SHORT NOTE

THREE-DIMENSIONAL QSAR MODELING BENZIMIDAZOLE ANALOGUES USING THE K-NEAREST NEIGHBOR METHOD

ABSTRACT

We undertook the three-dimensional (3D) QSAR studies of a series of benzimidazole analogues to elucidate the structural properties required for angiotensin II. The 3D-QSAR studies were performed using the stepwise, simulated annealing (SA) and genetic algorithm (GA) selection k-nearest neighbor molecular field analysis approach; a leave-one-out cross-validated correlation coefficient q^2 = 0.8216 and a pred_r^2 = 0.7852 were obtained. The 3D QSAR model is expected to provide a good alternative to predict the biological activity prior to synthesis as antihypertensive agents.

Keywords: angiotensin II, benzimidazole, k-nearest neighbor, Antihypertensive activity

Introduction

The Renin Angiotensin Aldosterone System (RAAS) is a proteolytic cascade that plays an important role in electrolyte homeostasis and in the regulation of blood pressure[1]. The physiological responses of Ang II are mediated through at least two receptor subtypes designated as AT_1 and AT_2[2]. Angiotensin II an octapeptide produced from angiotensin I by the action of angiotensin converting enzyme (ACE) localized on the endothelium of blood vessels in the lungs, kidneys, and many other organs, is the primary effector component of the RAS[3]. Angiotensin II receptor blockers (ARBs) have been developed to produce a more complete blockade of the action of Ang II compared to other drug classes as well as an improved side effect profile[4,5]. Computational chemistry, prediction of biological activity based quantitative structure activity relationship (QSAR) substantially increases the potential of work, avoiding time and resource consuming experiments[6]. 3D QSAR protocols have been selected with a view to understand the ligand-receptor interaction in the light of steric, electrostatic and hydrophobic properties. For the development of 3D-QSARs, molecular field analysis[7] has been applied to evaluate specific contributions of steric and electrostatic field effects necessary for the activity. The k-nearest neighbor (kNN) analysis with stepwise (SW), genetic algorithm (GA) and simulated annealing (SA) has been applied for the development of 3D-QSAR model. The molecular structure of the training set and test were sketched using V-Life MDS (Molecular Design Suite)™ 3.5, India 2006[8]. In continuation to our team's earlier efforts in developing a few QSAR models to predict the biological activities of different groups of compounds[9-30].

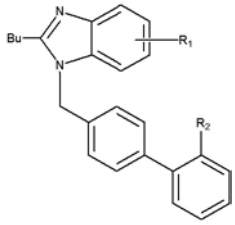
MATERIALS AND METHOD

The Angiotensin II AT_1 receptor activity of thirty three benzimidazole derivatives[31] was used in the present study (Table I). The inhibitor activities adrenal cortical membranes (IC_{50}) were converted negative values (pIC_{50}) was used in 3D QSAR. Each compound was energy minimized and batch optimized by using Merck Molecular Force Field[32] force field and charges followed Hamiltonian method available in MOPAC module with the convergence criterion 0.001 kcal/mol Å. fixing Root Mean Square Gradients (RMS) to 0.01 Kcal/mol Å. Energy minimized and geometry optimized structure of molecules were aligned by the template-based method[7] using VLife MDS 3.5 software. The template structure, i.e.benzimidazole ring was used for alignment by considering the common elements of the series, as shown in Fig. 1(a). The superimposition of all molecules based on minimizing root mean square deviation (RMSD) is shown in Fig. 1(b). The optimal test and training data set were generated using sphere exclusion method[33]. The dissimilarity level was set to 6.2, as the higher the dissimilarity level, the lesser the predictive ability of QSAR model.

Methodology

In 3D QSAR kNN-MFA common rectangular grid around the molecules was built. The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1. For calculation of field descriptor values, using Tripos force field[34] both electrostatic and steric field types, with cut-offs of 10.0 and 30.0 kcal/ mol, respectively, were selected and charge type was selected[35]. The dielectric constant was set to 1.0 considering the distance dependent dielectric function. This resulted in calculation of 3630 field descriptors (1210

Table I: The structures of 2-butylbenzimidazole derivatives with activities

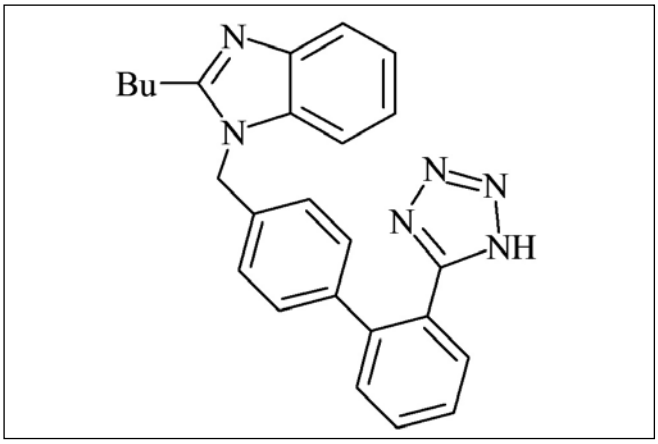


Comp	R^1	R^2	IC_{50} [a]	pIC_{50} [b]
1	H	Tet	9.0	6.045
2	5-OMe	Tet	9.1	6.043
3	6-OMe	Tet	11	5.958
4	5-Cl	Tet	15	5.823
5[c]	6-Cl	Tet	31	5.508
6	7-OMe	Tet	28	5.552
7	4-CO_2Me	Tet	72	5.142
8	5-CO_2Me	Tet	7.4	6.134
9	6-CO_2Me	Tet	4.4	6.356
10	7-CO_2Me	Tet	3.2	6.498
11[c]	5-Me-7-CO_2Me	Tet	8.7	6.064
12	5-Cl-7-CO_2Me	Tet	4.4	6.356
13	6-Me-7-CO_2Et	Tet	9.1	6.046
14	4-$CONH_2$	Tet	130	4.886
15[c]	7-CO_2Et	Tet	14	5.856
16	7-CO_2Bu	Tet	12	5.886
17	5-CO_2H	Tet	55	5.256
18[c]	6-CO_2H	Tet	90	5.045
19	7-CO_2H	Tet	5.5	6.259
20[c]	5-Me-7-CO_2H	Tet	13	5.886
21	5-Cl-7-CO_2H	Tet	11	5.958
22	6-Me-7-CO_2H	Tet	3.4	6.484
23[c]	H	CO_2H	11	6.958
24	7-CO_2H	CO_2H	6.6	6.187
25	7-CO_2H	1-Me-Tet	34	5.468
26[c]	7-CONHi-Pr	Tet	5.4	6.264
27	7-CH_2OH	Tet	4.5	6.346
28	7-CH_2OMe	Tet	6	6.221
29[c]	7-CH_2NMe_2	Tet	24	5.619
30	7-Me	Tet	3.3	6.481
31	7-CH_2CO_2Et	Tet	2.5	6.602
32	7-OH	Tet	11	5.958
33	7-CH_2CO_2H	Tet	26	5.558

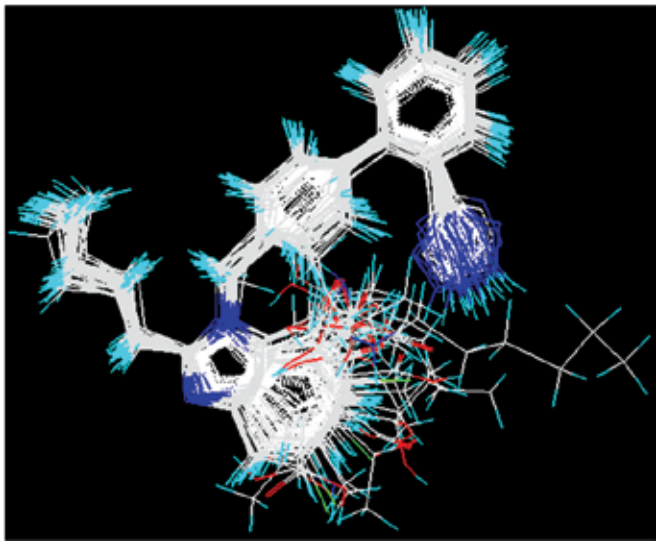
[a] Inhibition of ^{125}I angiotensin II binding to bovine adrenal cortex (IC_{50}), [b] -log IC_{50} to generate equation, [c] compounds belonging to test set.

Table II: Statistical results of 3D-QSAR models generated

Sr. No.	Statistical parameter	3D QSAR Result			
		Model 1	Model 2	Model 3	Model 4
1	q^2	0.8024	0.8216	0.7265	0.6471
2	pred_r^2	0.7264	0.7852	0.6984	0.5947
3	r^2_se	0.1652	0.4365	0.2765	0.1658
4	q^2_se	0.4218	0.6532	0.4326	0.3276
5	pred_r^2se	0.5427	0.2593	0.6532	0.2678
6	F test	73.652	82.653	25.654	16.476
7	$N_{training}$	25	25	25	25
8	N_{test}	8	8	8	8



(a) Common substructure



(b) Alignment of molecules

Fig 1: Benzimidazole with biphenyl tetrazole ring (template structure)

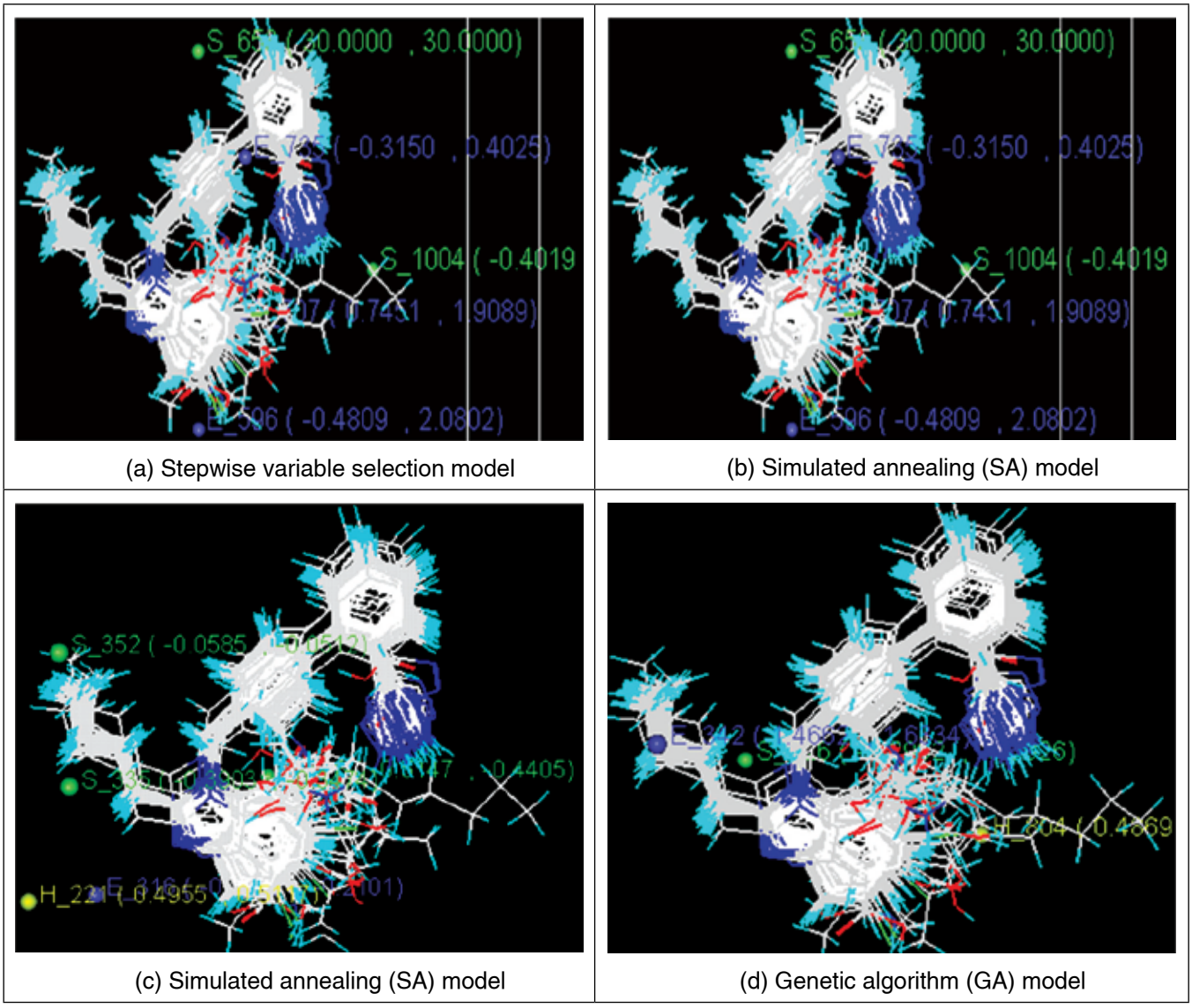


Fig:2 Contribution plot for steric, electrostatic and hydrophobic interactions.

for each steric, electrostatic and hydrophobic) for all the compounds in separate columns.

We hereby report the models, as generated by kNN-MFA in conjunction with stepwise (SW) forward-backward variable selection methods[36], simulated annealing (SA)[37] and genetic algorithm [38] have been applied for descriptor optimization. Partial least square (PLS) analysis has been applied for three-dimensional (3D) QSAR models development. In order to validate the generated QSAR models, leave one out (LOO) method[39] was used indicated as value of q^2 (cross validated explained variance) which is a measure of internal predictive ability of the model.

RESULTS AND DISCUSSION

The 3D QSAR models were evaluated using the following statistical measures: n, (the number of compounds in regression); k, (number of variables); DF, (degree of freedom); optimum component, (number of optimum PLS components in the model); r^2 (the squared correlation coefficient), F test (Fischer's value) for statistical significance, q^2 (cross-validated correlation coefficient); pred_r^2, (r^2 for external test set). To ensure a fair comparison, the same training and test sets are used for each of the models. The 3D-QSAR model-1 Stepwise (SW) variables (Fig 2a) were selected on the basis of statistical parameters and the values of the model 1 having q^2 = 0.8024 and pred_r^2 = 0.7264 (Table II). The respective relative contributions of steric

Table III: Observed and predicted activities of statistically significant models obtained by PLS

3D QSAR model-1		3D QSAR model-2		3D QSAR model-4	
Predicted activity	Res.*	Predicted activity	Res.*	Predicted activity	Res.*
6.012	0.033	6.114	-0.069	6.072	-0.027
6.136	-0.093	6.087	-0.044	6.098	-0.055
5.351	0.607	5.218	0.74	5.414	0.544
5.585	0.238	6.321	-0.498	5.816	0.007
5.021	0.487	4.887	0.621	4.768	0.74
6.131	-0.579	5.032	0.52	5.124	0.428
4.565	0.577	5.267	-0.125	5.765	-0.623
5.613	0.521	5.843	0.291	5.736	0.398
5.698	0.658	6.918	-0.562	6.822	-0.466
6.031	0.467	6.196	0.302	7.018	-0.52
6.012	0.052	7.057	-0.993	6.017	0.047
5.695	0.661	5.898	0.458	5.631	0.725
6.165	-0.119	5.583	0.463	5.447	0.599
4.124	0.762	4.076	0.81	4.129	0.757
6.375	-0.519	5.407	0.449	5.518	0.338
5.487	0.399	5.128	0.758	5.249	0.637
4.889	0.367	4.965	0.291	4.813	0.443
5.179	-0.134	4.579	0.466	4.497	0.548
5.654	0.605	5.736	0.523	5.759	0.5
5.321	0.565	6.032	-0.146	5.265	0.621
6.778	-0.82	5.521	0.437	5.915	0.043
6.167	0.317	6.119	0.365	6.756	-0.272
7.393	-0.435	7.183	-0.225	7.264	-0.306
5.827	0.36	5.804	0.383	5.696	0.491
5.769	-0.301	5.298	0.17	5.403	0.065
6.714	-0.45	6.422	-0.158	5.932	0.332
5.931	0.415	6.615	-0.269	6.076	0.27
6.721	-0.5	5.806	0.415	5.895	0.326
5.476	0.143	5.284	0.335	5.436	0.183
6.679	-0.198	6.283	0.198	6.738	-0.257
6.374	0.228	6.412	0.19	6.187	0.415
5.681	0.277	5.621	0.337	5.232	0.726
5.287	0.271	5.308	0.25	5.218	0.34

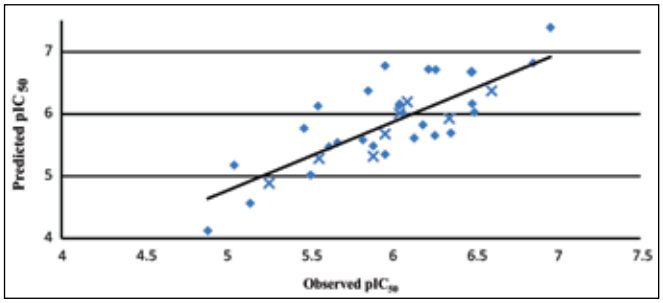


Fig. 3 (a) Contribution plot between selected descriptors for Model 1

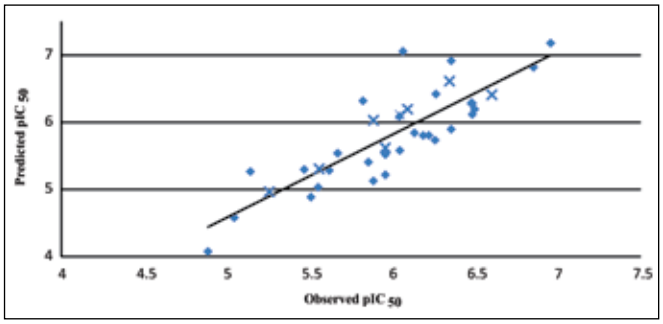


Fig. 3 (b) Contribution plot between selected descriptors for Model 2

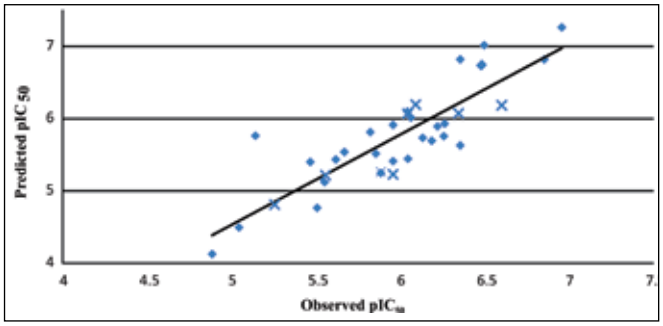


Fig. 3 (c): Contribution plot between selected descriptors for Model 4

field (green region), electrostatic field (blue region) and hydrophobic (yellow region) indicate that steric field is more predominant than electrostatic field. Descriptor range for the selected model negative range of descriptors S_1004 steric field showed a region where bulky substituents are disfavored for activity. Electrostatic field descriptor such as E_707 with positive coefficient represents regions where electropositive (electron-withdrawing) groups are favorable for activity. Negative range value of descriptors reveals that E_596, E_735 electrostatic indicates that less electronegative substituent would be favourable for the activity. The model is validated by predicting the biological activities of the test molecules, as indicated in Table III. The plot of observed versus predicted activities for the test compounds is represented in Fig.3 (a). The above steps were repeated Model-2 Simulated annealing

(Fig 2b) with four descriptors namely H_1027, E_1046, S_551 and E_274. As far as S_551 steric field is concerned, a negative range indicated that less bulky substituent group was preferred in that region. E_1046 and E_274 electrostatic field descriptors with positive coefficients represent regions where electropositive groups are favourable. Positive coefficients represent regions of H_1027 hydrophobic potential is favourable for increase in activity. The activity distribution plot for observed versus predicted activities for the test compounds are shown in Fig. 3(b). For the series of benzimidazoles derivatives, the template-based model from 3D-QSAR model-3 simulated annealing (Fig 2c), it is observed that electrostatic field with negative coefficient E_316, near moiety at 2-position of the electronegative group is preferred. As far as steric field S_352, S_618 is concerned, a negative range indicated that a negative steric potential was favourable for increased activity, and hence a less bulky substituent group was preferred in that region. The steric, electrostatic and hydrophobic contributions were 65, 20 and 15 %, respectively and exhibited good external prediction with r^2 *pred* of 0.6984. The LOO cross-validated value (q^2) of PLS analysis was found to be 0.6471, suggesting that the model could be useful for predicting antihypertensive activity for such benzimidazoles. In Model 4, genetic algorithm ((Fig 2d)) as shown in positive range of electrostatic field descriptor E_342. The steric contribution S_426 exhibiting negative range in green at the ring suggests if at bulky groups are not favorable for activity. The activity distribution plot for observed versus predicted activities for the test compounds are shown in Fig. 3(c).

CONCLUSION

In the present study, an attempt has been made to identify the necessary structural requirements of benzimidazole derivatives to be potential anti-hypertensive agents. The 3D-QSAR studies were carried out using SW, SA and GA kNNMFA based partial least-squares method provides the best results for training and test sets. 3D-QSAR suggested that substitution of less bulky, electron withdrawing and electron donating and hydrophobic groups around R1, and R2 positions. Thus the above QSAR study gives important information about the structural requirements of the moiety and further helps in generating new analogs to find app an antihypertensive activity.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Goodman G.: The Pharmacological Basis of Therapeutics, 7th Ed., McGraw-Hill, New York, 1985; 639-659.
2. Timmermans P.B, Wong P.C, Chiu A.T, Herblin W.F, Benfield P, Carini D.J, Lee R.J, Wexler R.R, Saye J.A. and Smith R.D.: Angiotensin II receptors and angiotensin II receptor antagonists. **Pharmacol. Rev.**, 1993,45(2)205–251.
3. Kubo K, Kohar Y, Yoshimura Y, Shibouta Y and Furukawa Y.: Nonpeptide angiotensin II receptor antagonists. Synthesis and biological activity of potential prodrugs of benzimidazole-7-carboxylic acids. **J. Med Chem.**, 1993,36,2343–2349.
4. Tsutsumi K, and Saavedra J.M.: Characterization and development of angiotensin II receptor subtypes (AT_1 and AT_2) in rat brain. **Am. J. Physiol.**, 1991,261(1 Pt 2) R209-16.
5. Martin J and Krum H.: Role of valsartan and other angiotensin receptor blocking agents in the management of cardiovascular disease. **Pharmacol. Res.**, 2002, 46(3) 203-12.
6. González-Díaz H and Prado-Prado F.J.: Unified QSAR and network-based computational chemistry approach to antimicrobials, part 1: multispecies activity models for antifungals. **J. Comput. Chem**, 2008,29(4)656-67.
7. Ajmani S, Jadhav K, Kulkarni S.A.:Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. **J. Chem. Inf. Model.**, 2006, 46:24–31
8. VLife Sciences Technology Pvt. Ltd. Pune-411045.
9. Sharma M.C, and Sharma S.: Quantitative structural–activity relationship (QSAR) study for antimycobacterial activities of pyrazine containing thiazoline and thiazolidinone derivatives, **Opt. Adv. Mat. Rap. Comm.**, 2010, 4 (3)415-421.
10. Sharma M.C, Sharma S, Kohli D.V.: Molecular Modeling Studies of Substituted 2, 4, 5-trisubstituted triazolinones aryl and nonaryl derivatives as angiotensin II AT_1 receptor antagonists. **J. Chem.**, 2013, doi org/10.1155/2013/427181
11. Sharma M.C, Sharma S.: Molecular Modeling Studies of thiophenyl C-aryl glucoside SGLT2 inhibitors as potential antidiabetic agents. **Int. J. Med. Chem.,** 2014, doi.org /10.1155 /2014/739646
12. Sharma M.C.: Structural features of substituted triazole linked chalcone derivatives as antimalarial activities against D10 strains of Plasmodium falciparum: A QSAR approach. **J. Cent. South. Univ.**, 2015, 22, 3738–3744
13. Bhadoriya K.S, Sharma M. C, Jain S.: Pharmacophore modeling and atom-based 3D-QSAR studies on amino derivatives of indole as potent isoprenylcysteine carboxyl methyltransferase (Icmt) inhibitors. **J. Mol. Struct.**, 2015,1081, 466-476
14. Sharma MC, Kohli DV, Sahu NK, Sharma S, Chaturvedi SC: 2D- QSAR studies of some 1, 3, 4-thidiazole-2yl

azetidine 2-one as antimicrobial activity. **Dig J. Nanomater Biostruct.** 2009,4, 339 – 347.

15. Sharma MC, Kohli DV, Sharma S, Chaturvedi SC: Two dimensional- quantitative Structure activity relationships -2,3diarylthiophenes as selective COX-1-2 inhibitors. **Dig J. Nanomater. Biostructures.** 2009,4 (3) 459-469.
16. Sharma MC, Kohli DV, Sahu,NK, Sharma S, Chaturvedi SC: QSAR, Synthesis and Biological activity studies of some thiazolidinones derivatives. **Dig. J. Nanomater. Biostruct.** 2009,4(1) 223- 232.
17. Sharma M.C, Sharma S, Sharma P, Kumar A, Bhadoriya K.S.: Structural insights for substituted acyl sulfonamides and acyl sulfamides derivatives of imidazole as angiotensin II receptor antagonists using molecular modelling approach. **J. Taiw. Inst. Chem. Eng.**, 2014,45 , 12–23.
18. Sharma M.C and Sharma S.:Molecular Modeling Studies of 3-acyl-2-phenylamino-1, 4-dihydroquinolin-4-one derivatives as phosphatase SerB653 inhibitors. **Med. Chem. Res.**, 2016, 25(10) 2119–2126.
19. Sahu N.K, ShahiS,Sharma M.C, Kohli D. V.: QSAR studies on imidazopyridazine derivatives as PfPK7 inhibitors. **Mol. Simul.**, 2011, 37(9) 752-765.
20. Sharma, M.C., Kohli, D.V.:Prediction of biological activity of substituted 2,5-Dibutyl-2, 4-dihydro derivatives by 3D-QSAR results. **J. Balk. Tribol. Assoc.**, 2018,24(1)95-105.
21. Sharma, M.C.:Structural insights into mode of actions of dipeptidyl peptidase IV inhibitors as anti-diabetic agents: Computational analyses. **J. Balk. Tribol. Assoc.**, 2017, 23(3) 515-526.
22. Sharma, M.C., Kohli, D.V.: Development of a credible QSAR studies for a series of sulphonamides derivatives. **J. Balk. Tribol. Assoc.,** 2017, 23(4):664-673.
23. Sharma M.C, Jain S, Sharma R.: In silico screening for identification of pyrolidine derivatives dipeptidyl peptidase-IV inhibitors using COMFA, CoMSIA, HQSAR and Docking studies. **In Silico Pharmac.** 2017, 5: 13.DOI: 10.1007/s40203-017-0032-2
24. Sharma, M.C.: A Structure-activity Relationship Study of Naphthoquinone Derivatives as Antitubercular Agents Using Molecular Modelling Techniques. **Interdiscip Sci Comput Life Sci**, 2015,7(4) 346-356
25. Sharma, M.C: Discovery of potent antihypertensive ligands Substituted imidazolyl biphenyl sulfonylureas analogs as angiotensin II AT_1 receptor antagonists by molecular modelling studies. **Inter. Sci. Comput. Life. Sci.** 2015,7(3), pp 221-232
26. Sharma, M.C.: Comparative Pharmacophore modeling and QSAR studies for structural requirements of some Substituted 2-Aminopyridines derivatives as Inhibitors Nitric Oxide Synthases. **Interdiscip. Sci. Comput. Life Sci**., 2015,7(2) 100-112
27. Sharma, M.C.: Structural insights into mode of actions of novel substituted 4- and 6-azaindole-3-carboxamides analogs as renin inhibitors: molecular modeling studies **Med. Chem. Res.**, 2015, 24:1038–1059
28 Bhadoriya K.S, Sharma M. C, Sharma S, Jain S. V and Avchar M. H**.:** An approach to design potent anti-Alzheimer's agents by 3D-QSAR studies on fused 5, 6-bicyclic heterocycles as γ-secretase modulators using kNN-MFA methodology. **Arab. J. Chem.,** 2014,7(6) 924-935
29. Sharma M.C, Sharma S, Kohli D.V.: QSAR Approach Insight the Structural Requirement of Substituted Quinazolinones Derivatives as Angiotensin II Receptor Antagonists. **Oxid. Comm.,** 2012, 35(3), 694-707, 2012
30. Sharma M.C and Kohli D.V.: Predicting 2, 3-dihydroquinazolinones Derivatives as Angiotensin II Receptor Antagonists: 2D QSAR approach.**Oxid.Comm.,** 2012, 35(3),721-734
31. Kubo K, Inada Y, Kohara Y, Sugiura Y, Ojima M, Itoh K, Furukawa Y, Nishikawa K, Naka T.: Nonpeptide angiotensin II receptor antagonists. Synthesis and biological activity of benzimidazoles. **J. Med. Chem.**, 1993;36(12):1772-84.
32. Halgren TA.: Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. **J. Comput. Chem.**,1996, 17:553–586
33. Golbraikh A, Tropsha A.: Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. **J. Comput. Aided. Mol. Des.** 2002,16:357–369
34. Clark M, Cramer RD III, Van ON.: Validation of the general purpose tripos 5.2 force field. **J. Comput. Chem.**,1989, 10:982–1012
35. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH.: Equation of state calculations by fast computing machines. **J .Chem. Phys.** 1953,21:1087–1092
36. Darlington RB.:Regression and linear models. McGraw-Hill,New York,1990
37. Zheng W and Tropsha A.: Novel variable selection quantitative structure–property relationship approach based on the k-nearest neighbor principle. **J. Chem. Inf. Comput. Sci.,** 2000, 40:185–194
38. Holland JH.: Genetic algorithms. **Sci. Am.** 1992, 267:66–72
39. Cramer RD, Patterson DE, Bunce JD.: Comparative molecular field analysis (CoMFA) 1. Effect of shape on binding of steroids to carrier proteins. **J. Am. Chem. Soc.,** 1988, 110,5959–5967

[a] Department of Pharmaceutical Sciences,
Dr. H.S. Gour University, Sagar - 470 003, Madhya Pradesh, India

Sharma M. C.[a]* and Kohli D. V.[a]

*** For Correspondence E-mail: mukesh2206@rediffmail.com**

(Received 13 November 2017) (Accepted 24 October 2018)

https://doi.org/10.53879/id.56.12.11234