14e** (*capped sticks*)
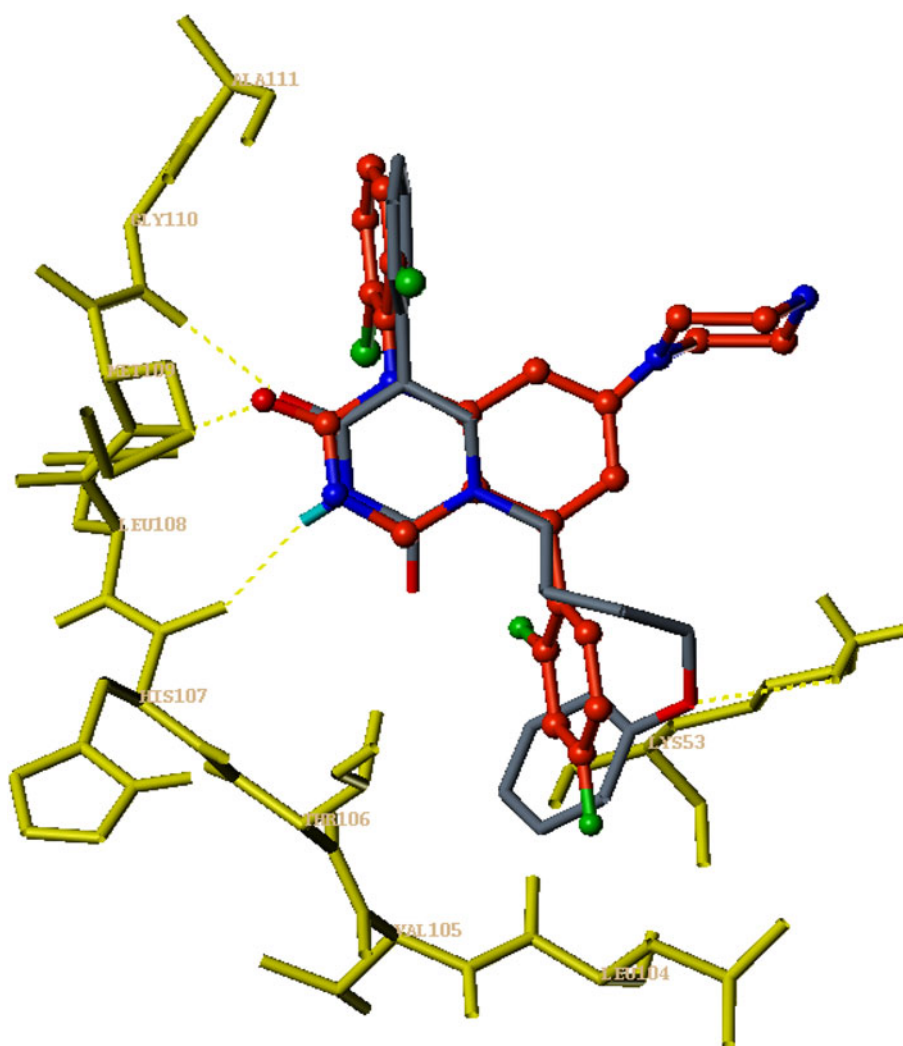
Conclusion

3D-QSAR analysis was performed for a series of potent p38 inhibitors. Ligand-based and receptor-guided alignment schemes were applied to develop models. Receptor-guided CoMFA and CoMSIA were explained, probably because the alignment using receptor information is more realistic. The contour maps revealed that bulky substitutions around piperidine ring of C7 carbon atom and electronegative groups around the phenyl ring of C5 carbon atom are desirable for potent activity. The information obtained can help to design new inhibitors of p38 kinases. We have also performed virtual screening analysis and we predicted the biological activities of obtained hits using the 3D-QSAR model generated by pharmacophore-based alignment. 1,182 hits were obtained using certain filters such as Lipinski's rule of five and number of rotatable bonds to a maximum of seven. These hits were further reduced to 268 compounds using QFIT. Molecular docking was employed as final filter and seven compounds were selected on the basis of dock score, hydrogen bonding interaction, and the biological activities were predicted using CoMFA and CoMSIA. The identified hit compounds were structurally different from already available inhibitors and we suggest that these compounds could serve as potential leads.

References

Andreakos ET, Foxwell BM, Brennan FM, Maini RN, Feldmann M (2002) Cytokines and anti-cytokine biologicals in autoimmunity: present and future. Cytokine Growth Factor Rev 13:299–313

Bohm M, Sturzebecher J, Klebe G (1999) 3D QSAR analyses using CoMFA and CoMSIA to elucidate selectivity differences of inhibitors binding to trypsin, thrombin and factor Xa. J Med Chem 42:458–477

Chen E, Keystone EC, Fish EN (1993) Restricted cytokine expression in rheumatoid arthritis. Arthritis Rheum 36:901–910

Cramer RD III, Patterson DE, Bunce JD (1988a) Comparative molecular field analysis CoMFA. 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110:5959–5967

Cramer RD III, Bunce JD, Patterson DE, Frank IE (1988b) Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. Quant Struct Act Relat 7:18–25

Dinarello CA (1991) Inflammatory cytokines: Interleukin-1 and tumor necrosis factor as effector molecules in autoimmune diseases. Curr Opin Immunol 3:941–948

Dominguez C, Powers DA, Tomayo N (2005) p38 MAP kinase inhibitors: many are made, but few are chosen. Curr Opin Drug Discov Dev 8:421–430

Feldmann M, Brennan FM, Maini RN (1996) Role of cytokines in rheumatoid arthritis. Cell 85:307–310

Geladi P (1988) Notes on the history and nature of partial least squares (PLS) modelling. J Chemometr 2:231–246

Jiang Y, Chen C, Li Z, Guo W, Gegner JA, Lins S, Han J (1996) Characterization of the structure and function of a new mitogen-activated protein kinase (p38β). J Biol Chem 271:17920–17926

Jiang Y, Gram H, Zhao M, New L, Gu J, Feng L, Padova F, Ulevitch R, Han J (1997) Characterization of the structure and function of the fourth member of p38 group mitogen-activated protein kinases, p38delta. J Biol Chem 272:30122–30128

John ES, Luping L, Sangita BP, James VP, Giovanna S, Suresh S, Cornelis ECAH, Zheng W, John RS, Patricia MC, Elizabeth AN, Stephen JO, Edward AO, Denniss MS, Cheryl DS, Chris MT, Dennis MZ, James BD (2003) Design and synthesis of potent, orally bioavailable dihydroquinazolinone Inhibitors of p38 MAP kinases. Bioorg Med Chem Lett 13:277–280

Jones G, Willett P, Glen RC (1995) A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. J Comput Aided Mol Des 9:532–549

Klebe G (1994) The use of composite crystal-field environments in molecular recognition and the de novo design of protein ligands. J Mol Biol 237:212–235

Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37:4130–4146

Kulkarni RG, Sastry GN, Achaiah G (2006) Novel targets for anti-inflammatory and anti-arthritic agents. Curr Pharm Des 12:2437–2454

Lee JC, Laydon JT, McDonnell PC, Gallagher TF, Kumar S, Green D, McNulty D, Blumenthal MJ, Heys JR, Landvatter SW (1994) A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. Nature 372:739–746

Li Z, Jiang Y, Ulevitch RJ, Han J (1996) The primary structure of p38γ: a new member of p38 group of MAP kinases. Biochem Biophys Res Commun 228:334–340

Martin YC, Bures MG, Danaher EA, DeLazzer J, Lico I, Pavlik PA (1993) A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. J Comput Aided Mol Des 7:83–102

Newton R, Holden N (2003) Inhibitors of p38 mitogen-activated protein kinase: potential as anti-inflammatory agents in asthma? BioDrugs 117:113–129

Patel Y, Gillet VJ, Bravi G, Leach AR (2002) A comparison of the pharmacophore identification programs: catalyst, DISCO and GASP. J Comput Aided Mol Des 16:653–681

Pearson G, Robinson F, Gibson TB, Xu B, Karandikar M, Berman K, Cobb MH (2001) Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. Endocr Rev 22:153–183

Pincus T (1995) Long-term outcomes in rheumatoid arthritis. Br J Rheumatol 34:59–73

Ruppert J, Welch W, Ajay NJ (1997) Automatic identification and representation of protein binding sites for molecular docking. Protein Sci 6:524–533

Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. J Chem Inf Comp Sci 29:163–172

Willett P, Winterman V (1986) A comparison of some measures of intermolecular structural similarity. Quant Struct Act Relatsh 5:18–25

Wold S (1978) Cross-validatory estimation of the number of components in factor and principal component model. Technometrics 20:397–405

Wold S, Ruhe A (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM J Sci Stat Comput 5:735

Wold S, Albano C, Dunn W, Edlund U, Esbensen K, Geladi P, Hellberg S, Johansson E, Lindberg W, Sjostrom M (1984) In: Kowalski BR (ed) Chemometrics mathematics and statistics in chemistry. Reidel publishing company, Dordrecht, p 17–95