J. of the Chosun Natural Science
Vol. 4, No. 4 (2011) pp. 266 – 272

Modeling Aided Lead Design of FAK Inhibitors

Thirumurthy Madhavan†

Abstract

Focal adhesion kinase (FAK) is a potential target for the treatment of primary cancers as well as prevention of tumor metastasis. To understand the structural and chemical features of FAK inhibitors, we report comparative molecular field analysis (CoMFA) for the series of 7H-pyrrolo(2,3-d)pyrimidines. The CoMFA models showed good correlation between the actual and predicted values for training set molecules. Our results indicated the ligand-based alignment has produced better statistical results for CoMFA (q^2 = 0.505, r^2 = 0.950). Both models were validated using test set compounds, and gave good predictive values of 0.537. The statistical parameters from the generated 3D-QSAR models were indicated that the data are well fitted and have high predictive ability. The contour map from 3D-QSAR models explains nicely the structure–activity relationships of FAK inhibitors and our results would give proper guidelines to further enhance the activity of novel inhibitors.

Key words : 3D-QSAR, CoMFA, FAK

1. Introduction

Focal adhesion kinase (FAK) is a protein-tyrosine kinase that is found at the sites of cellular contact and is phosphorylated in response to cell attachment[1-3]. FAK plays an important role in cellular movement and survival pathways. It is a potential target for the treatment of both primary cancers and the prevention of tumor metastasis[4-7]. It is phosphorylated in response to integrin engagement, growth factor stimulation, and the action of mitogenic neuropeptides[8,9]. Integrin receptors are heterodimeric transmembrane glycoproteins that cluster upon ECM engagement leading to FAK phosphorylation and recruitment to focal adhesions[10,11]. FAK is involved in cellular adhesion and spreading processes and it has been shown that when FAK was blocked, breast cancer cells became less metastastic due to decreased mobility[12]. Thus, the inhibition of FAK represents a promising approach for treatment of cancer.

Computational approaches like structure- and ligand-based methodology have been found to be valuable in further optimization and the development of novel inhibitors. Ligand-based three dimensional quantitative structure-activity relationship (3D-QSAR) approaches, including comparative molecular field analysis (CoMFA)[13] and comparative molecular similarity indices analysis (CoMSIA)[14], were reported to be effective for understanding the structure-activity relationships[15]. The 3D-QSAR modeling is useful to predict the activity of new molecules to be synthesized. 3D-QSAR methods serve as an important complement to the structure-based methods. CoMFA is one the 3D-QSAR method that has been successfully employed in drug design. This method was useful in the lead optimization and also in understanding the drug-target interaction[16-18]. In CoMFA, the biological activity of molecules is correlated with their steric and electrostatic interaction energies. The steric and electrostatic interaction energies are calculated using Lennard- Jones and Coulombic potentials, respectively. The 3D-QSAR method would give contour maps as output that can be used to get some general insights into the topological features of the binding site.

The main aim of our study is to optimize the reported FAK inhibitors (7H-pyrrolo(2,3-d)pyrimidines)[19] by Ha-Soon and coworkers using the 3D-QSAR methodology. We report 3D-QSAR model using CoMFA technique for FAK inhibitors to find the common structural features among them. The present work deals ligand-based technique using atom-by atom matching to gen-

Centre for Bioinformatics, Department of Biochemistry, School of life sciences, University of Madras, Guindy campus, Chennai-600025, India

†Corresponding author : thiru.murthyunom@gmail.com
(Received : September 28, 2011, Revised : December 15, 2011, Accepted : December 22, 2011)

erate a reliable 3D-QSAR models. We expect that our theoretical results give some useful reference for the experimentalists in the design of novel and more potent FAK inhibitors.

2. Computational methods

2.1. Inhibitor data set

The structures of the 7H-pyrrolo(2,3-d)pyrimidines derivatives and their biological activities of sixty compounds were taken from the literature[19]. All original IC_{50} value of each inhibitor was converted into pIC_{50} ($-logIC_{50}$) in order to use the data as dependent variable in CoMFA model. The test set molecule is truly representative molecule for training set molecules. The test set molecule should cover all the biological activity which is similar to the training set molecule. The total set of compounds was divided into a training set consist of 46 compounds and test set consist of 14 compounds. The selection of training and test sets were done manually so that low, moderate, and high FAK inhibitory activities were all represented. The structures and their activity values are displayed in Table 1.

Table 1. Structures and biological activities (pIC_{50}) of FAK inhibitors

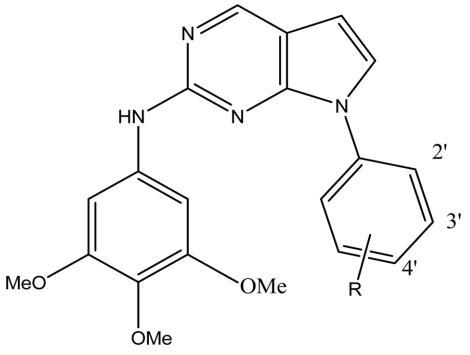
Compound	R	pIC_{50}
		
1	2'-N	6.674
16A*	2'-Me	5.551
16C*	2'-NO_2	5.393
16E	2'-CH_2OH	6.138
16G*	2'-CN	6.298
16H	2'-CO_2Me	6.503
16I	2'-CO_2H	5.550
17A*	3'-Me	6.656

Table 1. Continued

Compound	R	pIC_{50}
17B	3'-CF_3	6.629
17C	3'-CN	6.900
17D*	3'-CH_2CN	7.201
17E	3'-$(CH_2)_2CN$	7.032
17F	3'-CO_2Me	5.822
17G*	3'-CO_2H	7.420
17H	3'-CH_2CO_2H	7.432
17I	3'-$(CH_2)_2$ CO_2H	7.456
17J	3'-SO_2NHMe	5.567
18A	4'-Me	6.295
18B*	4'-Et	5.815
18C	4'-nBu	5.291
18D	4'-tBu	5.135
18E	4'-SMe	6.080
18F	4'-CF_3	5.821
18G	4'-NO_2	5.491
18H	4'-CO_2Me	5.187
18I	4'-CO_2H	5.547
22B*	2'-CONHMe	6.466
23B	3'-CONHMe	6.582
23C	3'-CONHEt	6.438
23D*	3'-CONHcPr	6.492
23E	3'-CONHiBu	5.296
23F	3'-CONHcPent	5.386
23G*	3'-$CONH(CH_2)_2NH_2$	7.538
23H	3'-$CONH(CH_2)_2NMe_2$	7.420
23J	3'-$CONH(CH_2)_2NHAc$	6.526
23K	3'-$CONH(CH_2)_3NMe_2$	7.268
23L*	3'-$CONH(CH_2)_4NMe_2$	6.007
23M	3'-$CONH(CH_2)_2OH$	6.363
23N	3'-$CONH(CH_2)_3OH$	6.055
24A	3'-$CH_2CONHMe$	5.742
24B	3'-$CH_2CONHEt$	6.376
24C	3'-$CH_2CONHcPr$	6.932
24D*	3'-$CH_2CONHcPent$	6.237
24E	3'-$CH_2CONH(CH_2)_2NH_2$	6.138
24F	3'-$CH_2CONH(CH_2)_3NH_2$	5.876
24H	3'-$CH_2CONHCH_2CO_2Me$	6.197
24I	3'-$CH_2CONH(CH_2)_2CO_2Me_2$	6.498
25A	3'-$(CH_2)_2CONHMe$	6.480
25B	3'-$(CH_2)_2CONHEt$	6.462
25C	3'-$(CH_2)_2CONHiBu$	7.201
25D	3'-$(CH_2)_2CONH(CH_2)_2NMe_2$	6.627

Table 1. Continued

Compound	R	pIC_{50}
25E*	3'-$(CH_2)_2CONHCH_2CO_2Me$	6.470
28A	2'-N, 3'-CONHMe	5.122
28C	2'-N, 3'-CONHiBu	6.740
28D	2'-N, 3'-$CONH(CH_2)_2NMe_2$	6.198
28E	2'-N, 3'-$CONH(CH_2)_3NH_2$	6.401
28F*	2'-N, 3'-$CONHCH_2CN$	6.556
28G	2'-N, 3'-$CONH(CH_2)_2CN$	6.616
28H	2'-N, 3'-$CONHCH_2CO_2Me$	6.264
32	2'-N, 3'-$(CH_2)_2CO_2H$	8.398

*Test set compounds

2.2. Ligand-based alignment method

In ligand based alignment, the most active molecule was used as template. All rotatable bonds were searched with incremental dihedral angle from 120° by using systematic search conformation method. Conformational energies were computed with electrostatic term, and the lowest energy conformer was selected as template molecule. Then the template was modified for other ligands of the series. The common moiety was constraint for each molecule and only the varying parts were energy minimized by Tripos force field with Gasteiger-Huckel charge by using conjugate gradient method, and convergence criterion was 0.05 kcal/mol at 10,000 iteration. The minimized structures were aligned over template using atom fit option in Sybyl and subsequently this alignment is used for CoMFA analysis. The aligned molecules are represented in Fig. 1.

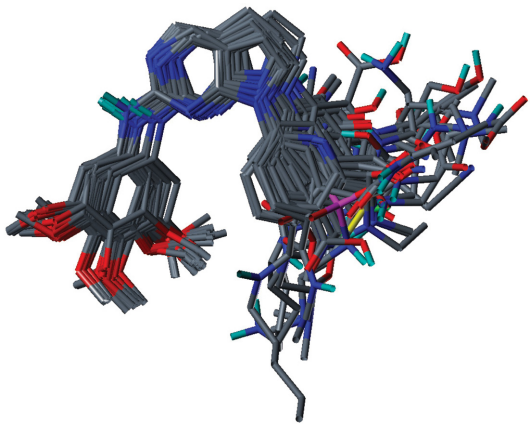


Fig. 1. Alignment of all molecules using atom-by-atom matching.

2.3. Generation of CoMFA Field

Sybyl8.1 molecular modeling package is used in this study[20]. CoMFA calculations were carried out by applying the default settings. The aligned molecules were placed in 3D cubic lattice with grid spacing of 2.0Å. The standard CoMFA field performing the Lennard-Jones potential and Coulombic potential for the steric and electrostatic fields respectively. A cut off value for both fields was set to 30 kcal/mol. Steric and electrostatic energies were calculated using sp^3 carbon atom with van der waals radius of 1.52Å and +1 charge at each lattice point.

2.4. Partial least square (PLS) analysis

The relationship between the structural parameters and the biological activities has been quantified by the PLS algorithm[21,22]. CoMFA and CoMSIA descriptors used as independent variables and pIC_{50} values used as dependent variables in PLS analysis for the generation of 3D-QSAR models. To select the best model, the cross-validation procedure was performed using leave one out (LOO) method, in this procedure one compound was removed from the data set and its activity was predicted using the model build from rest of the data set. It gives cross-validation correlation coefficient (q^2) and the optimum number of components. Final analyses i.e. non-cross-validation was performed to calculate conventional r^2_{ncv} using optimum number of components obtained from cross validation method. The cross-validated coefficient, q^2 is calculated using below Equation1.

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

$Y_{predicted}$, $Y_{observed}$, and Y_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively. $\Sigma(Y_{predicted}-Y_{observed})^2$ is the predictive sum of squares (PRESS) and the lowest PRESS value is used to derive the final PLS models.

3. Results and Discussion

3.1. CoMFA analysis

The optimum CoMFA model was derived with the combination of steric and electrostatic field contribution and Gasteiger-Hückel charge method with 2.0Å grid space. The Leave one out (LOO) analysis gave the

cross-validated q^2 of 0.505 with six components and noncross-validated PLS analysis resulted in a correlation coefficient r^2 of 0.950, F= 124.248, and an estimated standard error of 0.170. We further performed bootstrapping analyses to evaluate the robustness and statistical confidence of the final models (r^2_{boot}= 0.962, StdDev= 0.014). Statistical results obtained from the constructed model verified the predictive ability of the model (Table 2) and further implied that the steric and electrostatic factors contribute to the binding affinities. The predictive ability of the developed CoMFA model was assessed by the test set (fourten molecules) predictions, which were excluded during CoMFA model generation. The predictive ability of the test set was 0.537. Predicted and experimental activities and their residual values of all inhibitors are shown in Table 3, and the corresponding scatter plot is depicted in Fig. 2.

Table 2. Statistical result of CoMFA model

PLS statistics	CoMFA
q^2	0.505
N	6
r^2	0.950
SEE	0.170
F-value	124.248
r^2_{pred}	0.537
Bootstrap	
r2boot	0.962
StdDev	0.014
Field contribution (%)	
Steric	0.858
Electrostatic	0.142

q^2= cross-validated correlation coefficient; N= number of statistical components; r^2= non-cross validated correlation coefficient; SEE= standard estimated error; F= Fisher value; $r^2_{predictive}$= predictive correlation coefficient for test set.; r^2_{boot}= correlation coefficient after 100 runs of bootstrapping

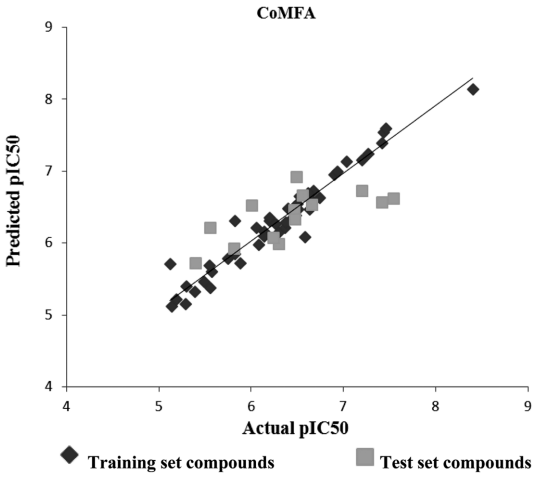


Fig. 2. Plot of actual versus predicted pIC_{50} values of the training set and test set compounds.

Table 3. Predicted activity of CoMFA models compared with the experimental pIC_{50} values

Compound	Actual pIC_{50}	CoMFA	
		Predicted	Residual
1	6.674	6.730	-0.0563
16A*	5.551	6.221	-0.6704
16C*	5.393	5.722	-0.3286
16E	6.138	6.164	-0.0262
16G*	6.298	5.995	0.3031
16H	6.503	6.470	0.0327
16I	5.550	5.385	0.1647
17A*	6.656	6.536	0.1205
17B	6.629	6.470	0.1594
17C	6.900	6.953	-0.0533
17D*	7.201	6.735	0.4656
17E	7.032	7.134	-0.102
17F	5.822	5.851	-0.0289
17G*	7.420	6.564	0.8564
17H	7.432	7.543	-0.1105
17I	7.456	7.596	-0.1396
17J	5.567	5.608	-0.041
18A	6.295	6.15	0.1447
18B*	5.815	5.926	-0.1112
18C	5.291	5.155	0.1359
18D	5.135	5.123	0.0124
18E	6.080	5.983	0.0971
18F	5.821	6.315	-0.4939
18G	5.491	5.467	0.0238
18H	5.187	5.224	-0.0375
18I	5.547	5.690	-0.1434
22B*	6.466	6.462	0.0038
23B	6.582	6.092	0.4899
23C	6.438	6.346	0.0923
23D*	6.492	6.921	-0.4295
23E	5.296	5.403	-0.1074
23F	5.386	5.331	0.0552
23G*	7.538	6.619	0.9187

Table 1. Continued

Compound	Actual pIC_{50}	CoMFA	
		Predicted	Residual
23H	7.420	7.396	0.0237
23J	6.526	6.653	-0.1267
23K	7.268	7.240	0.0278
23L*	6.007	6.530	-0.523
23M	6.363	6.215	0.1475
23N	6.055	6.218	-0.1634
24A	5.742	5.791	-0.0488
24B	6.376	6.290	0.0862
24C	6.932	7.001	-0.0693
24D*	6.237	6.072	0.1655
24E	6.138	6.094	0.0437
24F	5.876	5.726	0.15
24H	6.197	6.312	-0.1153
24I	6.498	6.545	-0.0473
25A	6.480	6.387	0.0929
25B	6.462	6.507	-0.0445
25C	7.201	7.159	0.0421
25D	6.627	6.620	0.007
25E*	6.470	6.335	0.1352
28A	5.122	5.715	-0.5934
28C	6.740	6.638	0.1017
28D	6.198	6.354	-0.156
28E	6.401	6.482	-0.081
28F*	6.556	6.665	-0.1085
28G	6.616	6.700	-0.0839
28H	6.264	6.273	-0.0086
32	8.398	8.145	0.2534

*Test set

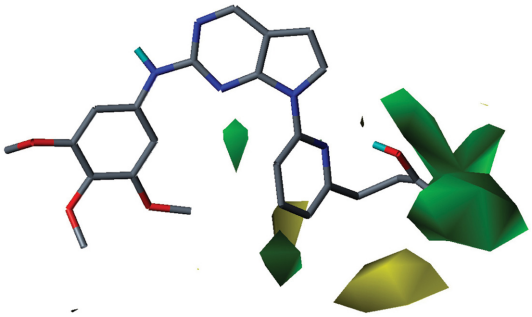


Fig. 3. CoMFA contour maps for steric field with highly active compound 32, where green contour indicates regions where bulky groups increases activity and yellow contours indicates bulky groups decreases activity.

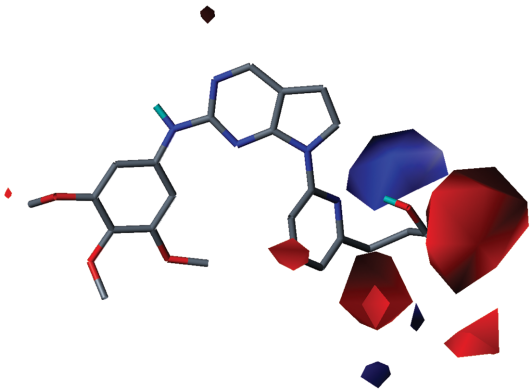


Fig. 4. CoMFA contour maps for electrostatic field with highly active compound 32, where blue contour indicates regions where electropositive groups increases activity and red contours indicates regions where electronegative groups increases activity.

3.2. CoMFA contour map

The CoMFA contour map was generated based on the ligand-based (atom-by atom matching) alignment method. The CoMFA result is usually represented as 3D 'coefficient contour' map. It shows regions where variations of steric and electrostatic nature in the structural features of the different molecules contained in the training set lead to increase or decrease in the activity. The steric interaction is represented by green and yellow contours, in which green colored regions indicate areas where increased steric bulk is associated with enhanced activity, and yellow regions suggest areas where increased steric bulk is unfavorable to activity. The steric contour map is displayed in Fig. 3. There is a big green steric contour near the 3 R position of the phenyl ring indicates that bulkier substituent is preferred at this position. Thus, compounds 17 (D, E, F, H, I), and 23 (G-N) with bulkier substituent at this position are more active. This may be the reason that compounds having bulkier substitution shows higher activity. There was one yellow contour region was observed near to the 4 R_1 position, the contour map indicated that substitution of bulkier groups would decrease the activity. This may be the reason that compounds (18C and D) having bulkier substitution shows lower activity. The electrostatic interaction is represented by red and blue contours (Fig. 4), among which blue colored regions show areas where more positively charged groups are favored, and red region high-

light areas where groups with more negative charges are favored. These contour maps give us some general insight into the nature of the receptor-ligand binding region. The electrostatic contour plot on the set of 60 compounds shows that there is a big red colored region situated close to the 4' R positions of phenyl ring. The negative charges in these regions are very important for ligand binding, and electro negative group linked to this position will enhance the biological activity. For example, compounds 32 having electro negative moiety shows more biological activity than other compounds.

According to the contour maps, steric, and electrostatic requirements are playing major role in order to enhance the biological activities of the compounds. The contour map analyses found that bulky substitution with more electronegative substitutions at the 3 R position of the phenyl ring is highly desirable to enhance the activity. Modification in these positions seems remarkably important to enhance the activity for FAK inhibitors.

4. Conclusion

We developed satisfactory 3D-QSAR model based on CoMFA technique using atom-by-atom matching alignment. The 3D-QSAR results revealed some important sites, such as steric, electrostatic modifications should significantly affect the bioactivities of the compounds. In this study, we found some implications could be drawn to improve the activity of 7H-pyrrolo(2,3-d)pyrimidines as FAK inhibitors, for example, modification of bulky substitution with more electronegative substitutions at the 3 R position of the phenyl ring is highly desirable to produce higher activity. The information's have gathered from CoMFA studies demonstrated the way to understand the structural and chemical features of 7H-pyrrolo(2,3-d)pyrimidines derivatives in designing potent and optimize new FAK inhibitors.

References

[1] A. Richardson, and J. T. Parsons, "A mechanism for regulation of the adhesion-associated protein tyrosine kinase pp125FAK", Nature., Vol. 380, pp. 538-540, 1996.

[2] A. P Gilmore, and L. H. Romer, "Inhibition of focal adhesion kinase (FAK) signaling in focal adhesions decreases cell motility and proliferation", Mol. Biol. Cell., Vol 7, pp. 1209-1224. 1996.

[3] R. O. Hynes, "Cell adhesion: Old and new questions", Trends Cell biology., Vol. 9, pp. M33-M37, 1999.

[4] D. Ilic, Y. Furuta, S. Kanazawa, N. Takeda, K. Sobue, N. Nakatsuji, S. Nomura, J. Fukimoto, M. Okada, T. Yamamoto, and S. Aizawa, "Reduced cell motility and enhanced focal adhesion contact formation in cells from FAK-deficient mice", Nature., Vol. 377, pp. 539-544, 1995.

[5] S. M. Frisch, K. Vuori, E. Ruoslahti, and P. Y. J. Chan-Hui, "Cell Biol., Vol. 134, pp. 793-799, 1996.

[6] T. Miyazaki, H. Kato, M. Nakajima, M. Sohda, Y. Fukai, N. Masuda, R. Manda, M. Fukuchi, K. Tsukada, and H. Br. J. Kuwano, "FAK overexpression is correlated with tumour invasiveness and lymph node metastasis in oesophageal squamous cell carcinoma", J. Cancer., Vol. 89, pp. 140-145. 2003.

[7] S. K. Mitra, D. A. Hanson, and D. D. Schlaepfer, "Focal adhesion kinase: in command and control of cell motility" Nat. Rev. Mol. Cell Biol., Vol. 6, pp. 56-68, 2005.

[8] H, Abedi, and I. Zachary, "Vascular endothelial growth factor stimulates tyrosine phosphorylation and recruitment to new focal adhesions of focal adhesion kinase and paxillin in endothelial cells", The Journal of biological chemistry., Vol. 272, pp. 15442-15451, 1997.

[9] I. Zachary, and E. Rozengurt, "Focal adhesion kinase (p125FAK): a point of convergence in the action of neuropeptides, integrins, and oncogenes", Cell., Vol. 71, pp. 891-894, 1992.

[10] K. Burridge, K. Fath, T. Kelly, G. Nuckolls, and C. Turner, "Focal adhesions: transmembrane junctions between the extracellular matrix and the cytoskeleton", Annual review of cell biology., Vol. 4, pp. 487-525, 1988.

[11] K. Burridge, C. E. Turner, and L. H. Romer, "Tyrosine phosphorylation of paxillin and pp125FAK accompanies cell adhesion to extracellular matrix: a role in cytoskeletal assembly", The Journal of cell biology. Vol. 119, pp. 893-903, 1992.

[12] S. E. Blackshaw, J. D. Kamal, J. M. Lackie, "The dictionary of cell and molecular biology", San Diego: Academic Press., 1999.

[13] R. D. Cramer, D. E. Patterson, and J. D. Bunce, "Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins", J. Am. Chem. Soc., Vol. 110, pp. 5959-5967. 1988.

[14] G. Klebe, U. Abraham, and T. Mietzner, "Molecular

similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity", J. Med. Chem., Vol. 37, pp. 4130-4146, 1994.

[15] H. Kubinyi, "Comparative molecular field analysis (CoMFA)", In The Encyclopedia of Computational Chemistry, John Wiley & Sons., Vol. 1, pp. 448-460, 1998.

[16] M. D. M. Abdul Hameed, A. Hamza, J. Liu, X. Huang, and C. G. Zhan, "Human microsomal prostaglandin E synthase-1 (mPGES-1) binding with inhibitors and the quantitative structure-activity correlation", J. Chem. Inf. Modeling., Vol. 48, pp. 179-185, 2008.

[17] C. L. Kuo, H. Assefa, S. Kamath, Z. Brzozowski, J. Slawinski, F. Saczewski, J. Buolamwini, and K. N. Neamati, "Application of CoMFA and CoMSIA 3D-QSAR and docking studies in optimization of mercaptobenzenesulphonamides as HIV-1 integrase inhibitors", J. Med. Chem., Vol. 47, pp. 385–399. 2004.

[18] G. F. Yang, H. T. Lu, Y. Xiong, C G. Zhan, "Understanding the structure-activity and structure-selectivity correlation of cyclic guanine derivatives as phosphodiesterase-5 inhibitors by molecular docking, CoMFA and CoMSIA analyses", Bioorg. Med. Chem., Vol. 14, pp. 1462-1473. 2006.

[19] H. S. Choi, Z. Wang, W. Richmond, X. He , K. Yang, T. Jiang, D. Karanewsky, X. J. Gu , V. Zhou, Y. Liu, J. Che, C. C. Lee, J. Caldwell, T. Kanazawa, I. Umemura, N. Matsuura, O. Ohmori, T Honda, N Gray, and Y. He, "Design and synthesis of 7H-pyrrolo[2,3-d]pyrimidines as focal adhesion kinase inhibitors. Part 2", Boorg Med Chem Lett., Vol. 16, pp. 2689-2692, 2006.

[20] Sybyl 8.1, Tripos Inc., St. Louis, MO 63144, USA.

[21] S. J. Cho, and A. Tropsha, "Cross-validated R2-guided region selection for comparative molecular field analysis: A simple method to achieve consistent results", J. Med. Chem., Vol. 38, pp.1060-1066, 1995.

[22] S. Wold, M. Sjostrom, and L. Eriksson, "PLS-regression: a basic tool of chemometrics", Chemometrics and intell. lab. Sy., Vol. 58, pp. 109-130, 2001.