

Internet **Electronic** Journal of **Molecular Design**

December 2007, Volume 6, Number 12, Pages 385–395

Editor: Ovidiu Ivanciuc

Comparative Molecular Field Analysis (CoMFA) for p38 Inhibitors

Munuganti Ravi Shashi Nayana¹ and Yadavalli Nataraja Sekhar¹

¹ Bioinformatics Division, Environmental Microbiology Lab, Department of Botany, Osmania University, Hyderabad 500007, Andhra Pradesh, India

Received: July 14, 2007; Revised: October 18, 2007; Accepted: November 27, 2007; Published: December 31, 2007

Citation of the article:

M. R. S. Nayana and Y. N. Sekhar, Comparative Molecular Field Analysis (CoMFA) for p38 Inhibitors, *Internet Electron. J. Mol. Des.* **2007**, *6*, 385–395, <http://www.biochempress.com>.

Comparative Molecular Field Analysis (CoMFA) for p38 Inhibitors

Munuganti Ravi Shashi Nayana^{1,*} and Yadavalli Nataraja Sekhar¹

¹ Bioinformatics Division, Environmental Microbiology Lab, Department of Botany, Osmania University, Hyderabad 500007, Andhra Pradesh, India

Received: July 14, 2007; Revised: October 18, 2007; Accepted: November 27, 2007; Published: December 31, 2007

Internet Electron. J. Mol. Des. 2007, 6 (12), 385–395

Abstract

Motivation. The p38 kinase plays a vital role in the inflammation mediated by tumor necrosis factor- α (TNF- α) and interleukin- 1β (IL- 1β) pathways and inhibitors of p38 kinase provide effective approach for the treatment of inflammatory diseases. Three dimensional quantitative structure activity relationship (3D QSAR) studies involving comparative molecular field analysis (CoMFA) were performed on 38 benzimidazolone derivatives as p38 kinase inhibitors.

Method. CoMFA was performed using Sybyl software 6.7v.

Results. Compounds were divided in to training and test set. The developed model based on training set containing 28 compounds gave leave one out cross validation q^2 value of 0.648 and non cross validation r^2 value of 0.980 and Standard error of estimate 0.119 for CoMFA. The steric and electrostatic contributions are 61.6% and 38.4%. Additionally, the binding mode of the high active compound at the active site of p38 MAP alpha kinase was explored using LigandFit docking program and hydrogen-bonding interactions were observed between the inhibitor and the target. The details of amino acid interactions of the active site are discussed briefly.

Conclusions. The CoMFA model provided the most significant correlation of steric and electrostatic fields with biological activities. The information rendered by 3D QSAR model may afford valuable clues to optimize the lead and design new potential inhibitors.

Keywords. QSAR; quantitative structure-activity relationships; CoMFA; comparative molecular field analysis; p38 kinase; benzimidazolone derivatives.

Abbreviations and notations

CoMFA, comparative molecular field analysis
LOO, leave-one-out

QSAR, quantitative structure-activity relationships
PLS, partial least squares

1 INTRODUCTION

Rheumatoid arthritis (RA) is characterized by the chronic inflammation of joints that lead to destruction of cartilage and deformation of bones [1,2]. The p38 MAP kinase regulates the release from leukocytes of IL-1 and TNF- α , two cytokines that are associated with the progression of rheumatoid arthritis (RA) [3]. Inhibitors of the p38 α through their downstream blockage of the

* Correspondence author; phone: +919985795735; E-mail: rsnayana@indiatimes.com.

production of TNF- α , IL-1 β , IL-6, COX-2 and arachidonic acid mobilization have therapeutic potential [4,5]. These inhibitors not only block the synthesis but also the signal cascades induced by these cytokines [6] and prevent activation of caspases and apoptosis of neuronal cells and neuronal progenitor cells [7]. The interest in the development of p38 kinase inhibitors is based on the expectation that p38 inhibiting drugs will treat the underlying cause of chronic inflammatory disease and cease their progression [8]. Since drug discovery process is in the new direction involving various computational approaches including 3D QSAR tools like CoMFA, which have been increasingly employed in rational drug discovery process to understand the drug receptor interaction and to design new molecules. A QSAR model establishes a statistical relationship between the biological activity exerted by a series of compounds and a set of parameters determined from their structures. The factors contributing to the biological activity can be understood through use of different physicochemical descriptors in the generation of QSAR models [9].

Two components are essential for QSAR models; the first component is the computation of the structural descriptors from the three-dimensional molecular structure. Various geometrical, quantum, or molecular field descriptors were proposed in recent years to substitute the Hansch substituent constants. The second component in a 3D QSAR model is an explicit mathematical structure activity relationship established between a dependent variable (biological activity) and a set of independent variables (3D structural descriptors); the mathematical 3D QSAR equations can be computed with the help of a large number of statistical models, such as multilinear regression, partial least squares (PLS), or neural networks. Some 3D QSAR models also contain a third component, a graphical representation of the three-dimensional information relative to the ligand-receptor interactions encoded into the structure-activity equation [10]. An important component of all QSAR models is a proper validation and evaluation of the prediction power. The modeling competition CoEPrA (comparative evaluation of prediction algorithms) was specially devised to compare QSAR models based on blind predictions [11]. The present study is aimed to gain insight into the steric and electrostatic properties of these compounds, their influence on the activity and to derive predictive 3D-QSAR models to design new class of inhibitors.

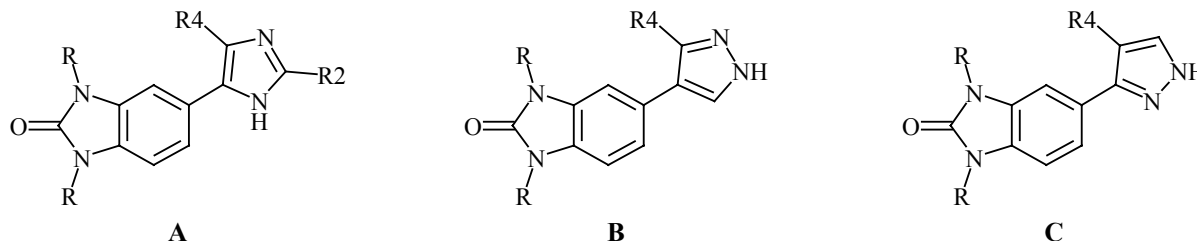
2 MATERIALS AND METHODS

2.1 Dataset and Molecular Modeling

For the current 3D QSAR studies of p38 kinase inhibitors, a reported data set of benzimidazolone derivatives were considered to derive 3D QSAR models [4]. The reported IC₅₀ values were converted into the corresponding pIC₅₀ using the following formula $pIC_{50} = \log IC_{50}$ (Table 1). The data set of 38 molecules was grouped into training set and test set containing twenty-eight and ten molecules with the following rules: (a) both training and test sets should cover structural diversity; (b) both the sets should cover the molecular bioactivities as wide as possible (c)

the most active and worst active compounds should be included in the training set because they provide critical information. All the molecular modeling calculations were performed using SYBYL programming package version 6.7 [12]. The structures were assigned with Gasteiger [13] partial atomic and further geometric optimization of these compounds was done using the semi-empirical program MOPAC 6.0 and applying the AM1 Hamiltonian [14].

Table 1. Structures and Activities of Benzimidazolone Derivatives



Comp.	Lit. [4]	R	R ₂	R ₄	IC ₅₀ (μM)	pIC ₅₀
1A	12b	Me	H	Ph	0.29	6.537
2A	12c	Et	H	Ph	0.14	6.853
3A	12d	Me	H	3-Me-Ph	0.10	7.000
4A	12e	Et	H	3-Me-Ph	0.05	7.301
5A	12f	Me	H	4-F-Ph	0.14	6.553
6A	12g	Et	H	4-F-Ph	0.10	7.000
7A	12h	Me	2-Thiophene	Ph	0.26	6.585
8A	12i	Me	2-Thiophene	2-F-Ph	0.37	6.431
9A	12j	Me	2-Thiophene	3-F-Ph	0.39	6.408
10A	12k	Me	2-Thiophene	4-F-Ph	0.27	6.568
11A	12l	Me	2-Thiophene	2-Cl-Ph	0.54	6.267
12A	12m	Me	2-Thiophene	3-Cl-Ph	0.38	6.420
13A	12n	Me	2-Thiophene	4-Cl-Ph	0.68	6.167
14A	12p	Me	2-Thiophene	3-CF ₃ -Ph	1.5	5.823
15A	12q	Me	2-Thiophene	2-Me-Ph	6.2	5.207
16A	12r	Me	2-Thiophene	3-Me-Ph	0.14	6.583
17A	12s	Me	2-Thiophene	4-Me-Ph	1.8	5.744
18A	12t	Me	2-Thiophene	3-Et-Ph	0.5	6.301
19A	12u	Me	2-Thiophene	4-Et-Ph	8.0	5.096
20A	12w	Me	3-Pyridyl	3-Me-Ph	0.20	6.698
21A	12x	Et	3-Pyridyl	3-Me-Ph	0.07	7.154
22A	12y	Et	4-Pyridyl	3-Me-Ph	0.07	7.154
23A	12z	Me	4-Methane Sulfinyl-Ph	4-F-Ph	4.0	5.397
24B	16a	Et	–	4-F-5-Me-Ph	0.03	7.522
25B	16b	Et	–	3-CF ₃ -Ph	0.06	7.221
26B	16c	Me	–	Ph	0.2	6.698
27B	16d	Et	–	Ph	0.20	6.980
28B	16e	Et	–	3-Me-Ph	0.009	8.045
29C	17a	Me	–	Ph	2.50	5.602
30C	17b	Et	–	Ph	0.68	6.167
31C	17c	Et	–	3-Me-Ph	0.73	6.136
32	11				14.4	4.841

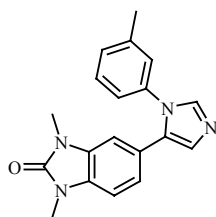
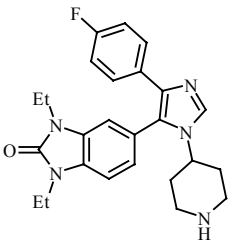
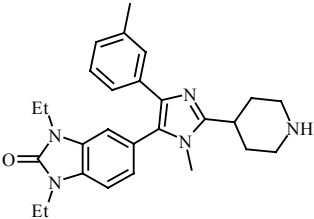
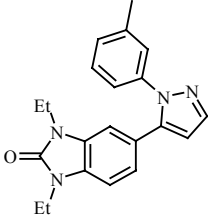
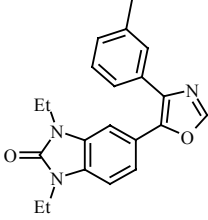
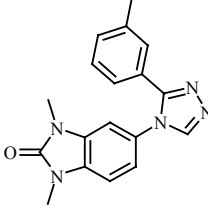
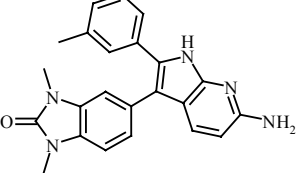


Table 1. (Continued)

Comp.	Lit. [4]	Compound	IC ₅₀ (μM)	pIC ₅₀
33	13		0.090	7.045
34	14		1.25	5.903
35	18		0.69	6.161
36	19		0.80	6.096
37	22		12.5	4.903
38	24		0.26	6.585

2.2 Alignment

Structural alignment is the most sensitive and vital part since the interaction energies depend upon the positioning of molecules in 3D fixed lattice. In the present study the MOPAC geometry optimized structures were aligned on the template **28B**, which, is the most active molecule among the given set. The aligned molecules are shown in Figure 1.

2.3 CoMFA

In CoMFA, Lennard–Jones (6–12 interactions), the steric interaction field and Columbic electrostatic potentials ($1/r$) were calculated at each lattice intersection. The CoMFA potential fields were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. The grid box dimensions were determined automatically in such a way that region boundaries were extended beyond 4 Å in each direction from co–ordinates of each molecule. The van der Waals potentials and Coulombic terms, which represent steric and electrostatic fields, respectively, were calculated using Tripos force field. An sp^3 hybridized carbon atom with radius 1.52 Å bearing +1 charge served as probe atom to calculate steric and electrostatic fields.

2.4 Partial Least Square (PLS) Analysis

Partial Least Square analysis implemented in SYBYL was employed to obtain correlation between the descriptors derived by CoMFA which were used as explanatory variables while pIC_{50} values as target dependant variables [15,16]. The minimum sigma (column filtering) was set to 2.0 kcal/mol accelerate the regression analysis and to reduce the noise. Initially, leave one out (LOO) cross validation method was carried out to check the predictivity of the derived model and to determine the optimum number of components with minimum standard error of estimate. In LOO method, one molecule is omitted from the data set and a model is derived involving the rest of the molecules and using this model the activity of omitted molecule is predicted. The q^2 is calculated by the following equation

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

The predictive correlation coefficient (r^2_{pred}) based on the test molecules, is computed by using formula

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

where, SD is the sum of the squared deviations between the biological activities of the test set and mean activities of training set molecules and PRESS is the sum of squared deviation between predicted and actual activity for every molecule in test set.

2.5 Molecular Docking

For the docking, an advanced molecular docking program Ligand Fit was used in this study. It is a shape–directed docking methodology for accurately docking ligands to protein active sites. The method uses a site detection algorithm for identifying candidate active sites within the protein. A Monte Carlo conformational search procedure is used for generating candidate ligand docking conformations. A shape comparison filter is then used to evaluate each ligand conformation against the active site shape. Ligand conformations satisfying the shape comparison filter are initially

docked into the active site via a shape alignment protocol based on principal axes and moments and are further refined via a grid-based energy calculation which rapidly evaluates protein–ligand interaction energies. A novel method of dramatically reducing the errors arising from interpolation of the grid energies has been presented resulting in an accurate energy evaluation. The docking method also employs a rapid rigid body minimization of the ligand with respect to the grid-based interaction energy. We have employed LigScore scoring function to prioritize docked ligands, which has been developed by using the Genetic Function Approximation [17]. For our studies, X-ray crystal structure of p38 kinase was taken from PDB entry 1ZZL, having resolution of 2.0 Å. Solvent molecules were deleted and bond order for crystal ligand and protein were adjusted.

3 RESULTS AND DISCUSSION

The CoMFA method was used for deriving 3D-QSAR model for 38 benzimidazolone compounds, which are reported as p38 kinase inhibitors. The CoMFA PLS yielded cross-validated q^2 -value of 0.648 and non-cross validated correlation coefficient r^2 of 0.980. The steric and electrostatic contributions are 61.6% and 38.4%. Table 2 lists experimental activities, estimated activities and residual values of the training set and test set by CoMFA model. The statistical parameters for the developed CoMFA model are presented in Table 3. These correlation coefficients suggest that our model is reliable and accurate. Figure 2 shows correlation between the experimental and predicted pIC_{50} values of training and test sets by this CoMFA model.

Table 2. Actual and calculated activities of compounds used in training and test set

Compound	Exp. IC_{50}	Pred. pIC_{50}	Residual				
1A	6.537	6.419	0.118	21A	7.154	7.003	0.151
2A	6.853	6.964	-0.111	22A	7.154	6.655	0.499
3A	7.000	7.109	-0.109	23A	5.397	5.839	-0.442
4A	7.301	7.329	-0.028	25B	7.221	7.633	-0.412
6A	7.000	6.608	0.392	26B	6.698	6.616	0.082
9A	6.408	6.459	-0.050	27B	6.980	7.430	0.450
12A	6.420	6.418	0.002	28B	8.045	7.050	0.995
13A	6.167	5.848	0.319	29C	5.602	5.659	-0.057
15A	5.207	5.835	-0.628	30C	6.167	5.917	0.250
16A	6.853	6.578	0.275	31C	6.136	6.194	-0.058
17A	5.744	5.940	-0.196	32	4.841	5.826	-0.985
18A	6.301	6.283	0.018	33	7.045	6.532	0.513
19A	5.096	6.015	-0.919	35	6.161	6.022	0.139
20A	6.698	6.780	-0.082	36	6.096	6.514	-0.418
Test set							
5A	6.853	6.500	0.353	14A	5.823	6.656	-0.833
7A	6.585	6.142	0.443	24B	7.522	7.288	0.234
8A	6.431	5.415	1.016	34	5.903	6.114	-0.211
10A	6.568	6.242	0.326	37	4.903	5.834	-0.931
11A	6.267	5.860	0.407	38	6.585	6.678	-0.093

The statistical indices presented in Table 3 are: q^2 , LOO cross-validated correlation coefficient; r^2 , non-cross validated correlation coefficient; r^2_{pred} , predictive correlation coefficient; n , number of

components used in the PLS analysis; SEE, standard error of estimate; F-value, F-statistic for analysis.

Table 3. CoMFA PLS Result Summary

Index	CoMFA
q^2	0.648
r^2	0.980
r^2_{pred}	0.765
n	6
F-value	173.835
SEE	0.119
Steric	0.616
Electrostatic	0.384

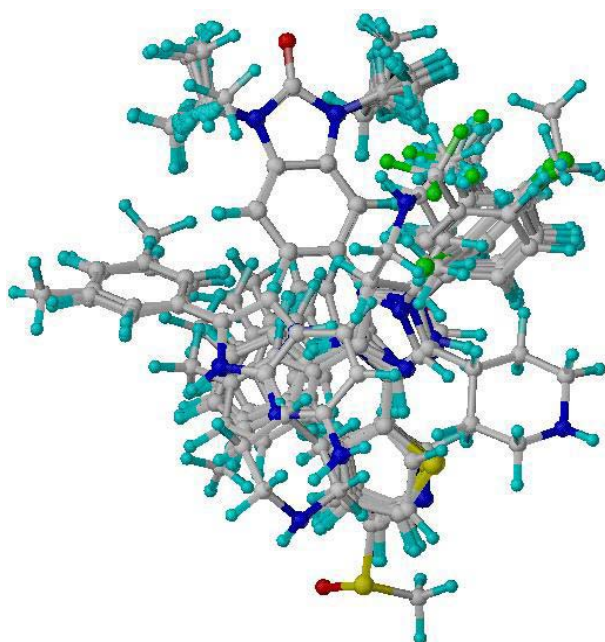


Figure 1. Alignment of all molecules used for CoMFA.

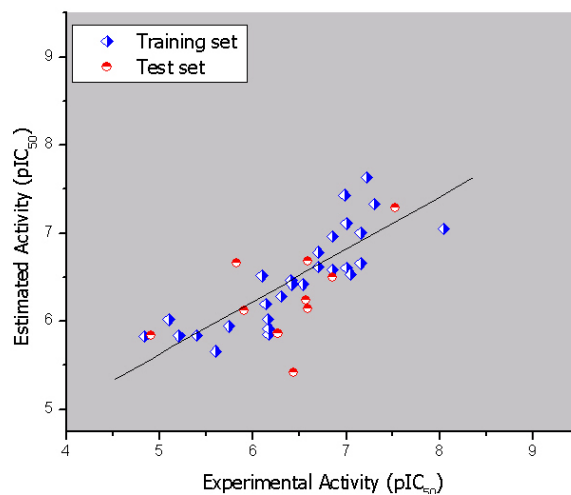


Figure 2. Plot of $pIC_{50,exp}$ versus $pIC_{50,calc}$ for CoMFA.

3.1 3D Contour Maps

To visualize the information content of the derived 3D-QSAR model, CoMFA contour maps were generated. The contour plots are the representation of the lattice points and the difference in the molecular field values at lattice points is strongly connected with difference in the receptor binding affinity. Molecular fields define the favorable or unfavorable interaction energies of aligned molecules with a probe atom traversing across the lattice grid points surrounding the molecules. The 3D colored plots suggest the modification required to design new molecules. The steric interactions are represented by green and yellow colored contours whereas electrostatic interactions are displayed as red and blue contours. Figure 3 and 4 (panel A and B) shows the contour maps derived from the CoMFA PLS model. Highly active compound **28B** (figure 3 A and B) and low active compound **19A** (figure 4 A and B), were embedded in the maps to demonstrate its affinity for the steric and electrostatic regions of inhibitors.

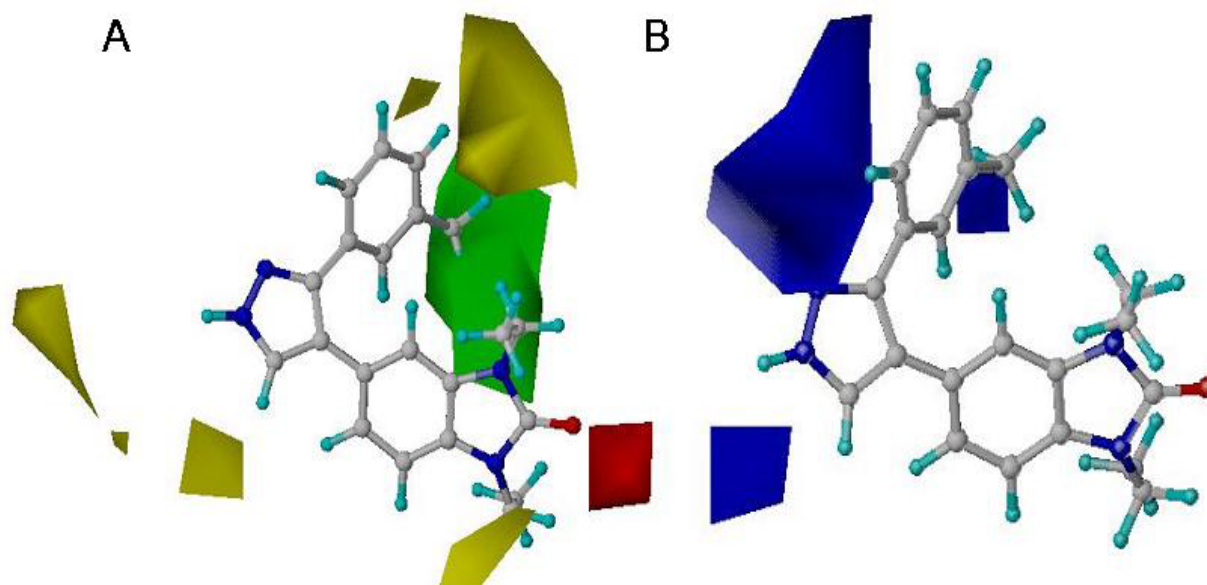


Figure 3. CoMFA contour maps for high active p38 alpha inhibitor (A) the steric field distribution and (B) the electrostatic field distribution for high active compound 28B. Green contour maps for sterically favored areas and sterically disfavored areas in yellow. Positive potential favored areas in blue; negative potential favored areas in red.

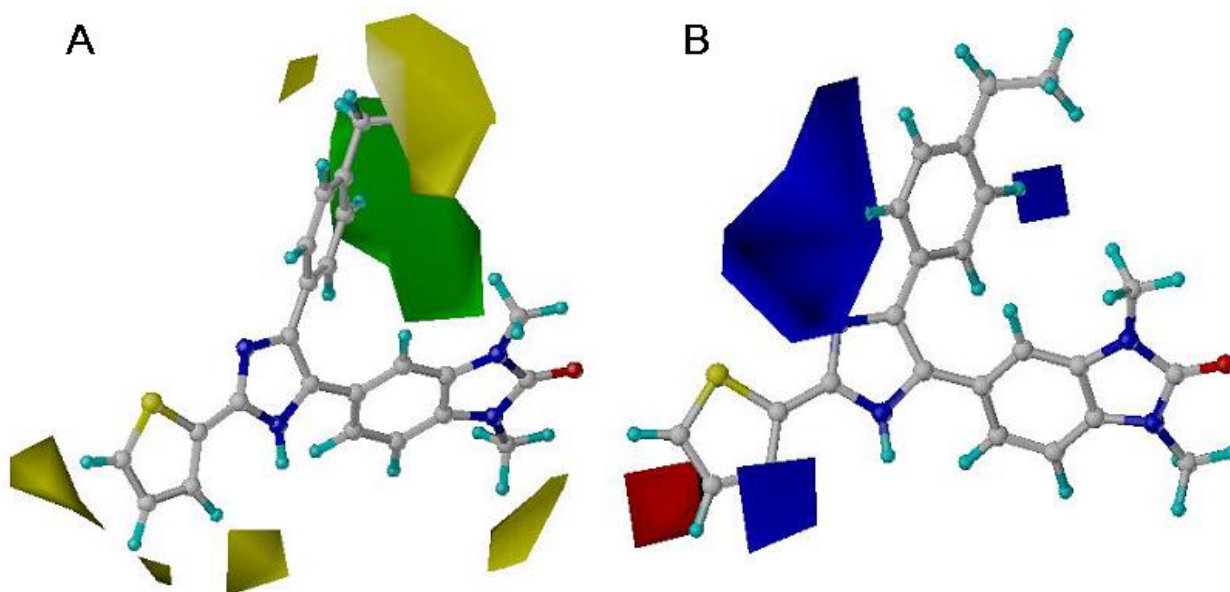


Figure 4. (A) The steric field distribution and (B) the electrostatic field distribution for low active compound 19A. Green contour maps for sterically favored areas and sterically disfavored areas in yellow. Positive potential favored areas in blue; negative potential favored areas in red.

3.2 Steric Interactions

In Figure 3 and 4, panel A shows the steric contours of CoMFA model. Yellow contours indicate regions of steric hindrance to activity, while green region indicate a steric contribution to potency (80% and 20% contribution). Big green plot situated near the 3rd position of the R₄ ring favor bulky groups for the activity. This was observed in the compounds **21A**, **22A**, **24B**, **25B** and **28B**. Presence of big yellow map in proximity to *ortho*-position of the phenyl group (R₄) indicates that bulky substituents may decrease the activity. This observation is confirmed by the fact that *ortho*-

methyl/ethyl phenyl containing ligands (**17A** and **19A**) display lower inhibitory potency. In Scaffold A compounds, thiophene group of **14A**, **15A**, **17A** and **19A** is pointed towards medium sized yellow contours, hence, has less activity.

3.3 Electrostatic Interactions

The electrostatic effects of the substituents are interpreted by the presence of blue and red colored plots. In figure 3 and 4, panel B shows the electrostatic contours of CoMFA model. The blue regions indicate positive electrostatic charge potential associated with increased activity, while regions of red show negative charge with increased activity. The large blue contour around the imidazole ring suggests that biological activity can be enhanced by introduction of more electropositive groups at this position. The conformational orientation of the imidazole ring towards the large blue contour map has lead to augmentation in the inhibitory potency. This fact is confirmed by observing the compounds **22A**, **24B**, **28B** and **33**. Rest of the blue plots endorses pyridyl and piperidyl groups of the compounds **21A**, **22A** and **33**. This is why these compounds show better activity than compounds **31C** and **37**. The red contour map in the vicinity of thiophene group signifies the presence of a negative charge to enhance the biological activity. Electronegative fragments at this position may increase the binding affinity.

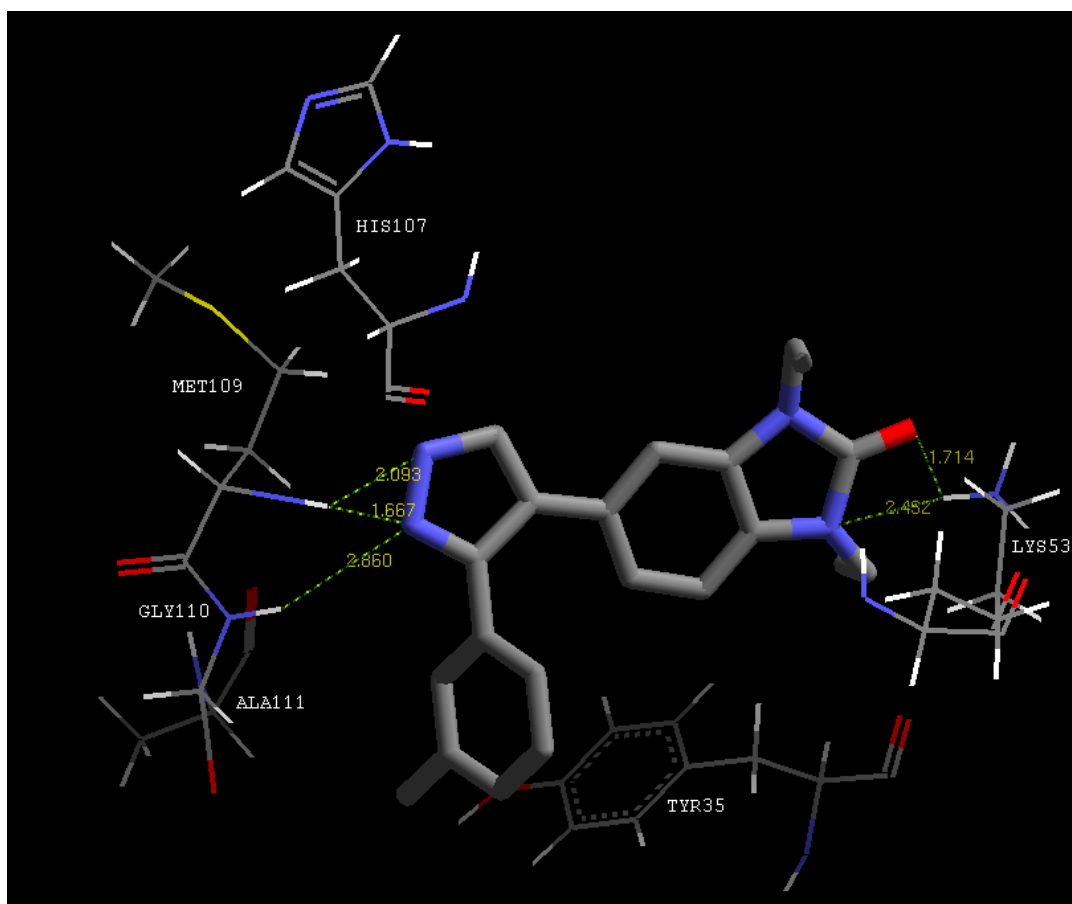


Figure 5. Hydrogen bond interactions between p38 binding pocket and compound 28B.

3.4 Docking Studies

Figure 5 shows the docking of highly active molecule **28B** in to the active site of p38. A high level of selectivity for specific kinase target, the inhibitors should have interactions with backbone amino acids present in the hinge region of protein [18]. The molecule is forming three strong H-bonds with Hinge region amino acids. Two electron rich nitrogen of pyrazole ring are forming three hydrogen bonds with main chain NH of amino acids Gly110, Met109 and Gly110, Ala111 present in the hinge region. Another two hydrogen bonds are formed between electron rich nitrogen of benzimidazolone and NH of Lys53.

4 CONCLUSIONS

3D QSAR analyses have been performed on 38 benzimidazolone derivatives as inhibitors for p38 alpha kinase inhibitors using CoMFA. Stable and statistically reliable predictive model of CoMFA have shown significant q^2 , r^2 and r^2_{pred} (0.648, 0.980 and 0.765, respectively), indicating a better statistical relationship between the activity and descriptors. Based on this model activities of ten test set molecules have been predicted and this model exhibited good prediction correlation with acceptable error range. The CoMFA results suggest that steric interactions (61.6%) as well as electrostatic interactions (38.4%) contribute to the activities of inhibitors. The 3D contour plots derived from the CoMFA and docking provided vital clues that can be used to design new molecules with improved kinase inhibitory activity.

Acknowledgment

Authors are thankful to the UGC Govt. of India for financial support. We are thankful to the management of PGRRCDE, Osmania University for providing the software facility and for continuous support to carry out this work.

Supplementary Material

An sd file containing the structures and activities of 38 benzimidazolone derivatives is available as supplementary material.

5 REFERENCES

- [1] E. Chen, E.C. Keystone, E.N. Fish, Restricted cytokine expression in rheumatoid arthritis, *Arthritis. Rheum.* **1993**, *36*, 901–910.
- [2] T. Pincus, Long-term outcomes in rheumatoid arthritis, *Br. J. Rheumatol.* **1995**, *34*, 59–73.
- [3] A.H. Julianne, K. Florida D.R. Rowena J.S. Peter, Ida Ita, X.M. Sherrrie V.P. James, E.C.A. Cornelis Hop; K. Sanjeev, W. Zhen J.O. Stephen, A.O Edward, P. Gene, E.T. James, W. Andrew, M.Z. Dennis, B.D. James, p38 Inhibitors: piperidine- and 4-aminopiperidine-substituted naphthyridinones, quinolinones, and dihydroquinazolinones, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 467.
- [4] A.D. Mark, A.L. Michael, F.M.C. Kim, T.B. John, J.C. Thomas, R.C. Santo, C. Csilla, J.D. Alan, C.E. Nancy, A.G. Christopher, K.J. Crystal, M.L. Jeff, H.M. William, M.P. Kevin, A.S. Ingrid, S. Linne, J.S. Francis, H.Y. Chul, Benzimidazolone p38 inhibitors, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 919–923.
- [5] J.L. Adams, A.T. Badger, S. Kumar, J.C. Lee, A novel Pd-catalyzed cyclization reaction of ureas for the synthesis of dihydroquinazolinones p38 kinase inhibitors, *Prog. Med. Chem.* **2001**, *38*, 1.

- [6] S. Kumar, P.C. McDonnell, R.J. Gum, A.T. Hand, J.C. Lee, P.R. Young, Novel homologues of CSBP/p38MAP Kinase: Activation, substrate specificity and sensitivity to inhibition by pyridinyl imidazoles, *Biochem. Biophys. Res. Commun.* **1997**, *235*, 533.
- [7] J. Harada, M. Sugimoto, An inhibitor of p38 and JNK MAP kinases prevents activation of caspase and apoptosis of cultured cerebellar granule neurons, *Japan. J. Pharmacol.* **1999**, *79*, 369–378.
- [8] R. Laszlo, E.D. Franco, B. Thomas, F. Roland, G. Hermann, H. Pete, M. Ute, W. Romain, G.Z. Alfred, SAR of 2,6-diamino-3, 5-difluoropyridinyl substituted heterocycles as novel p38 MAP kinase inhibitors, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2109.
- [9] S. M. Vadlamudi, V. M. Kulkarni, 3D-QSAR of Protein Tyrosine Phosphatase 1B Inhibitors by Genetic Function Approximation, *Internet Electron. J. Mol. Des.* **2004**, *3*, 586–609.
- [10] O. Ivanciuc, 3D QSAR Models. In: QSPR/QSAR Studies by Molecular Descriptors. Ed.: M. V. Diudea. Nova Science, Huntington, N.Y., **2001**, pp. 233–280.
- [11] <http://www.coepra.org>.
- [12] SYBYL, version 6.8; Tripos Associates: St. Louis, MO, **2000**.
- [13] J. Gasteiger, M. Marsilli, Iterative partial equalization of orbital electro negativity—a rapid access to atomic charges, *Tetrahedron.* **1980**, *36*, 3219.
- [14] J. J. J. Stewart, MOPAC: A semi empirical molecular orbital program, *Comput. Aided Mol. Des.* **1990**, *4*, 1.
- [15] P. Geladi, Notes on the history and nature of partial least squares (PLS) modeling, *J. Chemom.* **1998**, *2*, 231.
- [16] R. Cramer III, D. Bunce, D.E., Patterson, I.E. Frank, Cross validation, bootstrapping, and partial least squares compared with multiple regression Conventional. QSAR. *Quant. Struct. Act. Relat.* **1988**, *7*, 18–25.
- [17] C.M. Venkatachalam, X. Jiang, T. Oldfield, M. Waldman, LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, *J. Mol. Graph. Model.* **2003**, *21*, 289–307.
- [18] E.Perola, Minimizing False Positives in Kinase Virtual Screens, *Proteins*, **2006**, *64*, 422–435.

Biographies

M Ravi Shashi Nayana: Research Scholar working for Ph.D degree from Osmania University. His area of interest is Computer Aided Drug Desig.

Y Nataraja Sekhar: Post Graduate in M.Tech. Majoring in Industrial Biotechnology. Anna University, Tamilnadu.