



King Saud University
Journal of Saudi Chemical Society

www.ksu.edu.sa
www.sciencedirect.com



ORIGINAL ARTICLE

QSAR analysis of furanone derivatives as potential COX-2 inhibitors: kNN MFA approach



Ruchi Bhatiya, Ankur Vaidya, Sushil K. Kashaw, Abhishek K. Jain,
Ram K. Agrawal *

Pharmaceutical Chemistry Research Laboratory, Department of Pharmaceutical Sciences, Dr. Hari Singh Gour University,
Sagar 470 003, Madhya Pradesh, India

Received 7 November 2011; accepted 1 December 2011
Available online 17 December 2011

KEYWORDS

QSAR;
V-Life;
Furanone derivatives;
COX-2 inhibitors

Abstract A series of thirty-two furanone derivatives with their cyclooxygenase-2 inhibitory activity were subjected to quantitative structural–activity relationship analysis to derive a correlation between biological activity as a dependent variable and various descriptors as independent variables by using V-LIFE MDS3.5 software. The significant 2D QSAR model showed correlation coefficient (r^2) = 0.840, standard error of estimation (SEE) = 0.195, and a cross-validated squared correlation coefficient (q^2) = 0.773. The descriptors involved in the building of 2D QSAR model are retention index for six membered rings, total number of oxygen connected with two single bonds, polar surface area excluding P and S plays a significant role in COX-2 inhibition. 3D-QSAR performed via Step Wise K Nearest Neighbor Molecular Field Analysis [(SW) kNN MFA] with partial least-square (PLS) technique showed high predictive ability (r^2 = 0.7622, q^2 = 0.7031 and standard error = 0.3660) explaining the majority of the variance in the data with two principle components. The results of the present study may be useful in the design of more potent furanone derivatives as COX-2 inhibitors.

© 2011 Production and hosting by Elsevier B.V. on behalf of King Saud University.
Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction

Non-steroidal anti-inflammatory drugs (NSAIDs) are the most frequently prescribed drugs used in the treatment of

* Corresponding author. Tel.: +91 7582 233934; fax: +91 7582 264236.

E-mail address: ramkishoreagrawal@gmail.com (R.K. Agrawal).

Peer review under responsibility of King Saud University.



inflammation, rheumatic and arthritic diseases. The mechanism involves the non-selective inhibition of cyclooxygenase (COX), thereby preventing the biosynthesis of prostaglandins which are important mediators of inflammation as well as homeostatic physiological function (Vane et al., 1998; Botting, 1999). COXs exist in two isoforms, COX-1 and COX-2 (Vane et al., 1998). Selective inhibition of COX-2 provides therapeutic benefit in inflammation without gastric ulceration (Kurumbail et al., 1996; Simmons et al., 2004) leading to an improved safety profile allowing the use of these agents for long term prophylaxis in certain chronic diseases (Kargman et al., 1995).

Certain furanone derivatives with COX-2 inhibitory activity were proved to possess analgesic, anti-inflammatory and anti-rheumatic activities. Several QSAR works have been carried out leading to the development of selective COX-2 inhibitors (Gupta et al., 2008; Singh et al., 2008). QSAR has been traditionally perceived as a means of establishing correlation between chemical structure modifications and respective changes in biological activity (Yao et al., 2004; Ravichandran et al., 2007; Gupta et al., 1998; Hansch and Fujita, 1964). Traditionally, the classic 2D QSAR model is only a rough approximation to the actual relationship as it mainly uses 2D molecular descriptors. Attempts have been made to incorporate 3D descriptors into QSAR model, while stepwise kNN MFA method using V-Life M.D.S. has been introduced (Jain et al., 2008; Fegade et al., 2009; Ajmani et al., 2006).

The purpose of the present study was to explore the significant physicochemical parameters responsible for the anti-inflammatory and antirheumatic activities of furanone derivatives. Consequently 2D and 3D QSAR analyses were carried out by multiple linear regressions (MLRs) and partial least square (PLS) techniques, respectively.

2. Experimental section

2.1. Materials and methods

A dataset of thirty-two furanone derivatives (Sing et al., 2004; Lau et al., 1999) for their COX-2 inhibitory activity have been taken for present QSAR work given in Table 1. All molecular modeling techniques including 2D and 3D QSAR studies described herein were performed on molecular modeling software V-Life MDS 3.5 (V-Life Sciences). Each structure of the 32 furanone derivatives was built on a workspace, fully geometrically optimized using standard Merck molecular force field (MMFF) with a distant dependent ($1/r$) dielectric function and energy gradient correlation ratio of 0.001 kcal/mol. In 3D QSAR, structure alignment is the most subjective and critical step, because earlier studies showed that the resulting 3D QSAR model is often sensitive to the particular alignment technique. Compounds having highest pIC_{50} values were selected as the template molecules and each molecule had to be superimposed onto the template molecule (template based alignment). These aligned molecules were placed into a three dimensional cubic lattice of the 2 Å grid. The electrostatic and steric fields were generated at each grid point using methyl probe (sp^3 hybridized) of charge +1 by the kNN method with default energy of 30 and 10 kcal/mol respectively and partial atomic charges were generated using the Gasteiger–Marsili method.

2.2. Descriptors

2D QSAR study was performed using 239 physicochemical, alignment independent and atom type descriptors. The various 2D descriptors were individual, retention index (chi), atomic valence connective index (chiv), path count, chi chain, chiv chain, chain path count, cluster, path cluster, kappa, element count, dipole moment, distance based topological, estate number, estate contributions, information theory index, semi empirical, hydrophobicity XlogpA, hydrophobicity XlogpK,

hydrophobicity SlogpA, hydrophobicity SlogpK, polar surface area, alignment independent and atom type descriptors class. In 3D QSAR, electrostatic and steric fields were generated around the aligned molecules in the grid. Negative value in electrostatic field descriptors (blue points in the dialog box) indicates that negative electronic potential is favorable for activity and more electronegative substituent groups are preferred at that position, and a positive electronic potential range indicates vice-versa. Negative range in steric field (green points in the dialog box) signifies that negative steric potential is required for activity and a less bulky substituents group is preferred in that region, a positive value of steric descriptors reveals that a positive steric potential is favorable for increase in activity and the more bulky group is preferred in that region.

2.3. Regression analysis

Models were generated using the multiple regression technique

for 2D and the partial least square analysis technique for 3D QSAR study. The optimal training and test sets were selected by random selection method which uses the ratio of training to validation objects (test set) as 70%: 30% (Coi et al., 2009). The cross-validation analysis was performed using the leave-one-out method. The cross-correlation limit was set at 0.5 and the term selection criteria is r^2 . F value was specified to evaluate the significance of a variable. The higher the F value, the more stringent was the significance level. The variance cutoff was set at 0.0, and scaling was auto scaling in which the number of random iterations was set at 100. The following statistical parameters were considered for comparison of the generated QSAR models: correlation coefficient (r), squared correlation coefficient (r^2), internal cross validation (q^2), predictive r^2 for external test set ($pred_r^2$) for external validation, and Fischer's (F). Internal validation was carried out using leave-one-out (LOO) method. For calculating q^2 , each molecule in the training set was eliminated once and the activity of the eliminated molecule was predicted by using the model developed by the remaining molecules.

The q^2 was calculated using the equation which describes the internal stability of a model

$$q^2 = \frac{1 - \sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2} \quad (1)$$

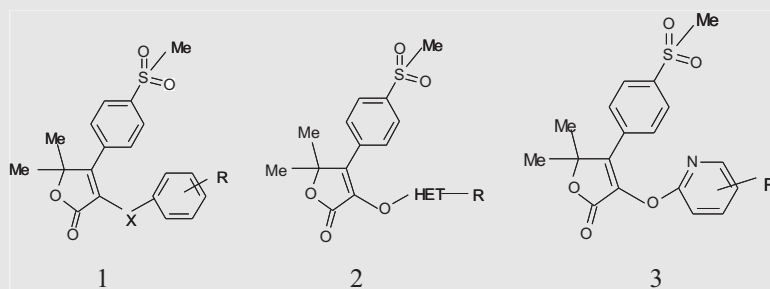
where y_i and \hat{y}_i are the actual and predicted activities of the i th molecule in the training set, respectively, and y_{mean} is the average activity of all molecules in the training set.

The $pred_r^2$ value indicates the predictive power of the current model for the external test set which is calculated as follows

$$pred_r^2 = \frac{1 - \sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2} \quad (2)$$

where y_i and \hat{y}_i are the actual and predicted activities of the i th molecule in the test set, respectively, and y_{mean} is the average activity of all molecules in the training set. Both summations are overall molecules in the test set.

Table 1 Structures and COX-2 inhibitory activity of furanone derivatives.



Compound no.	X	HET	R	<i>pIC</i> ₅₀ (μ M)
1a	O		H	1.699
1b	S		H	1.155
1c	NH		H	0.301
1d	CH ₂		H	0.523
1e	CO		H	0.222
2a	O		H	1.699
2b	O		H	2.000
2c	O		H	1.699
2d	O		H	1.698
2e	O		H	1.397
2f	O		H	2.000
2g	O		H	1.698
2h	O		H	1.698

Table 1 (continued)




















Compound no.	X	HET	R	$pIC_{50}(\mu M)$
2i	O		H	1.698
2j	O		H	1.698
2k	O		H	1.522
2l	O		H	0.481
2m	O		H	0.173
2n	O		H	-0.462
2o	O		H	0.259
2p	O		H	0.207
2q	O		H	0.568
2r	O		H	0.585
3a	O		3-F	0.920
3b	O		3,4-di F	1.698
3c	O		2,4-di F	0.455
3d	O		4-F	1.045
3e	O		3,5-di F	1.397
3f	O		4-Cl	0.824
3g	O		4-Me	1.523
3h	O		3-Me	1.398
3i	O		2-Me	1.699

Table 2 Unicolumn statistics of the training and test sets.

Data set	Average	Max.	Min.	Std. Dev.	Sum
2D QSAR					
Training	1.1970	2.0000	-0.4620	0.6911	29.924
Test	0.7965	1.5230	0.2080	0.4974	5.5690
3D QSAR					
Training	1.0352	2.0000	-0.4624	0.7139	22.7746
Test	1.3060	1.6990	0.2596	0.5424	10.4481

Table 3 Parameters value for the best 2D and 3D QSAR models generated.

Model	r^2	q^2	SE (r^2 se)	pred_ r^2	F-value	Descriptors
2D QSAR	0.840	0.773	0.195	0.618	36.852	chi6, SssO, PSA excluding P&S
3D QSAR	0.762	0.703	0.366	0.635	30.446	E_545, E_275, E_510, E_662

3. Results and discussion

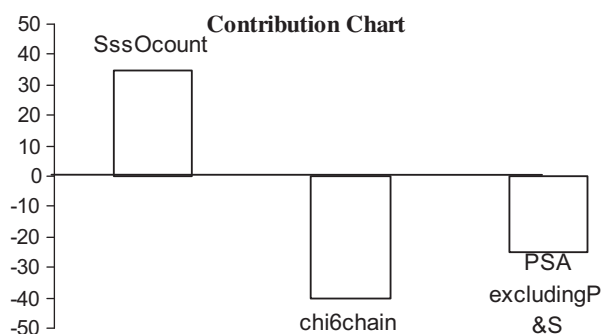
3.1. 2D results

QSAR studies on furanone series resulted in several QSAR equations using the multiple linear regression technique. Selection of test and training sets was based on uni-column statistics. Twenty-five compounds were placed in the training set and seven (**1b**, **2k**, **2o**, **2p**, **3a**, **3c**, and **3d**) compounds in the test set. Test and training sets were chosen randomly such that low, moderate and high-activity compounds were present in approximately the same proportions in both sets, which were confirmed by the results of uni-column statistics. The uni-column statistical analysis is summarized in **Table 2**. The uni-column statistics showed that the results are interpolative (i.e. derived within the minimum–maximum range of the training set).

The best equation obtained by multiple linear regressions is summarized here:

$$\text{Log}(1/\text{IC}_{50}) = -15.182(\pm 2.4221)\text{chi6} + 0.951(\pm 0.0468)\text{SssO} - 0.0321(\pm 0.0148)\text{PSA excluding P\&S} + 4.648.$$

Degrees of freedom = 21, $n = 25$, $r^2 = 0.840$, $q^2 = 0.773$, F test = 36.852, r^2 se = 0.195, q^2 se = 0.252, pred_ r^2 = 0.618, pred_ r^2 se = 0.158 (**Table 3**).

**Figure 1** Contribution chart of descriptors in 2D QSAR model.

IC_{50} is the molar concentration of the drug required for 50% inhibition of the COX-2 enzyme; n = number of data points; r^2 = squared correlation coefficient; SE = standard error of regression; F = F -ratio of calculated-to-observed value variances; q^2 = crossvalidated r^2 ; pred_ r^2 = external validation; chi6chain = retention index for six membered rings, SssOcount = total number of oxygen connected with two single bonds, PSA excluding P&S = polar surface area excluding P&S.

Above model was chosen as the best model as it showed a good correlation coefficient of 0.840 which explains 84% of the variance. The model showed an internal predictivity ($q^2 = 0.773$) of 77% and a predictivity for the external test (pred_ $r^2 = 0.618$) of 61%. The overall statistical significance level was found to exceed 99.9% with $F = 36.852$. The contribution of descriptor chart selected is mentioned in **Fig. 1**, which signifies the positive contribution of SssOcount (total number of oxygen connected with two single bonds), negative contribution of chi6chain (retention index for six membered ring) and polar surface area toward the biological activity.

The calculated and predicted (LOO) activities of the compounds by the above models are shown in **Table 4**. The graph between observed and predicted (LOO) activities is presented graphically in **Fig. 2**.

3.2. 3D results

In the present study the furanone derivative having the highest $p\text{IC}_{50}$ (compound no. **2b**) served as a template and each molecule was superimposed on the template via field fit alignment (**Fig. 3**). 3D QSAR model was derived using the kNN descriptors as independent variables and COX-2 inhibitor activity as a dependent variable. 3D-QSAR models were built in V-Life Molecular design Software using SWkNN method. Various models have been developed by changing training and test set data. The best selected model included compound no. **1a**, **2d**, **2o**, **3b**, **3d**, **3g**, **3i** set as test set data, while the remaining are training set compounds and **2i** and **2q** were removed as outliers. The results of uni-column statistics are summarized in **Table 2**, which shows that the test is interventive.

The final model was developed using the partial least square technique with an optimum number of components which

Table 4 The actual and predicted activities of the compounds by the above models.

Compound no.	Act. pIC_{50}	2D results MLR		3D results PLS	
		Pred. pIC_{50}	Res. pIC_{50}	Pred. pIC_{50}	Res. pIC_{50}
1a	1.699	1.3786	0.3204	1.561	0.138
1b	1.155	0.7242	0.4308	0.886	0.269
1c	0.301	0.338	-0.037	0.296	0.005
1d	0.523	0.7242	-0.2012	0.841	-0.318
1e	0.222	-0.0162	0.2382	0.795	-0.573
2a	1.699	1.6519	0.0471	0.774	0.925
2b	2.000	1.8948	0.1052	1.927	0.073
2c	1.699	1.8948	-0.1958	1.100	0.599
2d	1.699	1.6519	0.0471	2.024	-0.325
2e	1.398	1.8948	-0.4968	1.296	0.102
2f	2.000	1.6519	0.3481	1.728	0.272
2g	1.699	1.6519	0.0471	1.607	0.092
2h	1.699	1.6519	0.0471	1.529	0.17
2i	1.699	1.6519	0.0471	*	*
2j	1.699	1.8948	-0.1958	1.668	0.031
2k	1.523	1.8948	-0.3718	1.507	0.016
2l	0.481	0.2227	0.2583	0.678	-0.197
2m	0.174	0.2227	-0.0487	0.161	0.013
2n	-0.462	-0.1741	-0.2879	-0.196	-0.266
2o	0.260	0.585	-0.325	0.736	-0.476
2p	0.208	0.585	-0.377	-0.039	0.247
2q	0.569	0.9818	-0.4128	*	*
2r	0.585	0.9818	-0.3968	0.733	-0.148
3a	0.921	0.9818	-0.0608	1.189	-0.268
3b	1.699	1.2551	0.4439	1.147	0.552
3c	0.456	1.2551	-0.7991	0.765	-0.309
3d	1.046	1.2551	-0.2091	1.477	-0.431
3e	1.398	1.498	-0.1	1.575	-0.177
3f	0.824	1.2551	-0.4311	1.081	-0.257
3g	1.523	1.2551	0.2679	1.381	0.142
3h	1.398	1.2551	0.1429	1.953	-0.555
3i	1.699	1.2551	0.4439	1.891	-0.192

Notes: * – outliers, Act. pIC_{50} – actual activity, Pred. pIC_{50} – predicted activity, Res. pIC_{50} – residual activity.

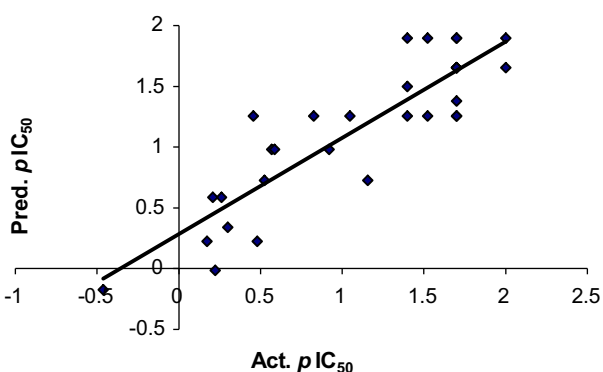


Figure 2 Correlation of actual and predicted activities in 2D QSAR model.

yields the highest q^2 . Selection of best model was performed on the basis of statistical parameters of the models.

For the selected kNN MFA model, the cross-validated r^2 (q^2) value of the training set was 0.703 with two principal components and standard error 0.409. The non-cross-validated r^2 value was 0.762 with standard error 0.366 and covariance ratio (F) of 30.446 (significant at 99% level). The predictive ability of model was also confirmed by external $pred_r^2$ having the

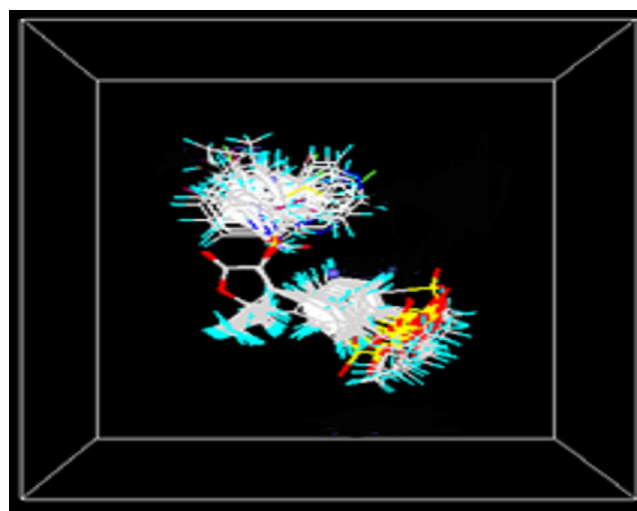


Figure 3 Aligned furanone molecules by template based alignment.

value 0.635 (Table 3). The correlation between experimental and predicted activities for both training and test sets of compounds are shown in Table 4 and represented graphically in

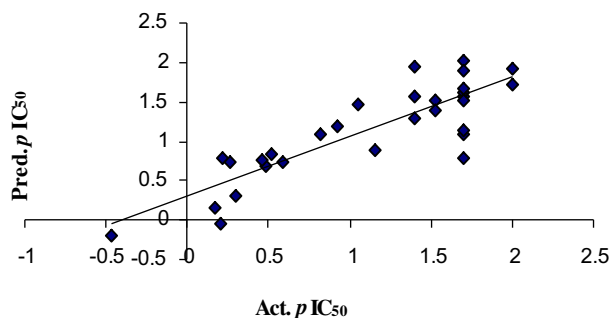


Figure 4 Correlation of experimental and predicted activities in 3D QSAR model.

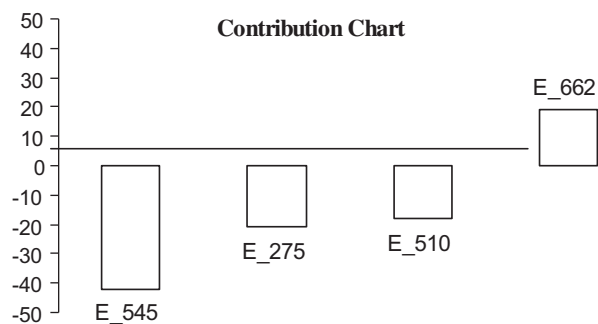


Figure 5 Contribution chart of steric and electrostatic parameters in 3D QSAR model.

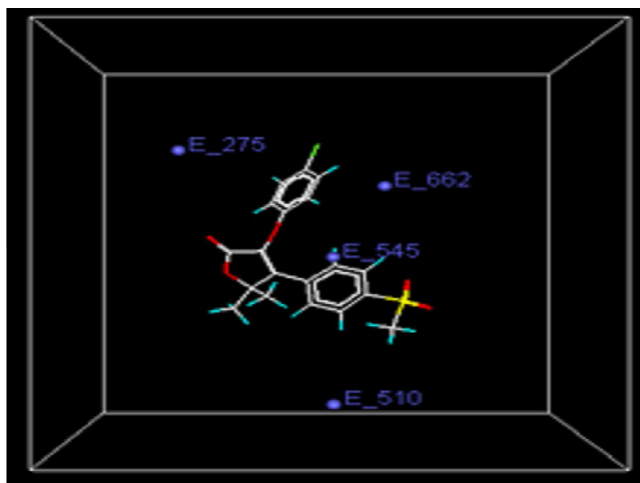


Figure 6 Model summary dialog box generated by [(SW) kNN MFA] method in 3D QSAR model shows steric and electrostatic regions.

Fig. 4 respectively. In kNN MFA model, the electrostatic parameter accounts for 100%. Contributions of electrostatic fields are shown in Fig. 5. The model summary dialog box, showed the relative positions of the local fields around the aligned molecule which were important for the activity variation in the model given in Figs. 6 and 7. In the dialog box, green points indicate the steric region and blue indicates the electronic region. Negative coefficients (-0.0996 , -0.0702 ,

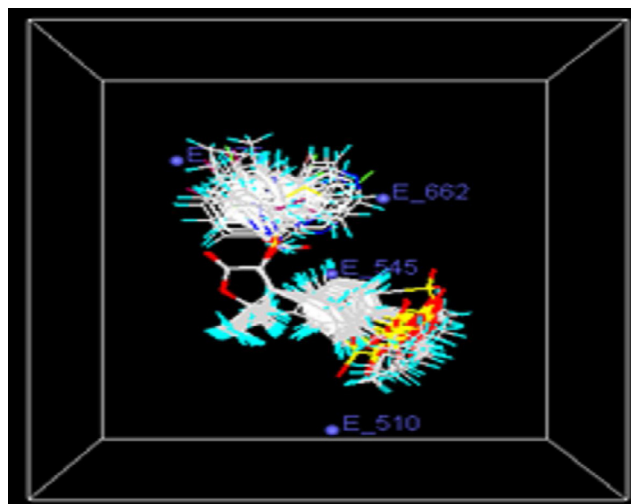


Figure 7 kNN MFA contour plots for steric and electrostatic regions. Green counters (S) indicate the steric region, whereas blue counters (E) indicate region where electrostatic contributions are required.

-0.0992) at positions E_545, E_275, E_510 respectively indicate that more electronegative groups like CF_3 , Cl, F, Br must be substituted at those positions while a positive coefficient (0.0639) at E_662 indicates that substitution of less electronegative group is favorable for activity.

4. Conclusions

The QSAR studies resulted in some significant 2D and 3D QSAR models. 2D QSAR model result showed the positive contribution of oxygen atoms and negative contribution of polarity and retention index toward biological activity. However 3D QSAR model showed only electrostatic constraint to be essential for biological activity. Hence, these models are very helpful for further optimization of anti-inflammatory activities. The current study provides better insight into the designing of more effective COX-2 inhibitors before their synthesis.

5. Conflict of interest

Authors report no declaration of interest.

References

- Ajmani, S., Jadhav, K., Kulkarni, S.A., 2006. Three-dimensional QSAR using the k-Nearest Neighbor method and its interpretation. *J. Chem. Inf. Model.* 46, 24–31.
- Botting, J.H., 1999. Nonsteroidal anti-inflammatory agents. *Drugs Today* 35, 225–235.
- Coi, A., Fiamingo, F.L., Livi, O., Calderone, V., Martelli, A., Massarelli, I., Bianucci, A.M., 2009. QSAR studies on BK channel activators. *Bioorg. Med. Chem.* 17, 319–325.
- Fegade, J.D., Rane, S.S., Chaudhari, R.Y., Patil, V.R., 2009. 3D-QSAR study of benzylidene derivatives as selective cyclooxygenase-2-inhibitors. *Digest. J. Nanomater. Biostruct.* 4, 145–154.
- Gupta, S.P., Babu, M.S., Garg, R., Sowmya, S., 1998. Quantitative structure-activity relationship studies on cyclic urea-based HIV protease inhibitors. *J. Enzyme Inhib.* 13, 399–407.

- Gupta, A.K., Gupta, R.A., Soni, L.K., Kaskhedikar, S.G., 2008. Exploration of physicochemical properties and molecular modeling studies of 2-sulfonyl-phenyl-3-phenyl-indole analogs as cyclooxygenase-2 inhibitors. *Eur. J. Med. Chem.* 43, 1297–1303.
- Hansch, C., Fujita, T., 1964. ρ - σ - π analysis. A method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.* 86, 1616–1626.
- Jain, A.K., Manocha, N., Ravichandran, V., Mourya, V.K., Agrawal, R.K., 2008. Three-dimensional QSAR study of 2, 4 – Disubstituted-phenoxy acetic acid derivatives as a CRTh2 receptor antagonist: using the k-Nearest Neighbor method. *Digest. J. Nanomater. Biostruc.* 3, 147–158.
- Kargman, S.L., O'Neill, G.P., Vickers, P.J., Evans, J.F., Mancini, J.A., Jothy, S., 1995. Expression of prostaglandin G/H synthase-1 and -2 protein in human colon cancer. *Cancer Res.* 55, 2556–2559.
- Kurumbail, R.G., Stevens, A.M., Gierse, J.K., McDonald, J.J., Stegeman, R.A., Pak, J.Y., Gildehaus, D., Iyashiro, J.M., Penning, T.D., Seibert, K., Isakson, P., Stalling, W.C., 1996. Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature* 384, 644–648.
- Lau, K.C., Brideau, C., Chan, S.K., Charelsion, S., Cromlish, W.A., Ethier, D., Gauthier, J.Y., Gordon, R., Guay, J., Kargman, S., Singh, L.C., Prasit, P., Riendeau, D., Therian, M., Visco, D.M., Lijing, X., 1999. Synthesis and biological evaluation of 3-heteroaryloxy-4-phenyl-2(5H)-furanones as selective COX-2 inhibitors. *Bioorg. Med. Chem. Lett.* 9, 3187–3192.
- Ravichandran, V., Mourya, V.K., Agrawal, R.K., 2007. QSAR study of novel 1, 1, 3-trioxo [1,2,4]-thiadiazine (TTDs) analogues as potent anti-HIV agents. *Arkivoc* 14, 204–212.
- Simmons, D.L., Botting, R.M., Hla, T., 2004. Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. *Pharmacol. Rev.* 56, 387–437.
- Sing, L.C., Black, W.C., Brideau, C., Chan, C.C., Charelsion, S., Cromlish, W.A., Claveau, D., Gauthier, Y.J., Gordon, R., Greig, G., Grimm, G., Guay, J., Lau, C.K., Riendeau, D., Shahapurkar, S., Pandya, T., Kawathekar, N., Chaturvedi, S.C., 2004. Quantitative structure activity relationship studies of diaryl furanones as selective COX-2 inhibitors. *Eur. J. Med. Chem.* 39, 383–388.
- Singh, P., Mittal, A., Kaur, S., Kumar, S., 2008. 2,3,5-Substituted tetrahydrofurans: COX-2 inhibitory activities of 5-hydroxymethyl-/carboxyl-2, 3-diaryl-tetrahydro-furan-3-ols. *Eur. J. Med. Chem.* 43, 2792–2799.
- Vane, J.R., Bakhle, Y.S., Botting, R.M., 1998. Cyclooxygenases 1 and 2. *Annu. Rev. Pharmacol. Toxicol.* 38, 97–120.
- V-Life Molecular Design Suite 3.0, V Life Sciences Technologies Pvt. Ltd., Baner Road: Pune, Maharashtra, India.
- Yao, X.J., Panaye, A., Doucet, J.P., Zhang, R.S., Chen, H.F., Liu, M.C., Hu, Z.D., Fan, B.T., 2004. Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks, and multiple linear regressions. *J. Chem. Inf. Comput. Sci.* 44, 1257–1266.