4. STUDIES ON SELECTIVE COX-II INHIBITORS

4.1 Quantitative structure activity relationship studies on diarylimidazoles as selective COX-2 inhibitors

4.1.1 Experimental data

4.1.1.1 Data sets used

The structures (**Figure 4.1.1**) and activities in IC_{50} (IC_{50} is the concentration in μM for 50% inhibition of the COX-2 or COX-1 enzymes) for diarylimidazoles were extracted from literature[1] and gathered in **Tables 4.1.3** and **4.1.4**.

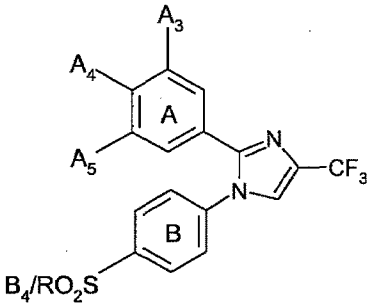


Figure 4.1.1 – Lead compounds for present study [A_3, A_4, A_5 and B_4 represent various positions of substituents on the basic skeleton].

Table 4.1.1 Compounds with ring 'A' and 'B' substitutions and values of descriptors used in training set for COX-2 inhibitory activity

C. N	A-ring substitution			B-ring substi tution	$\sum\prod$ (A_3 + A_4 + A_5)	V_wA_3	σ A_4	σ(A_3 + A_5)	σ B_4
	A_3	A_4	A_5						
1	H	F	H	Me	0.14	7.238	0.06	0.0	0.72
2	H	H	H	Me	0.0	7.238	0.0	0.0	0.72
3	H	Me	H	Me	0.56	7.238	-0.17	0.0	0.72
4	H	NMe_2	H	Me	0.18	7.238	-0.83	0.0	0.72
5	H	SMe	H	Me	0.61	7.238	0.0	0.0	0.72
6	H	SO_2Me	H	Me	-1.63	7.238	0.72	0.0	0.72
7	H	Cl	H	NH_2	0.71	7.238	0.13	0.0	0.57
8	H	H	H	NH_2	0.0	7.238	0.0	0.0	0.57
9	H	Me	H	NH_2	0.56	7.238	-0.17	0.0	0.57

10	Cl	H	H	Me	0.71	7.238	0.0	0.37	0.72
11	F	H	H	Me	0.14	7.238	0.0	0.34	0.72
12	Br	H	H	Me	0.86	22.44	0.0	0.39	0.72
13	CF_3	H	H	Me	0.88	31.05	0.0	0.43	0.72
14	OMe	H	H	Me	-0.02	26.60	0.0	0.12	0.72
15	SMe	H	H	Me	0.61	38.98	0.0	0.15	0.72
16	CH_2OMe	H	H	Me	-0.78	34.52	0.0	0.02	0.72
17	NMe_2	H	H	Me	0.18	43.57	0.0	-0.15	0.72
18	NO_2	H	H	Me	-0.28	52.42	0.0	0.71	0.57
19	Cl	H	H	NH_2	0.71	38.30	0.0	0.37	0.57
20	F	H	H	NH_2	0.14	21.02	0.0	0.34	0.57
21	Br	H	H	NH_2	0.86	31.43	0.0	0.39	0.57
22	Me	H	H	NH_2	0.56	22.44	0.0	-0.07	0.57
23	Cl	OMe	H	Me	0.69	12.24	-0.27	0.34	0.72
24	Cl	NMe_2	H	Me	0.89	26.67	-0.83	0.37	0.72
25	F	NMe_2	H	Me	0.32	22.44	-0.83	0.34	0.72
26	Cl	NHMe	H	Me	0.24	22.44	-0.84	0.37	0.72
27	F	Me	H	Me	0.70	22.44	-0.17	0.34	0.72
28	Me	F	H	Me	0.70	22.11	0.06	-0.07	0.72
29	Me	Cl	H	Me	1.27	22.44	0.23	-0.07	0.72
30	OMe	Cl	H	Me	0.69	12.24	0.23	0.12	0.72
31	NMe_2	Cl	H	Me	0.89	26.60	0.23	-0.15	0.72
32	F	F	H	Me	1.12	26.59	0.06	0.34	0.72
33	Me	H	Cl	Me	0.70	55.73	0.0	0.30	0.72
34	Me	H	F	Me	0.12	12.24	0.0	0.27	0.72
35	CF_3	H	F	Me	1.42	26.57	Out	0.77	0.72
36	Cl	H	Cl	Me	0.69	26.61	0.0	0.74	0.72
37	F	OMe	H	NH_2	0.69	34.53	-0.27	0.37	0.57
38	Cl	OMe	H	NH_2	0.84	38.96	-0.27	0.37	0.57
39	Br	OMe	H	NH_2	1.32	22.44	-0.27	0.39	0.57
40	Cl	SMe	H	NH_2	1.27	22.44	0.0	0.37	0.57
41	OMe	Cl	H	NH_2	0.28	22.44	0.23	0.12	0.57
42	F	F	H	NH_2	1.27	22.44	0.06	0.34	0.57
43	Me	H	F	NH_2	0.12	36.59	0.0	0.27	0.57
44	OMe	H	F	NH_2	0.26	12.24	Out	0.46	0.57
45	F	OMe	F	Me	1.40	26.52	-0.27	0.68	0.72
46	Cl	OMe	Cl	Me	1.70	26.68	-0.27	0.74	0.72

47	Br	OMe	Br	Me	1.10	34.45	-0.27	0.78	0.72
48	Cl	NMe_2	Cl	Me	0.26	12.24	-0.83	0.74	0.72
49	F	OMe	F	NH_2	0.26	31.05	-0.27	0.68	0.57

Table 4.1.2 Compounds with ring 'A' and 'B' substitutions and values of descriptors used in test set for COX-2 inhibitory activity.

C.N	A-ring substitution			B-ring substitution	$\sum\prod$ ($A_3 + A_4 + A_5$)	V_wA_3	σA_4	σ ($A_3 + A_5$)	σB_4
	A_3	A_4	A_5						
50	H	Cl	H	Me	0.71	7.238	0.23	0.0	0.72
51	H	OMe	H	Me	-0.02	7.238	-0.27	0.0	0.72
52	H	NHMe	H	Me	-0.47	7.238	-0.84	0.0	0.72
53	H	SOMe	H	Me	-1.58	7.238	0.49	0.0	0.72
54	H	F	H	NH_2	0.14	7.238	0.06	0.0	0.57
55	Me	H	H	Me	0.56	12.24	0.0	-0.07	0.72
56	NHMe	H	H	Me	-0.47	52.42	0.0	-0.30	0.72
57	NH_2	H	H	Me	-1.23	55.62	0.0	-0.16	0.72
58	F	OMe	H	Me	0.12	31.05	-0.27	0.34	0.72
59	Cl	SMe	H	Me	1.32	12.24	0.0	0.37	0.72
60	Cl	Me	H	Me	1.27	12.24	-0.17	0.37	0.72
61	Me	Me	H	Me	1.27	12.24	-0.17	-0.07	0.72
62	OMe	H	F	Me	1.02	26.53	0.0	0.46	0.72
63	Cl	Me	H	NH_2	0.69	31.05	-0.17	0.37	0.57

4.1.1.2 Results and discussion

The 63 compounds belonging to the diarylimidazole category were divided in two sets, 49 compounds were taken into training set and 14 compounds constituted the test set. The IC_{50} values for both COX-1 and COX-2 were transformed into pIC_{50} as described earlier. Stepwise regression analysis was performed by taking pIC_{50} value as dependent variable and different structural descriptors as independent variables. A large number of equations were generated; one of the best equations out of them was Equation 1, but even this equation was not statistically significant with large standard error of prediction (0.734).

$$pIC_{50}\ (COX\text{-}2) = (0.8723 \pm 0.210)\ \Sigma\ \Pi\ (A_3 + A_4 + A_5) - (0.560 \pm 0.126)\ [\Sigma\ \Pi\ (A_3 + A_4 + A_5)]^2 + (0.575 \pm 0.280)\ \sigma\ [A_3 + A_5] - (0.807 \pm 0.610)\ \sigma\ B_4 - (0.0026 \pm 0.0020) V_w A_3 + (0.208 \pm 0.186)\ \sigma A_4 + (7.667 \pm 0.516) \quad \ldots\ldots (1)$$

$n = 49,\ r = 0.583,\ S.E = 0.734,\ S.D = 0.60,\ r^2_{cv} = 0.274,\ F_{ratio} = 3.61$

This model was capable of explaining only 34 % of the variations. The cause of poor statistical figures were two compounds **(16** and **35),** whose calculated activities were showing too much deviation from the observed values and hence were considered to be outliers **(Table 4.1.3).** After excluding these two compounds a much more robust model (Equation 2) with the same descriptors was obtained which could explain 85.7% of the variance in the observed activity values. The correlation matrix for descriptors influencing COX-2 inhibitory activity is shown in **Table 4.1.5.** The predicted activity for the training set is shown in **Table 4.1.3.**

Table 4.1.3 Experimental and calculated biological activity of molecules used in training set for COX-2 inhibitors, with their selectivity ratio.

C. N	COX-2 inhibitory activity			COX-1 inhibition		Selectivity Log(IC_{50}COX-1/ IC_{50} COX-2)[a]	Selectivity Log(IC_{50}COX-1/ IC_{50} COX-2)[c]
	IC_{50}	pIC_{50} [a]	pIC_{50} [b]	IC_{50}	pIC_{50} [a]		
1	0.10	7.0	6.71	36	4.44	2.55	2.68
2	0.12	6.92	6.57	7.23	3.10	3.77	Outlier
3	0.16	6.79	6.77	26	4.58	2.21	1.98
4	0.70	6.15	6.17	5	5.30	0.85	0.81
5	0.16	6.79	6.90	2.1	5.67	1.11	Outlier
6	5.70	5.24	4.85	>100	-	-	-
7	0.01	8.00	7.95	1.6	5.79	2.20	2.37
8	0.04	7.39	7.45	19.3	4.71	2.68	2.50
9	0.04	7.39	7.65	4.6	5.33	2.06	1.71
10	0.06	7.22	7.29	360	3.44	3.77	3.35
11	0.12	6.92	7.01	>1000	-	-	-
12	0.08	7.00	7.02	>100	-	-	-
13	0.21	6.67	6.88	>100	-	-	-
14	0.35	6.45	6.27	>100	-	-	-
15	0.35	6.45	6.21	>100	-	-	-

16	68.1	4.16	Outlier	>100	-	-	-
17	3.20	5.49	5.79	42.2	4.37	1.12	1.59
18	0.58	6.23	6.39	>100	-	-	-
19	0.0008	8.00	7.88	6.20	5.20	2.88	2.95
20	0.03	7.52	7.38	67.7	4.16	3.35	3.19
21	0.007	8.10	7.90	4.50	5.34	2.80	3.00
22	0.03	7.52	7.58	3.20	5.49	2.02	1.92
23	0.13	6.86	6.71	296	3.52	3.35	2.69
24	0.32	6.49	6.48	1.56	5.80	0.68	1.37
25	0.33	6.48	6.27	17.1	4.76	1.71	Outlier
26	0.66	6.18	6.25	>100	-	-	-
27	0.11	6.95	6.84	>100	-	-	-
28	0.17	6.76	6.57	24.1	4.61	2.15	1.91
29	0.09	7.04	6.92	7.84	5.10	1.94	1.95
30	0.25	6.60	6.78	>100	-	-	-
31	1.04	5.98	6.55	>100	-	-	-
32	0.12	6.92	6.32	>100	-	-	-
33	0.08	7.09	7.12	>1000	-	-	-
34	0.11	6.95	6.52	>100	-	-	-
35	0.96	4.00	Outlier	67	4.17	-	-
36	0.17	6.76	7.09	>100	-	-	-

[a] experimental value, [b] predicted value by equation 2, [c] predicted value by equation 5

$$pIC_{50}\ (COX\text{-}2) = (0.639 \pm 0.173)\ \Sigma\,\Pi\ (A_3 + A_4 + A_5) - (0.3617 \pm 0.077)\ [\Sigma\,\Pi\ (A_3 + A_4 + A_5)]^2 + (0.994 \pm 0.212)\ \sigma\ [A_3 + A_5] - (5.834 \pm 0.648)\ \sigma\ B_4 - (0.020 \pm 0.0041)\ V_wA_3 + (0.6095 \pm 0.105)\ \sigma\ A_4 + (10.928 \pm 0.440) \quad \ldots\ldots (2)$$

$n = 47,\ r = 0.927,\ S.E = 0.302,\ S.D = 0.279,\ r^2_{cv} = 0.546,\ F_{ratio} = 40.04$

This equation showed a positive correlation of $\Sigma\,\Pi\ (A_3 + A_4 + A_5)$ and $\sigma\ [A_3 + A_5)$ with COX-2 inhibition and negative correlation of V_wA_3 with COX-2 inhibitory activity. Equation 2 was used to predict the activity of the test set **(Table 4.1.4)**. The comparative graph of experimental *versus* predicted activity for the test set is shown in **Figure 4.1.2.** B_4 position is very critical for hydrogen bonding of the ligand in the

COX-2 active site. Hence, only the reported substituents SO_2NH_2 and SO_2CH_3 at this position were considered for analyses.

Table 4.1.4 Experimental and calculated COX-2 inhibiting activity of molecules used in test set with their selectivity ratio.

C. N	COX-2 inhibitory activity			COX-1 inhibition		Selectivity [a] Log(IC_{50}COX-1/ IC_{50} COX-2)	Selectivity [c] Log(IC_{50}COX-1/ IC_{50} COX-2)
	IC_{50}	pIC_{50} [a]	pIC_{50} [b]	IC_{50}	pIC_{50} [a]		
43	0.11	6.95	7.07	23	4.63	2.32	2.40
44	1.47	5.83	5.60	53.5	4.27	1.56	0.99
45	0.96	4.0	4.80	>100	-	-	-
46	0.01	8.0	7.59	1.9	5.72	2.27	2.45
47	0.92	6.03	5.00	>100	-	-	-
48	5.89	5.22	4.03	>100	-	-	-
49	0.15	6.82	6.33	49	4.30	2.51	2.62
50	0.04	7.39	7.20	>100	-	-	-
51	0.03	7.52	7.10	12	4.9	2.60	2.88
52	0.33	6.48	6.82	30	4.52	1.95	1.77
53	0.96	6.01	7.02	>100	-	-	-
54	0.003	8.52	7.56	0.57	5.45	2.27	2.90
55	0.04	7.39	7.52	>100	-	-	-
56	0.72	6.14	6.22	91	4.0	-	-

[a] experimental value, [b]predicted value by Equation 2, [c] predicted value by equation 5

Table 4.1.5 Correlation matrix for descriptors influencing COX-2 inhibitory activity

	pIC_{50}	σA_4	$\sum\prod(A_3 + A_4 + A_5)$	$[\sum\prod(A_3 + A_4 + A_5)]^2$	$\sigma(A_3 + A_5)$	σB_4	V_wA_3
pIC_{50}	1.0						
σA_4	0.029	1.0					
$\sum\prod(A_3+A_4+A_5)$	0.518	0.194	1.0				
$[\sum\prod(A_3+A_4+A_5)]^2$	-0.039	0.219	0.379	1.0			
$\sigma(A_3+ A_5)$	0.228	0.390	0.318	0.226	1.0		
σB_4	-0.696	0.111	0.121	0.032	0.006	1.0	
V_wA_3	-0.205	0.069	0.191	0.127	0.366	0.091	1.0

The positive contribution of $\Sigma\ \Pi\ (A_3 + A_4 + A_5)$ indicates that substituents with higher Π values at position A_3, A_4 and A_5 would be favorable for the COX-2 inhibitory activity. The negative contribution of V_wA_3 clearly indicates that the vanderWaal's volume of the subsituent at position A_3 must be as low as possible in order to be favorable for COX-2 inhibitory activity. The positive correlation of σA_4 also highlighted the importance of electrophilicity of substituents at A_4, hence the substituents at this position should be more powerful electron withdrawing groups. As per this model, the COX-2 inhibitory activity was calculated for the test set **(Table 4.1.4)**. A comparison of the experimental and calculated values (using Equation 2) for COX-2 inhibitory activity in the form of a graph is depicted in **Figure 4.1.2**.

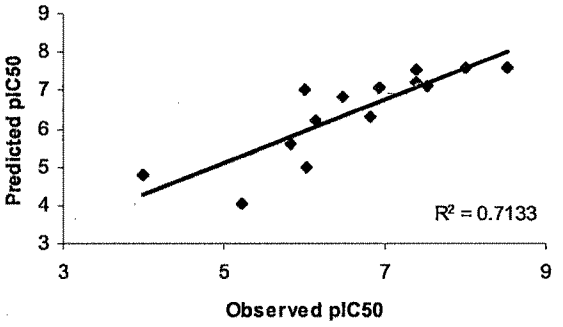


Figure 4.1.2 Graph between experimental and predicted COX-2 inhibitory activity for compounds of test set.

When COX-1 inhibitory data was subjected to regression analysis with those variables previously used for COX-2, a good correlation could not be obtained. The correlation obtained here is shown in Equation 3.

$$pIC_{50}\ (COX\text{-}1) = (0.1848 \pm 0.168)\ \Sigma\ \Pi\ (A_3 + A_4 + A_5) - (0.2261 \pm 0.143)\ [\Sigma\ \Pi\ (A_3 + A_4 + A_5)]^2 + (1.437 \pm 0.394)\ \sigma\ [A_3 + A_5] - (6.938 \pm 1.215)\ \sigma\ B_4 - (0.0024 \pm 0.007)\ V_wA_3 + (0.9209 \pm 0.322)\ \sigma\ A_4 + (9.0126 \pm 0.824) \quad \ldots\ldots (3)$$

$n = 35$, $r = 0.645$, $S.E = 0.5612$, $S.D = 0.515$, $F_{ratio} = 8.32$, $r^2_{cv} = 0.274$

This model explains only 55.5 % variance with high standard error of regression. Here, it can be seen that the contribution of V_wA_3 is very low. The role of summed Π values at A_3, A_4 and A_5 positions was also ambiguous as the standard error was quite high (± 0.168).

One of the aims of the present study was to identify the structural features which impart selectivity to these compounds for COX-2 enzyme over COX-1. To achieve this aim, structural descriptors used in Equation 2 were regressed against selectivity ratio [log (IC_{50} COX-1/IC_{50} COX-2)].

$$\log (\text{COX-1/COX-2}) = (0.7242 \pm 0.230)\ \Sigma\ \Pi\ (A_3 + A_4 + A_5) - (-0.115 \pm 0.08)\ [\Sigma\ \Pi\ (A_3 + A_4 + A_5)]^2 + (3.225 \pm 0.562)\ \sigma\ [A_3 + A_5] - (1.038 \pm 0.392)\ \sigma\ B_4 - (0.0125 \pm 0.0041)\ V_wA_3 + (2.050 \pm 0.32)\ \sigma\ A_4 + (1.997 \pm 0.462)$$

$$n = 26,\ r = 0.750,\ S.E = 0.46,\ S.D = 0.570,\ r^2_{cv} = 0.546,\ F_{ratio} = 6.05 \quad \ldots\ldots (4)$$

The number of descriptors used in Equation 4 are on a higher side. After removing outliers (**2, 5** and **25**) and optimizing the number of descriptors Equation 5 was obtained.

$$\log (\text{COX-1/COX-2}) = (2.710 \pm 0.20) - (0.722 \pm 0.230)\ \Sigma\ \Pi\ (A_3 + A_4 + A_5) + (3.44 \pm 0.510)\ \sigma\ [A_3 + A_5] - (0.012 \pm 0.0084)\ V_wA_3 + (2.050 \pm 0.32)\sigma\ A_4$$

$$n = 23,\ r = 0.8539,\ S.E = 0.40,\ S.D = 0.4148,\ r^2_{cv} = 0.600 \quad \ldots\ldots (5)$$

This equation offered a much better correlation in terms of statistics. It is interesting to note that summed Π values of three positions in ring-A have now a negative contribution towards selectivity and the positive contribution of electronic parameter (σA_4) further increased to impart higher selectivity to the compounds. Equation 5 was used to predict the selectivity of test compounds **(Table 4.1.4)** and a plot of the experimental *versus* calculated values is shown in **Figure 4.1.3**.

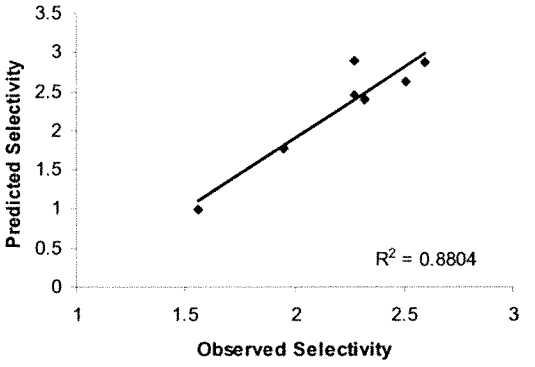


Figure 4.1.3 Graph between experimental and predicted COX-2 selectivity for test set of compounds.

Based on the correlations [Equation 2 (for COX-II inhibitory activity) and 5 (for selectivity ratio)] obtained above, it was planned to design new molecules having the diarylimidazole skeleton with much higher selectivity than for those compounds which are reported. Earlier it was observed that SO_2CH_3 group at B_4 position had higher contribution over SO_2NH_2 for COX-II selectivity. So, all these compounds **(D1-10, Table 4.1.6)** have been designed with SO_2CH_3 group at B_4 position. All the designed compounds have much higher selectivity for COX-2 enzyme with retention of high inhibitory activity.

It can be noted that the reported compound (10) has the highest selectivity ratio of 3.35 while all the designed compounds have higher selectivity than compound (10) with the best ratio for the designed compound (D1) with a selectivity ratio of 5.24. In the reported series, compound (54) has the highest COX-II inhibitory activity but its selectivity towards the COX-II enzyme is poor (selectivity ratio 2.27).

Table 4.1.6 Structures, descriptors, selectivity ratio and COX-2 inhibitory activity of the designed compounds

C. N	A_3	A_4	A_5	(Pi)*	$[Pi)^*]^2$	$\sigma(A_3 + A_5)$	V_wA_3	σA_4	(S)*	pIC_{50} COX-2
D1	Cl	SO_2CF_3	F	1.4	1.96	0.71	22.449	0.93	5.24	7.93
D2	Cl	SCOEt	F	1.49	2.22	0.71	22.449	0.44	4.12	7.54
D3	Cl	SO_2CF_3	Cl	1.97	3.88	0.74	22.449	0.93	4.92	7.72
D4	Br	SO_2CF_3	Cl	2.12	4.49	0.77	31.059	0.93	4.77	7.47
D5	Br	SO_2CF_3	F	1.55	2.4	0.73	31.059	0.93	5.06	7.72
D6	Br	SO_2CF_3	Br	2.27	5.15	0.78	31.059	0.93	4.69	7.36
D7	SMe	SO_2CF_3	F	1.30	1.69	0.49	43.578	0.93	4.35	7.55
D8	SMe	SO_2CF_3	Cl	1.87	3.49	0.52	43.578	0.93	4.02	7.38
D9	SMe	SCOEt	Cl	1.96	3.84	0.52	43.578	0.44	3.95	6.70
D10	SMe	SO_2CF_3	Br	2.02	4.08	0.55	43.578	0.93	4.0	7.32

(S)*- Refers to selectivity ratio calculated by Equation 5; $(Pi)^* = \sum \prod (A_3 + A_4 + A_5)$

In nutshell a good correlation (Equation 2) could be obtained between COX-II inhibitory activity and certain structural descriptors in diarylimidazoles. The selectivity ratio could also be correlated well (Equation 5) with structural descriptors like summed hydrophobic constants for substituents A_3, A_4 and A_5, the vander Waal's volume of A_3 substituent, and Hammett sigma constants for the three substituent positions. On the basis of this equation (Equation 5) some compounds have been designed which are predicted to have a very high selectivity for COX-II enzyme with high to very high inhibitory activity (predicted by Equation 2).

4.2 3D-QSAR CoMFA/CoMSIA studies on 5-aryl-2,2-dialkyl-4-phenyl-3(2*H*)- furanone derivatives, as selective COX-2 inhibitors

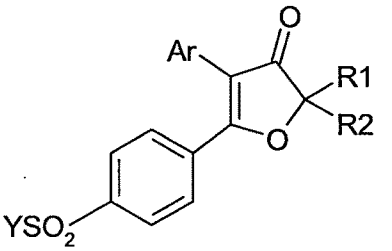
4.2.1 Experimental

4.2.1.1 Data set and alignment rules

Data set

A series of 5-aryl-2,2-dialkyl-4-phenyl-3(2*H*)-furanone derivatives have been reported as selective COX-2 inhibitors[2]. *In vitro* assay for COX-2 inhibition has been reported to be performed using human whole blood method. **Table 4.2.1** defines structure and biological activity of training set and test set compounds.

Table 4.2.1 Structures and biological activities of training and test set of molecules.

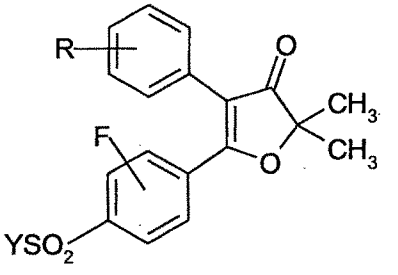


1-29; Ar = phenyl/substituted phenyl

C.N	Ar	R_1	R_2	Y	Biological activity [a]		
					Actual	Calculated	
						CoMFA [b]	CoMSIA
1	phenyl	CH_3	CH_3	CH_3	7.30	7.24	7.11
2	3-fluorophenyl	CH_3	CH_3	CH_3	7.69	7.59	7.32
3	3-chlorophenyl	CH_3	CH_3	CH_3	8.00	6.28	6.60
4	4-bromophenyl	CH_3	CH_3	CH_3	7.52	7.59	7.70
5[c]	3-chloro-4-fluorophenyl	CH_3	CH_3	CH_3	6.52	7.81	7.67
6	3,5-difluorophenyl	CH_3	CH_3	CH_3	7.30	7.51	7.34
7	3,5-dichlorophenyl	CH_3	CH_3	CH_3	7.52	7.71	7.68
8[c]	3-(trifluoromethyl)-phenyl	CH_3	CH_3	CH_3	7.30	7.96	7.63
9	3-isopropylphenyl	CH_3	CH_3	CH_3	7.52	7.52	7.59

10	4-isopropylphenyl	CH_3	CH_3	CH_3	7.52	7.47	7.43
11[c]	4-n-butylphenyl	CH_3	CH_3	CH_3	7.69	7.34	7.25
12	3-acetylphenyl	CH_3	CH_3	CH_3	7.30	7.44	7.15
13	4-acetylphenyl	CH_3	CH_3	CH_3	6.52	6.68	6.68
14	3-methoxyphenyl	CH_3	CH_3	CH_3	7.00	7.03	7.03
15	phenyl	CH_3	C_2H_5	CH_3	6.69	6.74	6.84
16	2-fluorophenyl	CH_3	C_2H_5	CH_3	6.52	6.72	6.79
17	3-chlorophenyl	CH_3	C_2H_5	CH_3	7.52	7.31	7.32
18	3,5-dichlorophenyl	CH_3	C_2H_5	CH_3	7.52	7.34	7.40
19	3-(trifluoromethyl)-phenyl	CH_3	C_2H_5	CH_3	7.30	7.33	7.40
20[c]	3-methoxyphenyl	CH_3	C_2H_5	CH_3	6.52	7.68	6.79
21	phenyl	C_2H_5	C_2H_5	CH_3	6.69	6.82	6.92
22	4-(trifluoromethyl)-phenyl	C_2H_5	C_2H_5	CH_3	7.52	7.55	7.42
23[c]	3-methylphenyl	C_2H_5	C_2H_5	CH_3	7.30	7.09	7.19
24	4-methylphenyl	C_2H_5	C_2H_5	CH_3	7.52	7.33	7.17
25	phenyl	CH_3	CH_3	NH_2	8.09	7.96	8.13
26	3-fluorophenyl	CH_3	CH_3	NH_2	8.52	8.53	8.35
27	3,4-difluorophenyl	CH_3	CH_3	NH_2	8.15	7.05	7.25
28[c]	3-(trifluoromethyl)-phenyl	CH_3	CH_3	NH_2	7.95	8.89	8.67
29	4-acetylphenyl	CH_3	CH_3	NH_2	7.88	7.61	7.70

[a] Biological activity expressed as -log IC_{50} against human COX-2 enzyme in M/L^{-1}
[b] Calculated activity from alignment IV; [c] Test set molecules

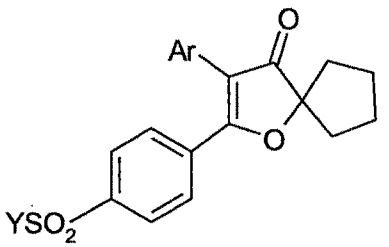


30-43

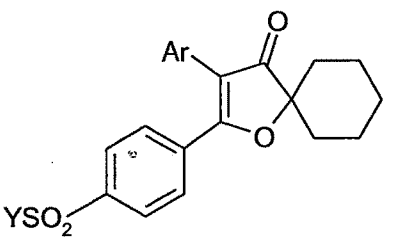
C.N	Fluoride Position	R	Y	Biological activity [a]		
				Actual	Calculated	
					CoMFA [b]	CoMSIA
30[c]	2	H	CH_3	7.52	7.21	7.24
31[c]		3-F	CH_3	7.52	7.18	7.32

32		3-Cl	CH_3	7.69	7.74	7.68
33		3-F, 5-F	CH_3	7.52	7.58	7.42
34[c]	3	H	CH_3	7.52	6.64	7.11
35		3-F	CH_3	7.52	7.09	7.32
36		3-Cl	CH_3	7.36	7.20	7.59
37[c]		3-F, 5-F	CH_3	7.52	6.95	7.33
38[c]	2	H	NH_2	7.52	7.96	8.40
39		3-F	NH_2	8.52	8.55	8.62
40		3-F, 5-F	NH_2	8.52	8.45	8.64
41	3	H	NH_2	7.52	7.45	7.45
42		3-F	NH_2	7.52	7.89	7.66
43[c]		3-F, 5-F	NH_2	7.69	7.79	7.67

[a] Biological activity expressed as -log IC_{50} against human COX-2 enzyme in M/L^{-1} [b] Calculated activity from alignment IV.
[c] Test set molecules



44-47



48-49

Ar - phenyl/substituted phenyl

C.N	Ar	Biological activity[a]		
		Actual	Calculated	
			CoMFA [b]	CoMSIA
44[c]	phenyl	7.52	7.08	7.12
45	3-fluorophenyl	6.47	7.93	7.80
46	3-methylphenyl	7.52	7.36	7.39

47	4-isopropylphenyl	7.30	7.40	7.44
48	phenyl	6.30	6.16	6.15
49	3-methylphenyl	6.30	6.46	6.46

[a] Biological activity expressed as -log IC_{50} against human COX-II enzyme in M/L^{-1}
[b] Calculated activity from alignment IV; [c] Test set molecules

Alignment rules

The atoms and centroids selected in different 3D QSAR studies are depicted separately, after the datasets. **Figure 4.2.1** below shows the superimposition of 49 furanone derivatives obtained by alignment IV. The atoms and centroids used for alignments are defined in **Figure 4.2.2**. These alignments were subsequently used in CoMFA/CoMSIA probe interaction energy calculations.

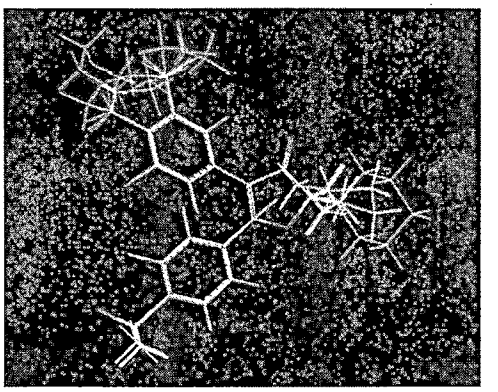
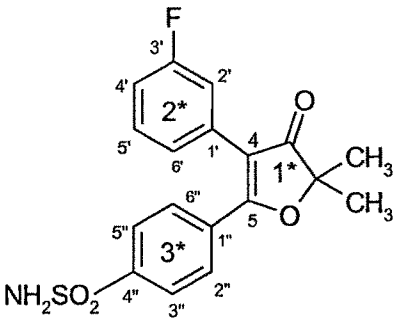


Figure 4.2.1 Superimposition of 49 molecules obtained by alignment IV



Alignment	Atoms/centroids
I	Database alignment
II	2, 4, 1', 5, 1"
III	1*, 2*, 3*
IV	1*, 4, 1', 5, 1''

Figure 4.2.2 Alignment rules with template molecule **29**

4.2.2 Results and discussion

3D-QSAR has certain advantages over the 2D-QSAR technique. Hence, CoMFA and CoMSIA techniques were used to derive 3D-QSAR models for 5-aryl-2,2-dialkyl-4-phenyl-3(2*H*)-furanones (**Table 4.2.1**) as selective COX-2 inhibitors. **Figure 4.2.1** and **4.2.2** depicts the alignment rules employed in the study and superimposition of the molecules, respectively. The negative logarithm of the *in vitro* inhibitory activity (pIC_{50}) was used as a dependent variable, as described earlier. The lower energy conformers obtained from MULTISEARCH option in SYBYL were used in the study. All the molecules were aligned employing atom/shape based RMS fitting and RMSD based database fitting techniques. Various 3D-QSAR models were generated and the best one was selected on the basis of statistically significant parameters.

Preliminary studies performed using 39 molecules in training set revealed the significance of CoMFA parameters on final results. PLS analysis was performed using

varying column filtering values. Finally column filtering was set to 0.0 and used for further calculations in PLS analysis.

Analysis A shows CoMFA results **(Table 4.2.2)** obtained from four different alignments using 39 molecules in training set. The database alignment I showed cross-validated r^2 0.416 with four components, non cross-validated r^2 0.809, F value 18.850, bootstrapped r^2 0.909, the steric and electrostatic contributions were 70.1 % and 29.9 %, respectively. The CoMFA model generated from atom based RMS alignment II **(Table 4.2.2)** showed cross-validated r^2 of 0.389 with four components, non cross-validated r^2 0.846, F value 43.806, bootstrapped r^2 0.899 with 60.2 % steric and 39.8 % electrostatic contributions. The shape based alignment III yielded **(Table 4.2.2)** cross-validated r^2 0.322 with five components, non cross-validated r^2 0.759, F value 35.670 and bootstrapped r^2 0.806. The steric and electrostatic contributions were observed to be 68.3 % and 31.7 %, respectively.

Table 4.2.2. Summary of CoMFA results (analysis A)

Statistical parameters	**Alignments**			
	I [a]	II [b]	III [c]	IV [d]
r^2_{cv} [e]	0.416	0.389	0.322	0.429
Nc [f]	4	4	5	5
SEP [g]	0.493	0.543	0.540	0.470
r^2_{ncv} [h]	0.809	0.846	0.759	0.822
SEE [i]	0.284	0.273	0.292	0.262
F value	18.850	43.806	35.670	28.664
P $r^2 = 0$	0.0	0.0	0.0	0.0
Contrib. steric	70.1	60.2	68.3	64.3
Contrib. elect	29.9	39.8	31.7	35.7
r^2_{bs} [j]	0.909	0.899	0.806	0.890
SD [k]	0.034	0.044	0.060	0.039

[a]Alignment by RMSD database; [b]Alignment by atom-based RMS fit; [c]Alignment by shape-based RMS fit; [d]Alignment by atom and shape-based RMS fit; [e]Cross-validated correlation coefficient; [f]Number of components; [g]Standard error of prediction; [h]Non cross-validated square of correlation coefficient; [i] Standard error of estimate; [j] bootstrapped correlation coefficient obtained from 100 bootstrapping runs; [k]Standard deviation; P $r^2 = 0$: Probability of the value of correlation coefficient = 0

The alignment IV (atom and shape based) yielded a cross-validated r^2 of 0.429 with five components, non cross-validated r^2 0.822, F value 28.664 and bootstrapped r^2 0.890. The steric and electrostatic contributions were 64.3 % and 35.7 %, respectively.

Since, CoMFA technique is alignment sensitive, differences in statistical values are observed with different alignments **(Table 4.2.2)**. All the CoMFA models obtained from analysis **'A'** demonstrated moderate internal and external predictivity. Thus, in order to increase the predictive power of derived models, further experiments were performed. Based on the results of PLS analysis three molecules (**3, 27** and **45**) with high residual values were omitted. Analysis **'B'** (after removing the outliers) **(Table 4.2.3)** shows the CoMFA results obtained from the said four different alignment rules.

In the new CoMFA analysis (analysis B) RMSD based database alignment I showed improved cross-validated r^2 value from 0.416 to 0.502 with five components, non cross-validated r^2 from 0.809 to 0.922, F value from 18.850 to 34.292 and bootstrapped r^2 from 0.909 to 0.969. The steric and electrostatic contributions were found to be 58.2 % and 41.8 %, respectively.

Table 4.2.3 Summary of CoMFA results (analysis B)

Statistical parameters	**Alignments**			
	I [a]	II [b]	III [c]	**IV [d]**
r^2_{cv}	0.502	0.559	0.652	**0.664**
Nc	5	5	6	**6**
SEP	0.430	0.405	0.361	**0.360**
r^2_{ncv}	0.922	0.893	0.914	**0.916**
SEE	0.183	0.203	0.179	**0.180**
F value	34.292	40.416	48.127	**47.341**
P $r^2 = 0$	0.0	0.0	0.0	**0.0**
Contrib. steric	58.2	59.4	59.9	**58.1**
Contrib. elect.	41.8	40.6	40.1	**41.9**
r^2_{bs}	0.969	0.913	0.946	**0.947**
SD	0.017	0.034	0.025	**0.029**

[a]Alignment by RMSD database; [b]Alignment by atom-based RMS fit; [c]Alignment by shape-based RMS fit; [d]Alignment by atom and shape-based RMS fit

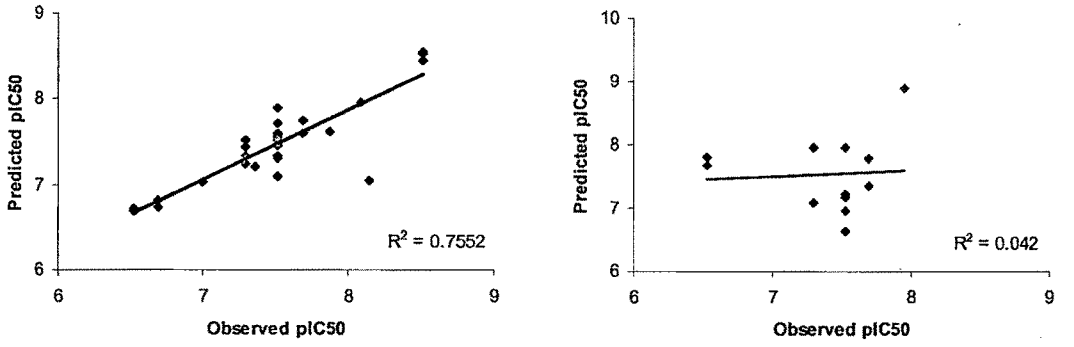
The same trend was observed in rest of the alignments (**Table 4.2.3**) except for atom-based alignment II, where the F value slightly decreased from 43.806 to 40.416. This model displayed an improved cross-validated r^2 of 0.389 to 0.559 with five components, non cross-validated r^2 from 0.846 to 0.893 and bootstrapped r^2 from 0.899 to 0.913 with 59.4 % and 40.6 % steric and electrostatic contributions, respectively.

The shape based alignment III (**Table 4.2.3**) yielded cross-validated r^2 0.652 with six components, non cross-validated r^2 of 0.914, r^2 value 48.127 and bootstrapped r^2 of 0.946. The steric and electrostatic contributions were 59.9 % and 40.1 %, respectively.

The alignment IV (**Table 4.2.3**) yielded highest cross-validated r^2 of 0.664 with six components, non cross-validated r^2 of 0.916, F value 47.347 and bootstrapped r^2 of 0.947. The steric and electrostatic contributions were found to be 58.1 % and 41.9 %, respectively.

Thus, all the CoMFA models derived from analysis B showed higher contributions of steric parameters towards the activity of these compounds similar to that of analysis A. Based on the predictive ability of the four CoMFA models from analysis B, alignment IV was further selected for further analysis and all the CoMFA contours were generated using this model. The graphs of actual *versus* fitted/predicted activities for training and test set of molecules is depicted in **Figure 4.2.3.**

The field values generated at each grid point were calculated as the scalar product of the associated QSAR coefficient and the standard deviation of all values in the corresponding column of data table (STDDEV*COEFF), plotted as the percentage contributions to QSAR equation. The CoMFA steric and electrostatic contour maps developed using the alignment IV analyses are shown in **Figures 4.2.4a and 4.2.4b,** respectively.



(a) (b)

Figure 4.2.3 A graph of actual *versus* calculated activities of training (a) and test (b) set molecules from alignment IV (analysis B)

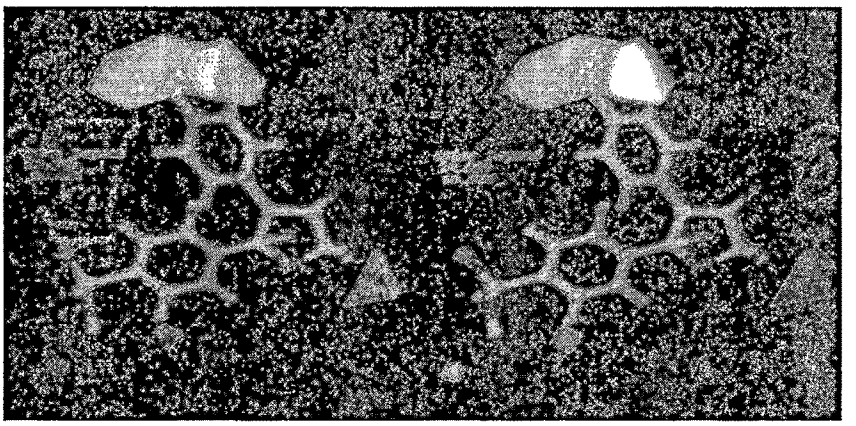


Figure 4.2.4a Stereoview of CoMFA steric STDDEV*COEFF contour plots from alignment IV (analysis B). Sterically favored areas (contribution level 80 %) are represented by contours at the top of the molecule as shown above. Sterically disfavored areas (contribution level 20 %) are represented by polyhedra in the vicinity of five membered ring. The active compound **26** is shown in ball and stick model.

The CoMFA steric map (**Figure 4.2.4a**) encompasses sterically unfavorable contours (80% contribution) corresponding to regions in space where steric bulk envisages the decrease in activity and the polyhedron bordering the furanone ring suggests that bulkier substituents are not favorable in that region. Conversely, the sterically favorable regions (20% contribution) observed at the upper side of 4-phenyl ring in the

vicinity of 5′-F of compound **26** reveals that, an increase in activity is anticipated due to increased steric bulk.

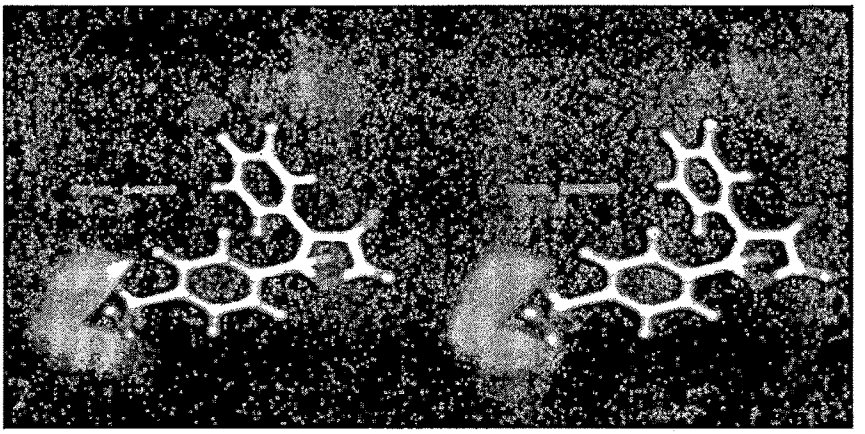


Figure 4.2.4b Stereoview of CoMFA electrostatic STDDEV*COEFF contour plots from alignment IV (analysis B). Positive charged favored areas (contribution level 80 %) are represented by polyhedra (left corner). Negatively charged favored areas (contribution level 20 %) are represented by polyhedra at the top. The active compound 26 is shown in ball and stick model.

Steric substituents at 3′ and 4′ position favor the activity (compound **11**). In this compound the structural flexibility of the n-butyl moiety seems to allow itself to maneuver in to the ideal space for COX-2 activity and selectivity. While sterically unfavorable contours surrounding 2-position of furanone nucleus suggested that high steric bulk reduces the activity (compounds **48, 49**). The compounds containing spirocyclopentyl moiety (compounds **44-47**) at 2-position in furanone nucleus showed better activity as compared to compounds **21-24**. This may be because spirocyclopentyl moiety occupies smaller space as compared to diethyl moiety in compounds **21-24**. Upon the basis of reported structure of COX-II enzyme, there should be a cavity of limited size available to the 2,2-dialkyl moiety of 3(2*H*)-furanone COX-2 inhibitors. Spirocyclohexyl moiety containing compounds (**48** and **49**) show even poor activity, which supports the rationale for "the cavity of limited size near the entrance of COX-2 active site".

The CoMFA electrostatic map (**Figure 4.2.4b**), displays contours in the vicinity of 4'-H and 5'-H where the partial negative charge is associated with increased activity (80 % contribution). Contours (20 % contribution) observed in the vicinity of sulfonamide group of 5-phenyl ring of compound **26** indicates areas where the electropositive properties of molecules describe an increase in activity. The electronegative group favoring contours (**Figure 4.2.4b**) in the vicinity of 3'-H and 4'-H suggests that increased activity is anticipated by electronegative substituents at position 3' and 4' (compounds **26** and **40**), whereas electropositive group favoring contour indicates low electron density substituents at 4''-position favoring the activity.

In addition to the steric and electrostatic fields, CoMSIA also defines the lipophilicity, hydrogen bond donor and acceptor fields that are not generally accessible with standard CoMFA. The atom and shape based alignment IV used in CoMFA studies served as an alignment for CoMSIA and the results of the study are summarized in **Table 4.2.4**.

The CoMSIA model (Model 9) with the combination of all fields yielded a cross-validated r^2 0.739 with five components, non cross-validated r^2 0.918, F value 60.145 and bootstrapped r^2 0.945. The contributions of steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor fields of this model were 8.2 %, 8.1 %, 25.1 %, 30.9 % and 27.7 %, respectively.

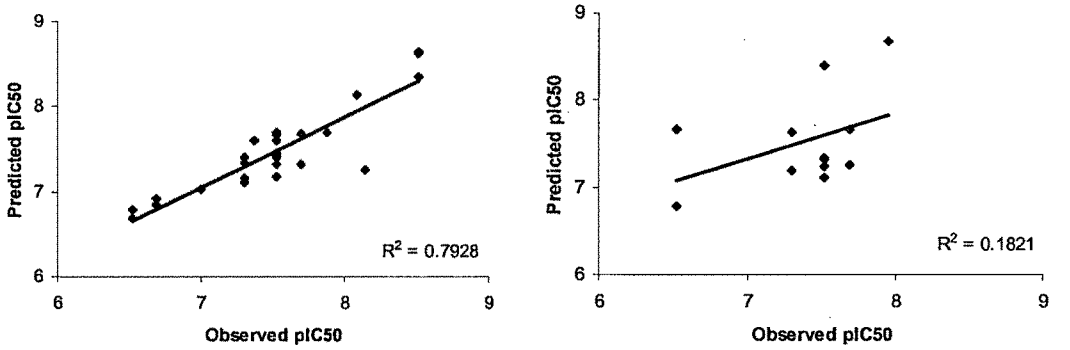
Model 6 with the combination of steric, electrostatic, hydrophobic and hydrogen bond donor fields yielded a highest cross-validated r^2 0.777 with four components, non cross-validated r^2 0.905, F value 66.322 and bootstrapped r^2 of 0.933. The steric, electrostatic, hydrophobic and hydrogen bond donor field contributions were found to be 12.6 %, 11.0 %, 34.4 % and 41.9 % , respectively. The combinations of steric, electrostatic and hydrogen bond donor fields yielded a CoMSIA model (Model 1) with cross-validated r^2 0.722 with four components, non cross-validated r^2 0.852, F value 40.280 and

bootstrapped r^2 0.896. The steric, electrostatic and hydrogen bond donor contributions were found to be 31.3 %, 21.9% and 46.9 %, respectively.

Table 4.2.4 Summary of CoMSIA results

Statistical parameters	**Models**								
	1	2	3	4	5	**6**	7	8	9
	S+E+ H	S+E+ D	S+E+ A	S+E+D+ A	H+D+ A	**S+E+ D+H**	S+E+ H+A	S+D+ A	All
r^2_{cv}	0.635	0.722	0.507	0.705	0.648	**0.777**	0.641	0.710	0.739
Nc	3	4	7	3	4	**4**	6	5	5
SEP	0.355	0.316	0.445	0.319	0.355	**0.282**	0.372	0.328	0.311
r^2_{ncv}	0.844	0.852	0.890	0.831	0.901	**0.905**	0.902	0.819	0.918
SEE	0.232	0.230	0.210	0.242	0.188	**0.185**	0.194	0.259	0.175
F value	52.48	40.280	28.766	47.548	63.65	**66.322**	40.092	24.47	60.145
Pr2=0	0.0	0.0	0.0	0.0	0.0	**0.0**	0.0	0.0	0.0
r^2_{bs}	0.880	0.896	0.944	0.848	0.925	**0.933**	0.946	0.882	0.945
S D	0.044	0.039	0.031	0.055	0.030	**0.028**	0.036	0.055	0.026

S=Steric; E=Electrostatic; H=Hydrophobic; D=Hydrogen bond donor; A=Hydrogen bond acceptor; All = S+E+D+A+H.



(a) (b)

Figure 4.2.5 A graph of actual *versus* predicted activities of training set (a) and test set (b) molecules from CoMSIA analysis using Model 6

The models generated by various combinations of CoMSIA fields (**Table 4.2.4**) show statistically significant moderate to high internal and external predictions in which the

hydrogen bond donor fields were observed to be predominant than hydrophobic and hydrogen bond acceptor fields. The graphs of actual *versus* predicted activities for the training and test set molecules from Model 6 are shown in **Figure 4.2.5.**

The CoMSIA steric and electrostatic contours (not shown) were placed almost similar to that of the CoMFA model. The hydrogen bond donor contour maps of CoMSIA (STDDEV*COEFF) are displayed in **Figure 4.2.6.**

The contour region (80 % contribution) near 3''- H of 5- phenyl ring and also in the vicinity of 3'-F and 4'-H of 4-phenyl ring indicate the disfavored regions for hydrogen bond donor substituents. While the contour region at the lower region of 5-phenyl ring, in the vicinity of 5''-H of compound **26** indicates favored regions for hydrogen bond donor substituents.

The hydrophobic contour maps of CoMSIA (STDDEV*COEFF) are displayed in **Figure 4.2.7.** The contours covering the five membered furanone ring indicate disfavored regions, while the contours in the vicinity of 3'-F, 4'-H and 5'-H of 4-phenyl ring of compound **26** indicate favored regions.

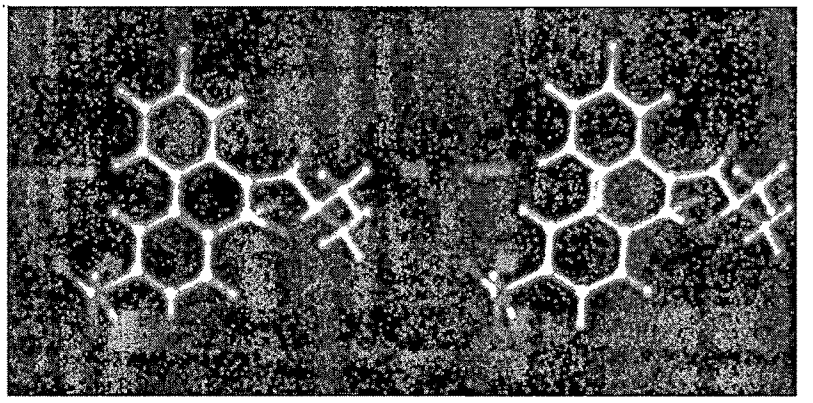


Figure 4.2.6 Stereoview of CoMSIA hydrogen bond donor fields.

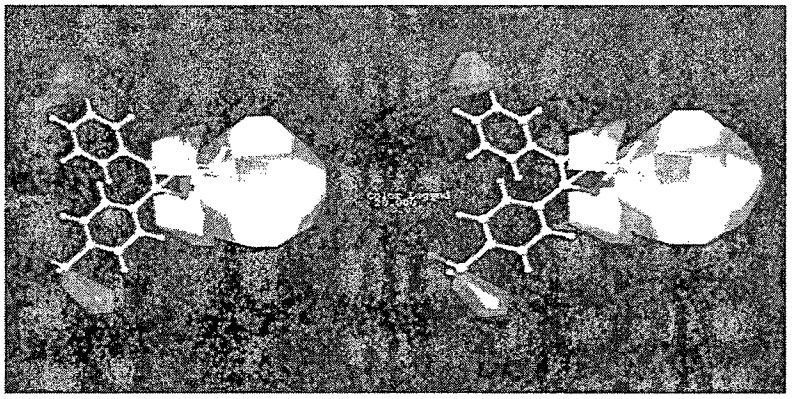


Figure 4.2.7 Stereoview of CoMSIA hydrophobic fields.

The CoMSIA QSAR studies give an additional structural insight in the study of the probable binding sites of the ligand-receptor. The hydrogen bond donor favorable contour in the vicinity of the amino group of sulfonamide moiety indicates significance of proton in the hydrogen bond formation of receptor surface. The disfavorable contours in the upper region of 4-phenyl ring indicate that hydrogen bond donor substituents on this ring decrease activity (compound **13** and **23**). The hydrophobic field favoring contours (**Figure 4.2.7**) in the upper region of 4-phenyl ring highlight the significance of hydrophobic groups on 3'-H, 4'-H and 5'-H of 4-phenyl ring in the biological activity. While the disfavoring contour appears surrounding the furanone ring, indicating that hydrophobic substitutions on this five membered ring decrease activity (compound **45-49**).

Table 4.2.5 shows comparison of the statistical parameters derived from the best CoMFA and CoMSIA Model 6. The cross-validated r^2 is on higher side (0.777) for CoMSIA model as compared to 0.664 for CoMFA. The standard error of prediction (0.282) is also low for the CoMSIA model. Results from CoMSIA suggest that apart from steric and electrostatic parameters, hydrogen bond donor and hydrophobic parameters also influence the COX-II inhibitory activity of furanone derivatives.

As CoMFA is alignment sensitive, four different alignments were tried in which alignment IV (atom and shape-based) yielded good results. Initially, analysis 'A' was performed using 39 molecules yielded poor results. Compounds which were interfering the analysis were removed as outliers and analysis 'B' was carried out using 36 molecules which showed improved statistics.

Table 4.2.5 Statistical comparison of the best models obtained by CoMFA and CoMSIA

Technique	r^2_{cv}	SEP	N_c	r^2_{ncv}	SEE	r^2_{bs}	S.D
CoMFA	0.664	0.360	6	0.916	0.180	0.947	0.029
CoMSIA	0.777	0.282	4	0.905	0.185	0.933	0.028

In brief, CoMSIA Model 6 derived using steric, electrostatic, hydrophobic and hydrogen bond donor fields can be used to design better COX-II inhibitors belonging to furanone class. It can be concluded from the study that apart from steric and electrostatic parameters, hydrogen bond donor and hydrophobic parameters also influence the COX-II inhibitory activity.

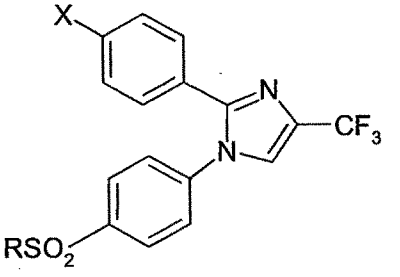
4.3 Integrating Ligand and Target-Based Information to Assess the Influence of Calculated Enzyme-Inhibitor Binding Affinities and Partial Charges on the Statistical Quality and Predictive Ability of 3D-QSAR CoMFA models: A Case Study with Selective COX-2 Inhibitors

4.3.1 Experimental

4.3.1.1 Data sets and alignment rule

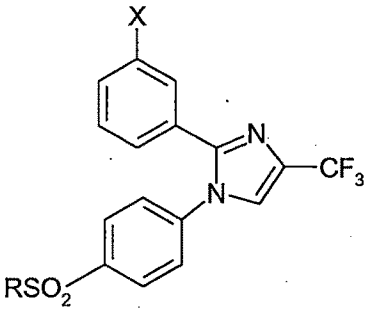
A data set of 66 compounds, whose structures and biological activities are given in **Table 4.3.1**, has been taken from literature[1]. The compounds were divided into training and test sets based on variations in the structures and biological activities. COMFA models were developed using 57 training set compounds and the predictive ability of the models were assessed using an external set of 9 compounds.

Table 4.3.1 Comparison of experimentally observed and CoMFA-predicted activities (pIC_{50}) for COX-2 inhibitors.

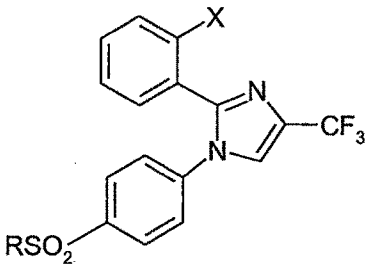


C. N	X	R	Calculated Binding Energy (kcal/mol)	pIC_{50}		Residual
				Act	Pred[a]	
1	H	Me	-88.5	6.92	6.82	0.10
2	Me	Me	-87.9	6.79	6.73	0.06
3*	OMe	Me	-95.8	6.24	6.44	-0.20
4	NHMe	Me	-99.2	5.83	5.98	-0.15
5	NMe_2	Me	-91.8	6.15	6.17	-0.02
6	SMe	Me	-94.4	6.79	6.70	0.09

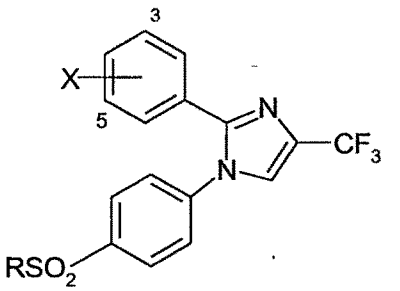
7	SO_2Me	Me	-85.5	5.24	5.26	-0.02
8	Cl	NH_2	-97.4	8.00	7.92	0.08
9	H	NH_2	-95.4	8.00	7.75	0.25
10*	F	NH_2	-93.7	7.39	7.65	-0.26
11	Me	NH_2	-93.5	7.39	7.57	-0.18



C. N	X	R	Calculated Binding Energy (kcal/mol)	pIC50		Residual
				Act.	Pred[a]	
12	Cl	Me	-93.7	7.22	7.08	0.14
13	F	Me	-89.8	6.92	6.89	0.03
14	Br	Me	-93.5	7.09	6.95	0.14
15	CF_3	Me	-94.5	6.67	6.65	0.02
16	OMe	Me	-92.4	6.45	6.29	0.16
17	SMe	Me	-94.8	6.45	6.39	0.06
18	CH_2OMe	Me	-100.3	4.16	4.07	0.09
19*	NMe_2	Me	-91.0	5.49	5.29	0.20
20	NHMe	Me	-89.4	6.03	6.16	-0.13
21	NO_2	Me	-103.1	6.23	6.55	-0.32
22	Cl	NH_2	-98.5	8.09	7.94	0.15
23	F	NH_2	-95.0	7.52	7.77	-0.25
24	Br	NH_2	-98.0	8.15	7.79	0.36
25*	Me	NH_2	-95.5	7.52	7.55	-0.03

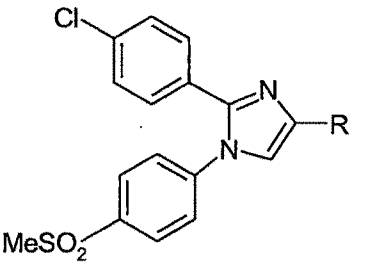


C. N	X	R	Calculated Binding Energy (kcal/mol)	pIC50		Residual
				Act.	Pred[a]	
26	Cl	Me	-92.9	6.04	6.16	-0.12
27	F	Me	-90.0	6.39	6.38	0.01
28	Me	Me	-89.4	6.09	5.84	0.15
29	OMe	Me	-91.7	4.00	3.82	0.18
30	F	NH_2	-85.3	7.00	7.17	-0.17
31*	Me	NH_2	-94.1	6.69	6.62	0.07



C. N	X	R	Calculated Binding Energy (kcal/mol)	pIC50		Residual
				Act.	Pred[a]	
32	4-NMe_2,3-Cl	Me	-84.6	6.49	6.44	0.05
33	4-NMe_2,3-F	Me	-91.2	6.48	6.34	0.14
34	4-NHMe,3-Cl	Me	-99.8	6.18	6.25	-0.07
35*	4-Me,3-Cl	Me	-94.8	7.52	7.19	0.33
36	4-Me,3-F	Me	-92.1	6.95	7.04	-0.09
37	3-Me,4-F	Me	-91.5	6.76	6.86	-0.10
38	3-Me,4-Cl	Me	-90.9	7.04	7.12	-0.08
39	3-OMe,4-Cl	Me	-96.4	6.60	6.73	-0.13
40*	3,4-OCH_2O-	Me	-98.3	6.76	7.10	-0.34
41	3,4-diF	Me	-92.9	6.92	6.80	0.12
42	3,4-diMe	Me	-87.9	6.48	6.83	-0.35
43	3-Me,5-Cl	Me	-88.2	7.09	6.80	0.29
44	3-Me,5-F	Me	-90.9	6.95	6.63	0.32

45	3-OMe,5-F	Me	-96.1	6.01	5.88	0.13
46	3,5-diCl	Me	-89.4	6.76	7.20	-0.44
47	4-OMe,3-F	Me	-101.1	7.52	7.51	0.01
48	4-SMe,3-Cl	Me	-102.8	8.00	7.96	0.04
49	4-Me,3-Cl	Me	-100.3	8.52	8.04	0.48
50	3-OMe,4-Cl	Me	-102.4	7.69	7.56	0.13
51	3,4-diF	Me	-98.3	7.52	7.66	-0.14
52*	3-Me,5-Cl	Me	-94.2	7.39	7.64	-0.25
53	3-OMe,5-F	Me	-100.5	6.33	6.74	-0.41
54	4-OMe,3,5-diF	NH_2	-95.8	6.76	6.51	0.25
55	4-OMe,3,5-diCl	NH_2	-94.9	6.85	6.99	-0.14
56	4-OMe,3,5-diBr	NH_2	-92.3	7.04	6.82	0.22
57*	4-OMe,3,5-diMe	NH_2	-91.3	6.14	6.16	-0.02
58	4-OMe,2,5-diMe	NH_2	-89.7	4.91	5.20	-0.29

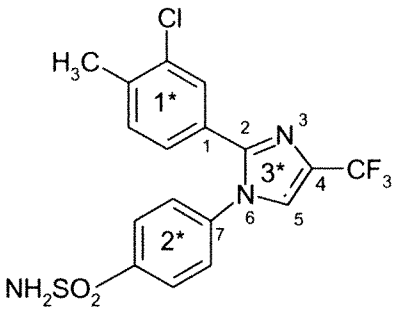


C. N	R	Calculated Binding Energy (kcal/mol)	pIC_{50}		Residual
			Act.	Pred[a]	
59	CHF_2	-89.2	6.21	6.18	0.03
60	CH_2F	-88.5	6.38	6.35	0.03
61	CHO	-95.8	5.79	5.55	0.04
62	CN	-90.7	6.63	6.79	-0.16
63	CO_2Et	-91.7	5.24	5.27	-0.03
64	CH_2OH	-97.6	5.07	5.22	-0.15
65	CH_2SMe	-79.2	6.49	6.50	-0.01
66	CH_2CN	-95.4	5.81	5.77	0.04

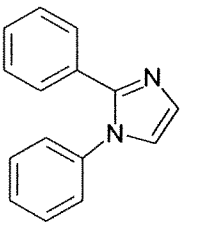
* test set compounds; [a] predicted using model $AI1_{iv}$

Alignment rules

Atom-based, atom-shape-based and database alignment approaches (**Figure 4.3.1**) have been employed in order to calculate the CoMFA interaction energies. The most active molecule **49** was used as a template and rest of the molecules were aligned to it using the below mentioned alignment rules.



Template molecule (**49**)



Fragment for database alignment

Alignment rule	atoms/centroids
AI	1 2 3 4 5 6 7
AII	1* 2* 1 2 6 7
AIII	Database alignment

Figure 4.3.1 Alignment rules

4.3.2 Results and discussion

4.3.2.1 CoMFA Models

To explore the effect of different variables on the predictive ability of the CoMFA model, separate CoMFA models were constructed that varied in terms of their

alignment scheme (atom-based, shape-based and database alignment) (**Figure 4.3.1**) and partial charge assignment schemes (Gasteiger-Marsili, Gasteiger-Huckel, Huckel and MMFF94). Results from various CoMFA models are reported in **Table 4.3.2.**

Initially, CoMFA model was developed by employing MMFF94 charge formalism. The model was validated for its statistical quality and predictive ability, not only by an external set of 9 compounds, but also by predicting the activity of currently available selective COX-2 inhibitors (celecoxib, valdecoxib, etoricoxib and rofecoxib). The activity (pIC_{50}) of these drugs predicted by the CoMFA model (**Table 4.3.3**) were in excellent correlation (r^2 = 0.874) with the experimentally determined pIC_{50} values as depicted in **Figure 4.3.4** confirming its (the developed model) utility in the designing of selective COX-2 inhibitors. The training (r^2 = 0.960) and test set (r^2 = 0.920) compounds predicted, too were well in accordance with the *in vitro* data. Standard coxibs were also studied to assess and confirm the influence of partial charges on the predictive ability of the CoMFA models (**Table 4.3.2**).

Proper prediction of the 3D-QSAR model depends on the method by which partial atomic charges are calculated in a particular data set. Therefore, to optimize the partial charges to be used, the inhibitors were assigned partial charges using four different schemes: Gasteiger-Marsili (i), Gasteiger-Huckel (ii), Huckel (iii) and MMFF94 (iv) charges. In all the four cases atom-based, shape-based and database alignment schemes were employed. To examine the influence of including the calculated enzyme inhibitor binding energies on the statistical quality of these CoMFA models, (models $AI1_{i\text{-}iv}$, $AII1_{i\text{-}iv}$ and $AIII1_{i\text{-}iv}$) were reconstructed after adding the binding energy as an additional independent descriptor to yield models $AI2_{i\text{-}iv}$, $AII2_{i\text{-}iv}$ and $AIII2_{i\text{-}iv}$, respectively.

Table 4.3.2 Influence of molecular alignment, partial atomic charge scheme and inclusion of calculated binding energy on the statistical quality and predictive ability of CoMFA models

Statistics	Charge formalisms	Models					
		AI		AII		AIII	
		1	2	1	2	1	2
r^2_{ncv}	GM (i)	0.925	0.953	0.901	0.933	0.922	0.952
q^2		0.422	0.426	0.348	0.352	0.415	0.427
r^2_{pred}		0.812	0.846	0.726	0.808	0.825	0.850
r^2_{ncv}	GH (ii)	0.939	0.945	0.922	0.933	0.922	0.915
q^2		0.358	0.413	0.376	0.404	0.388	0.373
r^2_{pred}		0.855	0.852	0.831	0.851	0.847	0.748
r^2_{ncv}	H (iii)	0.907	0.914	0.929	0.934	0.947	0.957
q^2		0.530	0.511	0.502	0.475	0.542	0.507
r^2_{pred}		0.785	0.818	0.839	0.850	0.867	0.808
r^2_{ncv}	MMFF94 (iv)	**0.955**	0.954	0.944	0.912	0.955	0.935
q^2		**0.543**	0.558	0.483	0.595	0.533	0.565
r^2_{pred}		**0.929**	0.907	0.891	0.835	0.785	0.790

AI: atom-based rms alignment; AII: atom and shape-based alignment; AIII: database alignment; 1 - CoMFA without binding energy; 2 - CoMFA with binding energy; CoMFA models developed using GM: Gasteiger-Marsili (i); GH: Gasteiger-Huckel (ii); H: Huckel (iii) and MMFF94 (iv) charge formalisms

Table 4.3.3 Influence of partial charge formalisms on the predicted pIC_{50} values of standard coxibs

Compounds	pIC_{50}				
	Actual	Predicted (Models AII_{i-iv})			
		Gasteiger Marsili (i)	Gasteiger Huckel (ii)	Huckel (iii)	MMFF94 (iv)
Celecoxib	6.05	**5.51**	5.18	4.90	**5.87**
Rofecoxib	6.27	5.97	**6.18**	**6.41**	4.96
Etoricoxib	6.16	6.45	6.51	5.05	5.47
Valdecoxib	6.57	**5.01**	5.33	4.93	**5.64**
Template(**49**)	8.52	**8.04**	9.80	7.98	8.00

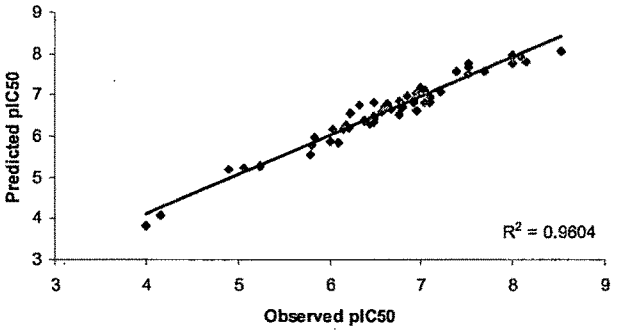


Figure. 4.3.2 Plot of experimentally observed *vs* fitted values of pIC_{50} from best CoMFA model $AI1_{iv}$ for training set compounds

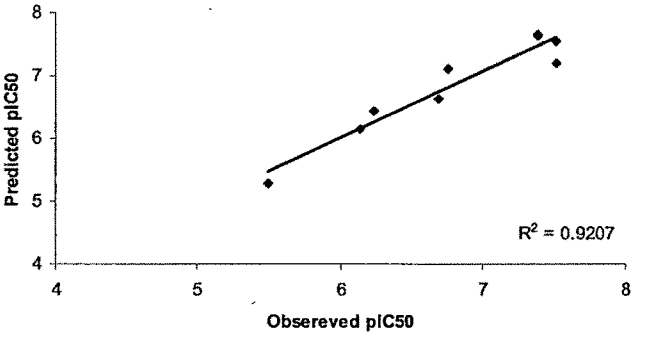


Figure 4.3.3 Plot of experimentally observed *vs* fitted values of pIC_{50} from best CoMFA model $AI1_{iv}$ for test set compounds

Considering the binding energy as an important descriptor involving structure-based data, its contribution to the resulting model may be substantial. Preparation of the COX-2 enzyme and docking methodology are mentioned in previous chapter. This is not surprising from a physical point of view in the sense that the binding energy would be expected to capture the true nature of the enzyme-inhibitor interactions. On the other hand, the steric-electrostatic fields from CoMFA are hypothetical in the sense that they are derived from the ligands themselves and not from any explicit information about the enzyme.

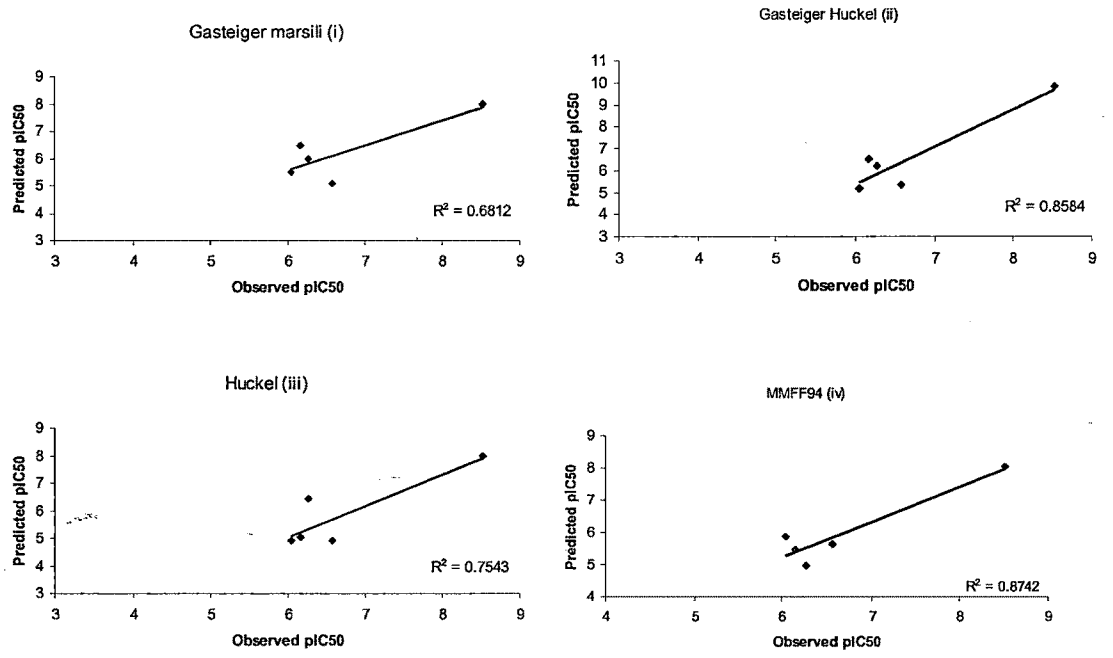
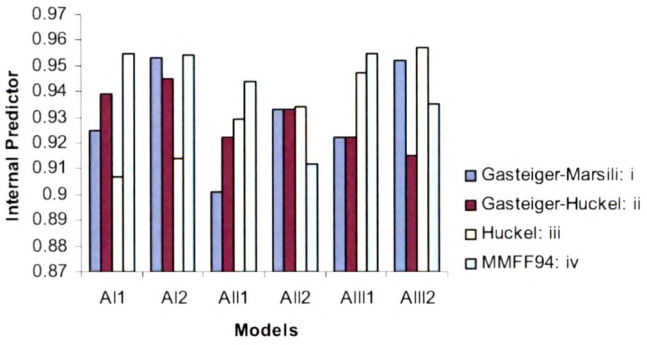


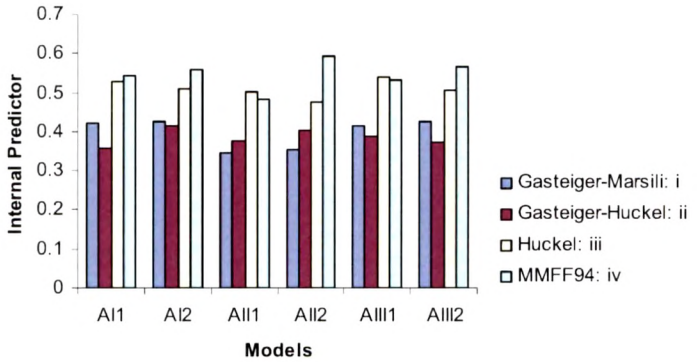
Figure 4.3.4 Correlation between actual and predicted activities by CoMFA models $AI1_{i-iv}$ employing different charge formalisms for standard coxibs and template

All the 24 models developed exhibited good statistical quality between the predicted and experimentally determined values of pIC_{50}: models $[(AI\text{-}III)1\text{-}2]_i$ (r^2 = 0.901-0.953, q^2 =0.348-0.427, r^2_{pred} = 0.726-0.850), $[(AI\text{-}III)1\text{-}2]_{ii}$ (r^2 = 0.922-0.945, q^2 = 0.358-0.414, r^2_{pred} = 0.748-0.855), $[(AI\text{-}III)1\text{-}2]_{iii}$ (r^2 = 0.907-0.957, q^2 = 0.475-0.542, r^2_{pred} = 0.785-0.867), $[(AI\text{-}III)1\text{-}2]_{iv}$ (r^2 = 0.912-0.955, q^2 = 0.483-0.595, r^2_{pred} = 0.785-0.929). A visual comparison of the results from these four partial-charge assignment schemes is given in **Figure 4.3.5**.

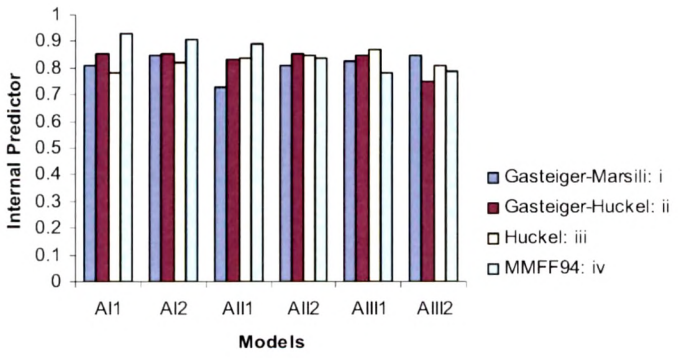
Among the parameter variations considered, the combination of MMFF94 charges (iv) and atom-based alignment (AI) without binding energy (1) (**Figure 4.3.2**) yielded the best CoMFA model ($AI1_{iv}$). The training and test set predictions using this model are depicted in **Table 4.3.1.**



(a) Comparison of non-cross-validated r^2 (r^2_{ncv}) values



(b) Comparison of cross-validated q^2 values



(c) Comparison of predictive r^2 (r^2_{pred}) values

Figure 4.3.5 Visual comparison of (a) r^2_{ncv}, (b) q^2 and (c) r^2_{pred} values for CoMFA models in which different partial-charge formalisms are applied.

Inclusion of binding energy (2) as an additional descriptor did influence the predictive power of CoMFA models but not to a great extent and varied according to the partial charges employed. Further, a comparative study was performed to assess the influence of different partial charge formalisms on the predictive ability of the developed CoMFA models (**Table 4.3.4 and 4.3.5**).

Table 4.3.4 Comparison of the predicted activities of training set of compounds employing different partial charge formalisms using alignment rule AI

C.N	pIC50				
	Actual	Predicted with Models $AI1_{i-iv}$			
		Gasteiger Marsili (i)	Gasteiger Huckel (ii)	Huckel (iii)	MMFF94 (iv)
1	6.92	6.93	6.85	6.81	6.82
2	6.79	6.74	6.75	6.73	6.73
4	5.83	5.86	6.09	5.75	5.98
5	6.15	6.22	6.28	6.06	6.17
6	6.79	6.80	6.93	6.80	6.7
7	5.24	5.30	5.19	5.13	5.26
8	8.00	7.90	7.95	8.04	7.92
9	8.00	7.77	7.76	7.87	7.75
11	7.39	7.63	7.68	7.75	7.57
12	7.22	6.98	7.02	7.07	7.08
13	6.92	6.81	6.70	6.85	6.89
14	7.09	6.88	6.91	6.98	6.95
15	6.67	6.66	6.57	6.87	6.65
16	6.45	6.40	6.46	6.48	6.29
17	6.45	6.35	6.52	6.36	6.39
18	4.16	4.16	4.28	4.08	4.07
20	6.03	6.06	5.95	6.02	6.16
21	6.23	6.53	6.72	6.58	6.55
22	8.09	7.91	7.99	7.95	7.94
23	7.52	7.72	7.64	7.70	7.77
24	8.15	7.84	7.90	7.83	7.79
26	6.04	6.01	6.17	5.89	6.16
27	6.39	6.13	6.36	6.37	6.38
28	6.09	5.83	5.86	5.67	5.84
29	4.00	3.92	4.00	4.13	3.82

30	7.00	7.05	7.37	7.27	7.17
32	6.49	6.43	6.31	6.56	6.44
33	6.48	6.24	6.23	6.33	6.34
34	6.18	6.29	6.41	6.34	6.25
36	6.95	6.25	6.41	6.92	7.04
37	6.76	6.99	6.91	6.61	6.86
38	7.04	6.79	6.80	7.04	7.12
39	6.60	7.04	7.02	6.50	6.73
41	6.92	6.40	6.52	6.87	6.8
42	6.48	6.83	6.69	6.74	6.83
43	7.09	6.83	6.78	6.89	6.8
44	6.95	6.95	6.83	6.68	6.63
45	6.58	6.69	6.57	6.58	6.6
46	6.01	5.92	5.96	5.99	5.88
47	7.52	7.62	7.75	7.48	7.51
48	8.00	8.01	8.20	8.06	7.96
49	8.52	8.06	8.12	7.97	8.04
50	7.69	7.36	7.53	7.39	7.56
51	7.52	7.82	7.71	7.75	7.66
53	6.33	6.90	6.95	6.87	6.74
54	6.76	6.44	6.37	6.56	6.51
55	6.85	6.91	6.92	7.04	6.99
56	7.04	6.80	6.79	6.88	6.82
58	4.91	5.26	5.21	5.19	5.2
59	6.21	6.21	6.11	6.39	6.18
60	6.38	6.37	6.38	6.38	6.35
61	5.79	5.76	5.64	5.73	5.55
62	6.63	6.68	6.61	6.44	6.79
63	5.24	5.25	5.20	5.22	5.27
64	5.07	5.21	5.33	5.27	5.22
65	6.49	6.45	6.24	6.61	6.5
66	5.81	5.72	5.67	5.75	5.77

Table 4.3.5 Comparison of the predicted activities of test set of compounds employing different charges formalisms using alignment rule AI

C.N	pIC_{50}				
	Actual	Predicted with (Models $AI1_{i-iv}$)			
		Gasteiger Marsili (i)	Gasteiger Huckel (ii)	Huckel (iii)	MMFF94 (iv)
3	6.24	6.43	6.61	6.49	6.44
10	7.39	7.79	7.8	7.84	7.65
19	5.49	5.41	5.49	5.52	5.29
25	7.52	7.50	7.53	7.57	7.55
31	6.69	6.67	6.82	6.43	6.62
35	7.52	7.14	7.16	7.07	7.19
40	6.76	6.76	7.11	6.74	7.10
52	7.39	7.92	7.86	7.81	7.64
57	6.14	6.30	6.33	6.22	6.16

Least-squares-fit plots of experimental *versus* predicted activity (pIC_{50}) employing four different charge formalisms are depicted in **Figure 4.3.6**. The correlation coefficients varied from 0.932-0.960 for training set and from 0.861-0.920 for test set compounds, respectively confirming the influence of partial charges on the predictive ability of CoMFA models. The reason for varying predictive ability of models with different charge formalisms may be change in the energy and conformation of the molecule which finally affected the predictive power of the model. This is also in accordance with the fact that orientation of a compound in the active site of the target plays a vital role in determining its binding affinity and selectivity towards the target protein under study.

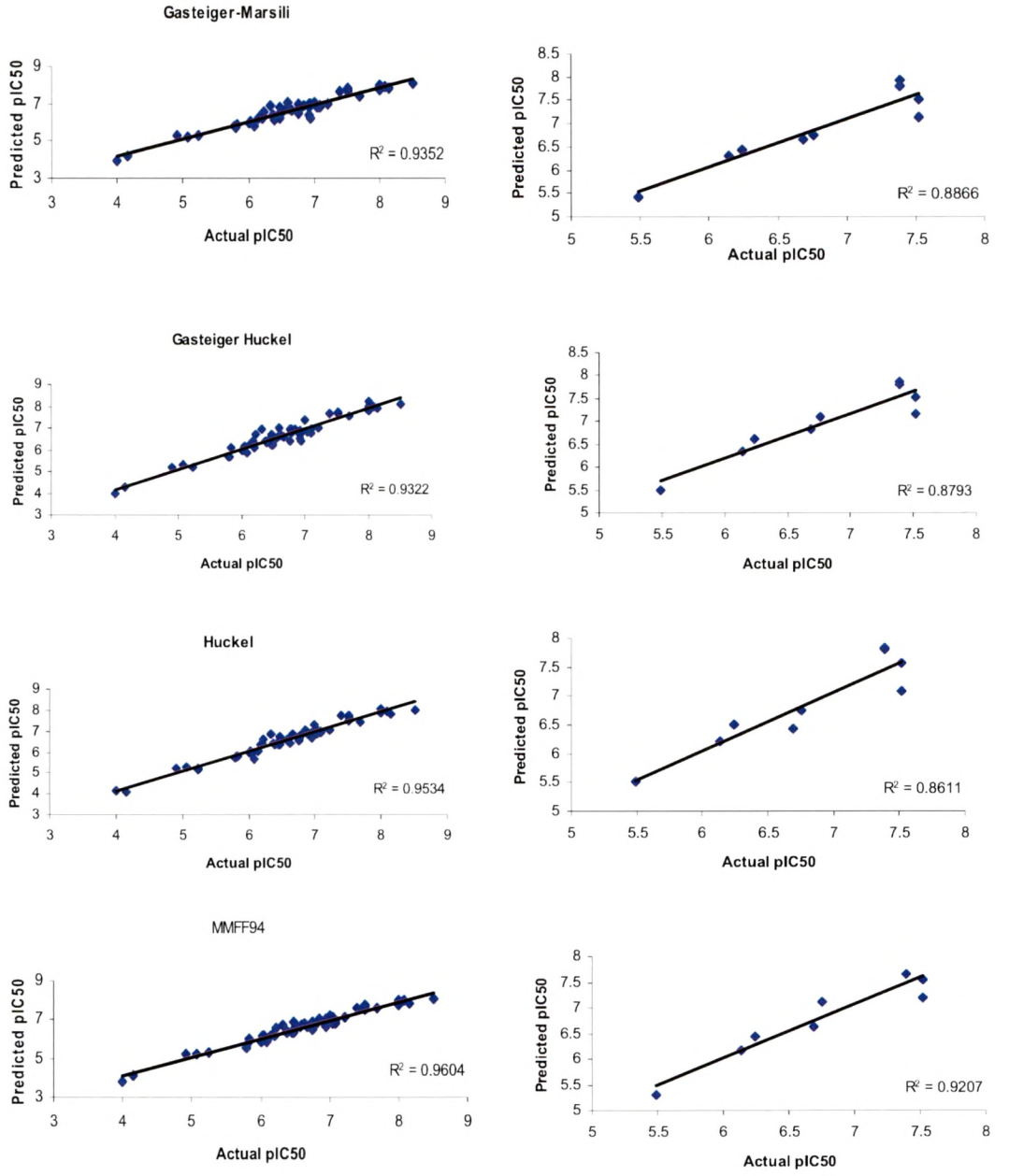


Figure 4.3.6 Correlation between experimental activities and predicted activities of 1,2-diarylimidazoles and test set compounds employing different charge formalisms by best CoMFA models $AI1_i$- $AI1_{iv}$

In this study we have identified the effect of charges on the statistical and predictive ability of the CoMFA models. MMFF94 charges were found to be effective in case of

selective COX-II inhibitors. The predictive abilities of the CoMFA models varied a lot with charges hence charge assignment was found to be one of the major factors governing the predictivity of CoMFA models. Inclusion of binding energy as an additional descriptor did not improve the CoMFA models significantly. Although, this helped us to incorporate structure-based information into the CoMFA analysis, but not with expected success.

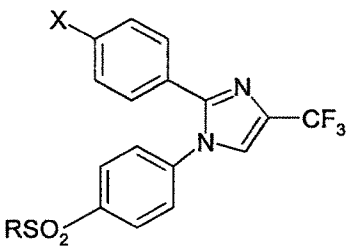
4.4 Development of universal 3D QSAR CoMFA model for COX-II inhibitors

4.4.1 Experimental

4.4.1.1 Data sets and alignment rule

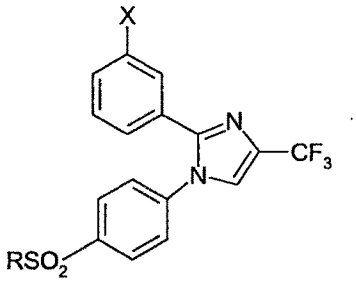
A data set of 143 diaryl COX-II inhibitors having eight different basic scaffolds (imidazole, pyrrole, pyrazole, furanone, spiroheptene, cyclopentenone, pyrazine, phenylpyran-2-one) was taken from literature[1-8] and their actual and CoMFA-predicted activity values are given in **Table 4.4.1.** Compounds with similar biological activity and less structural variations were omitted from the study to avoid redundancy. For developing CoMFA model for selective COX-II inhibition, training set and test set comprised of 123 and 20 compounds, respectively. Various alignment rules were tried, but a combination of atom and centroid-based alignment yielded best results and is only discussed. **Figure 4.4.1** displays the template structure and atoms/centroids chosen for alignment. Compound **117** was used as template to develop 3D CoMFA model for COX-II inhibitors.

Table 4.4.1 Structures and comparison of observed and CoMFA-predicted COX-II inhibitory values

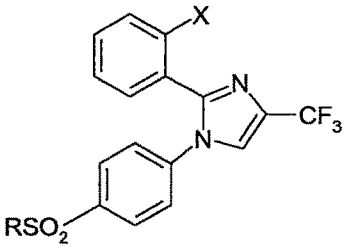


C. N	X	R	pIC_{50}	
			Actual	Predicted
1	Me	Me	6.79	6.99
2	OMe	Me	6.24	6.57
3	NMe_2	Me	6.15	6.33
4	SMe	Me	6.79	6.79

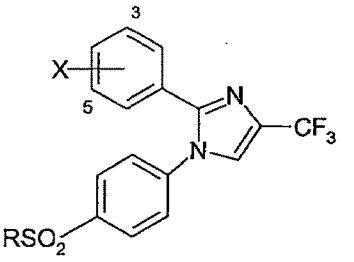
5	SO_2Me	Me	5.24	5.22
6*	H	NH_2	8.00	6.94



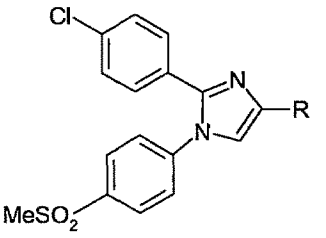
C. N	X	R	pIC_{50}	
			Actual	Predicted
7	F	Me	6.92	6.60
8	CF_3	Me	6.67	6.52
9	OMe	Me	6.45	6.53
10	SMe	Me	6.45	6.62
11	NMe_2	Me	5.49	5.68
12	NHMe	Me	6.03	6.06
13	F	NH_2	7.52	7.11



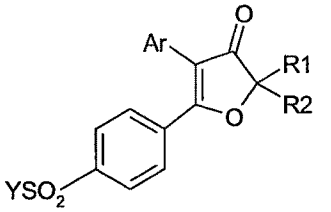
C. N	X	R	pIC_{50}	
			Actual	Predicted
14*	Cl	Me	6.04	6.44
15	F	Me	6.39	6.39
16	Me	Me	6.09	6.39
17	F	NH_2	7.00	6.84
18	Me	NH_2	6.69	6.86



C. N	X	R	pIC50	
			Actual	Predicted
19	4-NMe_2,3-Cl	Me	6.49	6.50
20	4-NMe_2,3-F	Me	6.48	6.47
21	4-Me,3-F	Me	6.95	6.98
22	3-Me,4-F	Me	6.76	6.65
23	3-Me,4-Cl	Me	7.04	6.79
24	3,4-OCH_2O-	Me	6.76	6.98
25*	3-Me,5-Cl	Me	7.09	6.92
26	3-Me,5-F	Me	6.95	6.92
27	4-OMe,3-F	Me	7.52	7.26
28	4-OMe,3,5-diF	NH_2	6.76	6.92
29	4-OMe,3,5-diCl	NH_2	6.85	7.12
30	4-OMe,3,5-diBr	NH_2	7.04	6.94

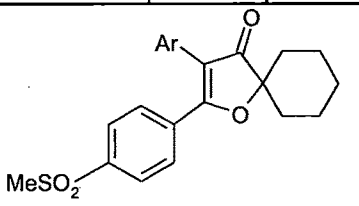


C. N	R	pIC50	
		Actual	Predicted
31	CHF_2	6.21	6.51
32*	CH_2F	6.38	6.58
33	CHO	5.79	5.95
34	CN	6.63	6.29
35	CO_2Et	5.24	4.84
36*	CH_2CN	5.81	5.97



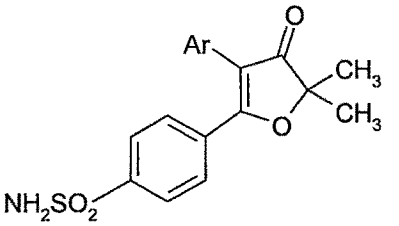
Ar = phenyl/substituted phenyl

C.N	Ar	R_1	R_2	Y	pIC_{50}	
					Actual	Predicted
37	phenyl	CH_3	CH_3	CH_3	7.30	7.21
38	3-fluorophenyl	CH_3	CH_3	CH_3	7.69	7.47
39	4-bromophenyl	CH_3	CH_3	CH_3	7.52	7.50
40*	3-(trifluoromethyl)-phenyl	CH_3	CH_3	CH_3	7.30	7.39
41	4-isopropylphenyl	CH_3	CH_3	CH_3	7.52	7.39
42	4-n-butylphenyl	CH_3	CH_3	CH_3	7.69	7.64
43	3-acetylphenyl	CH_3	CH_3	CH_3	7.30	7.47
44	2-furanyl	CH_3	CH_3	CH_3	6.69	6.70
45	3-thienyl	CH_3	CH_3	CH_3	6.69	6.59
46	4-(1-N-methyl-pyrazolyl)	CH_3	CH_3	CH_3	6.30	6.85
47*	phenyl	CH_3	C_2H_5	CH_3	6.69	6.81
48	2-fluorophenyl	CH_3	C_2H_5	CH_3	6.52	6.68
49	3,5-dichlorophenyl	CH_3	C_2H_5	CH_3	7.52	7.48
50	3-methylphenyl	C_2H_5	C_2H_5	CH_3	6.30	6.41

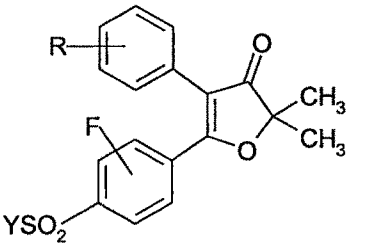


Ar - phenyl/substituted phenyl

C.N	Ar	pIC_{50}	
		Actual	Predicted
51	phenyl	5.69	5.92
52*	3-methylphenyl	6.30	6.07

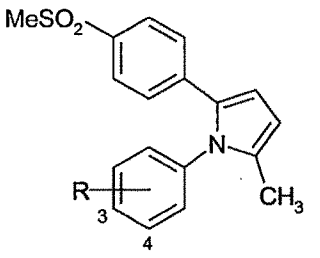


C.N	Ar	pIC50	
		Actual	Predicted
53*	phenyl	8.09	7.68
54	3-chlorophenyl	7.85	7.99
55	3,4-difluorophenyl	8.15	7.89
56	3-(trifluromethyl)-phenyl	7.95	7.82
57	4-acetylphenyl	7.88	7.89
58	2-furanyl	7.24	7.06
59	3-thienyl	6.92	6.97

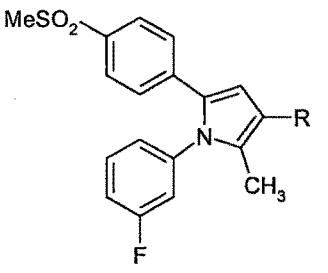


C.N	Fluoride Position	R	Y	pIC50	
				Actual	Predicted
60	2	H	CH_3	7.52	7.35
61*		3-F	CH_3	7.52	7.59
62		3-Cl	CH_3	7.69	7.64
63	3	3-F	CH_3	7.52	7.25
64		3-Cl	CH_3	7.30	7.30
65		3-F, 5-F	CH_3	7.52	7.53

66	2	3-F	NH_2	7.52	7.63
67		3-F, 5-F	NH_2	8.52	8.35
68	3	H	NH_2	7.52	7.45
69		3-F	NH_2	7.52	7.71
70		3-F, 5-F	NH_2	7.69	7.78

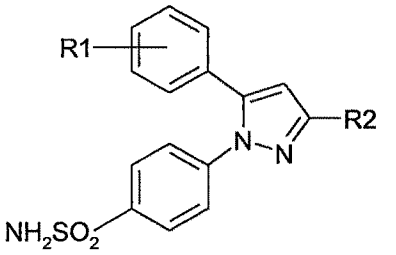


C.N	R	pIC_{50}	
		Actual	Predicted
71	4-CF_3	7.09	7.41
72	4-CH_3	7.39	7.37

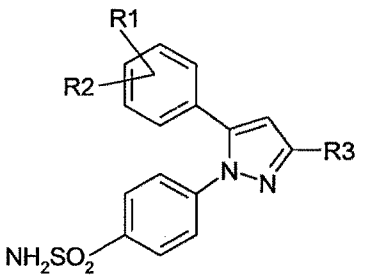


C.N	R	pIC_{50}	
		Actual	Predicted
73	$COCF_3$	6.92	6.66
74	COPh	5.99	5.70
75	SO_2CF_3	7.22	6.96
76	CH_2OAc	6.32	6.10

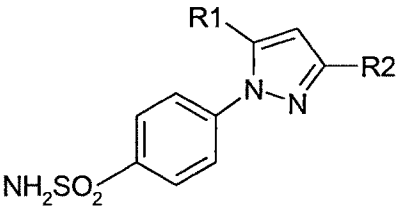
77*	$CH_2OPh(p\text{-}Cl)$	7.52	6.57
78	CH_3CF_3	6.85	7.04

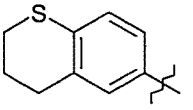
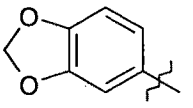


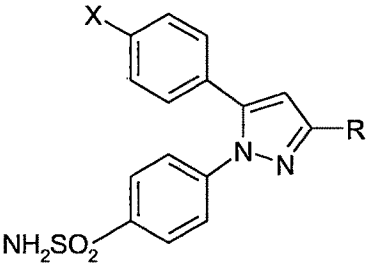
C.N	R1	R2	pIC_{50}	
			Actual	Predicted
79	4-Cl	CHF_2	8.00	7.86
80	2-Me	CF_3	7.16	7.55
81	3-Me	CF_3	6.95	6.88
82	4-Me	CF_3	7.39	7.19
83	4-Me	CHF_2	7.88	8.07



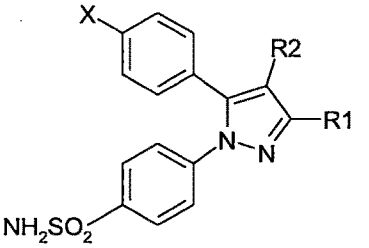
C.N	R1, R2	R3	pIC_{50}	
			Actual	Predicted
84	3-fluro-4-methoxy	CHF_2	7.30	7.36
85	3-chloro-4-methoxy	CHF_2	7.56	7.59
86*	3-ethyl-4-methoxy	CF_3	6.36	6.87
87	3,4-dimethoxy	CF_3	6.22	6.56
88	3-chloro-4-(methylamino)	CHF_2	7.56	7.45
89	3,5-dichloro-4-methoxy	CHF_2	7.67	7.79
90	3-chloro-4-methoxy-5-methyl	CF_3	7.18	7.15



C.N	R1	R2	pIC_{50}	
			Actual	Predicted
91		CF_3	6.63	6.85
92		CHF_2	7.61	7.83

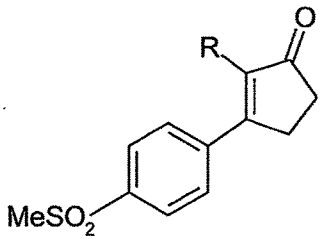


C.N	R	X	pIC_{50}	
			Actual	Predicted
93*	CH_2OH	Cl	6.08	6.75
94	CH_2OCH_2Ph	Cl	7.53	7.55
95	4-methoxyphenyl	Cl	7.00	6.85
96	CN	F	6.46	6.28

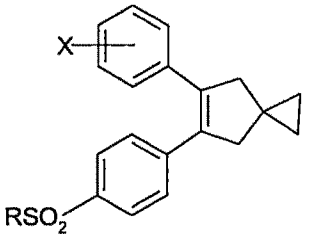


C.N	R1	R2	X	pIC_{50}	
				Actual	Predicted
97	CF_3	Me	Cl	7.65	7.87
98	CF_3	Et	Cl	7.55	7.53
99	CF_3	OMe	H	7.09	7.24

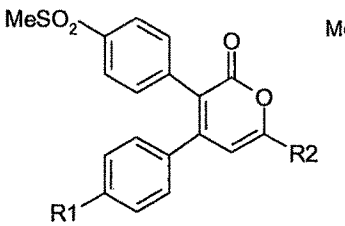
100	H	CN	CH_3	7.11	7.14
101*	H	NO_2	H	6.53	6.20
102	CH_3	Cl	H	7.55	7.44
103	CH_2OH	Cl	Cl	6.46	6.60
104	$CONH_2$	Cl	Cl	5.96	5.90



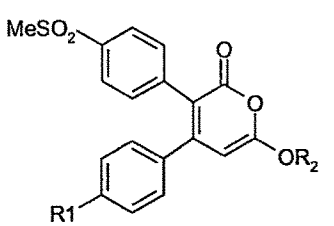
C.N	R	pIC_{50}	
		Actual	Predicted
105*	3-flurophenyl	6.46	6.33
106	3,4-diflurophenyl	6.67	6.32
107	3,5-dichlorophenyl	6.55	6.73
108	3-fluro-4-chlorophenyl	6.82	6.5 1
109	5-chloro-3-pyridyl	5.82	5.97
110	5-bromo-3-pyridyl	5.76	5.94
111	2-pyridyl	5.00	5.22



C.N	X	R	pIC_{50}	
			Actual	Predicted
112	4-fluro	CH_3	8.09	8.08
113	4-fluro	NH_2	8.52	8.33
114	4-methyl	CH_3	8.82	8.50
115*	4-methoxy	CH_3	8.30	7.93
116	4-trifluromethyl	CH_3	8.69	8.50
117	4-trifluromethyl	NH_2	9.00	8.97
118	3-fluro-4-methoxy	NH_2	8.69	8.62
119	3-bromo-4-methoxy	NH_2	8.69	8.80
120	3,4-difluro	CH_3	8.52	8.44
121	3,4-difluro	NH_2	8.52	8.68
122*	3,4-dichloro	CH_3	8.52	8.62
123	3,4-dichloro	NH_2	9.00	8.86
124	3-chloro-4-fluro	NH_2	8.82	8.71

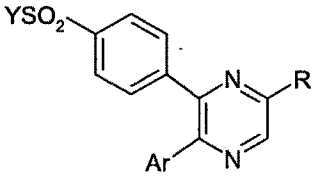


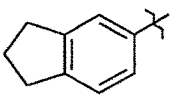
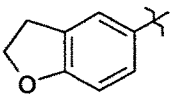
125-126



127-131

C.N	R1	R2	pIC_{50}	
			Actual	Predicted
125	H	Me	6.16	5.78
126	H	Et	5.82	5.62
127	H	Me	4.54	4.44
128	F	Me	4.22	4.42
129*	H	Et	5.88	5.45
130	H	i.Pr	5.69	5.45
131	F	t.Bu	4.69	4.84



C.N	Ar	R	Y	pIC_{50}	
				Actual	Predicted
132	4-OMe-phenyl	Me	Me	5.79	5.99
133	4-F-phenyl	Me	Me	5.26	5.15
134	3-F-phenyl	Me	Me	5.35	5.04
135	4-Me-phenyl	Me	NH_2	5.91	5.99
136*	4-OMe-phenyl	H	NH_2	5.71	6.09
137	4-OMe-phenyl	Me	NH_2	6.33	6.35
138	4-Cl-phenyl	H	NH_2	5.35	5.40
139	4-Br-phenyl	H	NH_2	5.13	5.46
140	4-OMe-3-Me-phenyl	H	NH_2	5.40	5.34
141	4-OMe-3-Me-phenyl	Me	NH_2	5.79	5.62
142*		H	NH_2	5.99	5.77
143		H	NH_2	5.35	5.22

Alignment rule

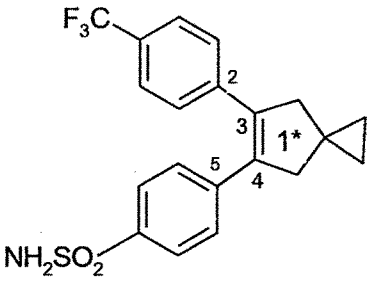


Figure 4.4.1 Template structure **117** with atoms (2,3,4,5) and centroid (1*) used for alignment

4.4.2 Results and discussion

After developing 3D QSAR models for selective COX-II inhibitors belonging to different chemical classes, it was planned to study the feasibility of developing a universal (i.e. any 1,2-diaryl-heterocyclic/carbocyclic category of compounds) 3D-QSAR model for selective COX-II inhibitors. If such a model could be developed then it would be employed in future for predicting the activity of a compound which is yet to be synthesized.

As all the COX-II inhibitors exert their action by blocking COX-II enzyme irrespective of their structural features. Hence, we developed a common 3D CoMFA model for these inhibitors using eight different data sets belonging to imidazole, pyrrole, pyrazole, furanone, cyclopentenone, spiroheptenes, phenylpyran-2-ones and pyrazine classes of compounds.

Various alignment rules were tried but only the combination of atom and centroid-based alignment yielded best results and hence only discussed. The CoMFA model generated using this model displayed good predictive power. Contour maps were

generated using this model which can be further used to design potent COX-II inhibitors.

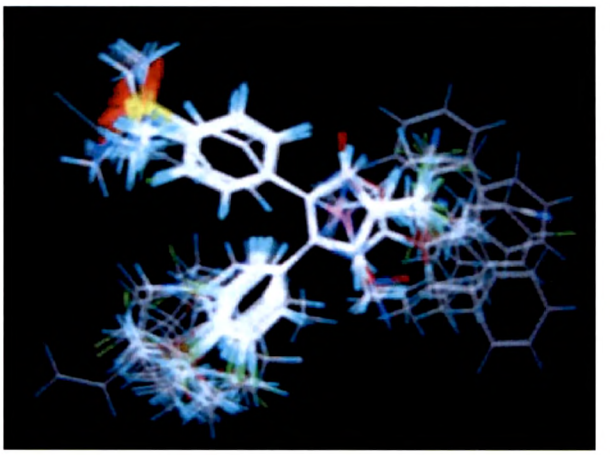


Figure 4.4.2 Superimposition of 143 COX-II inhibitors on template (**117**)

All the COX-II inhibitors were aligned (**Figure 4.4.2**) using atom and centroid-based alignment and further used for CoMFA analysis.

4.4.2.1 CoMFA model for COX-II inhibitors

3D-QSAR studies employing the most common and effective technique CoMFA was carried out on a set of 143 compounds belonging to different chemical classes. Compounds were pooled based on the structural and biological variations. The training set and test set comprised of 123 and 20 molecules, respectively. The negative logarithm of the *in vitro* inhibitory activity (pIC_{50}) was used as a dependent variable, as described earlier. Different alignment rules were tried but only the atom and centroid-based alignment (**Figure 4.4.1**) yielded best results and is discussed here.

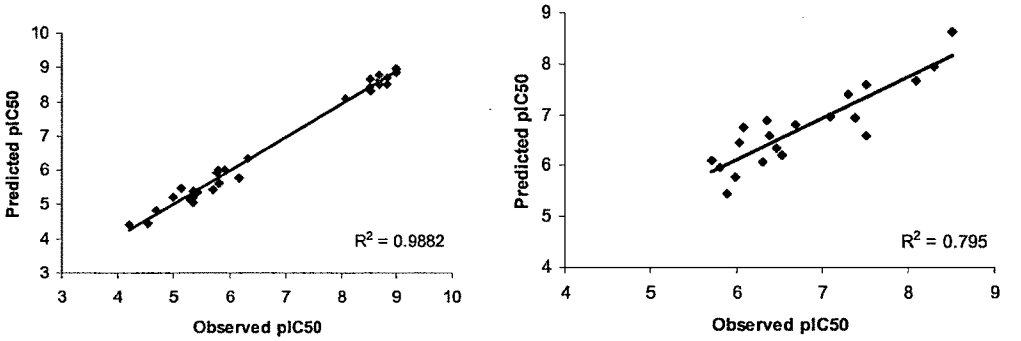
CoMFA model generated using combination of atom and centroid-based alignment showed cross-validated r^2 0.749 with seven components, non cross-validated r^2 0.962, F value 420.83, low standard error of estimate (0.199), bootstrapped r^2 0.981 and

predictive r^2 0.776. The results of CoMFA analysis are shown in **Table 4.4.2**. The predicted activities of training and test set molecules are given in **Table 4.4.1**. **Figure 4.4.3** depicts a graph of actual *versus* predicted activities of the molecules used in training and test set.

Table 4.4.2 Results of CoMFA analysis

Statistical parameters	Values
r^2_{cv}	0.749
SEP	0.513
N_c	7
r^2_{ncv}	0.962
SEE	0.199
F	420.83
Prob. $r^2 = 0$	0.0
r^2_{bs}	0.981
S.D	0.007
r^2_{pred}	0.776

A good correlation was observed within the actual and predicted activities for both training (r^2 0.988) and test set (r^2 0.794) compounds. The obtained model was validated by predicting the activity of coxibs which are in clinical use. Comparison of the actual and predicted activities for coxibs is shown in **Table 4.4.3**.



(a) (b)

Figure 4.4.3 A graph of actual *versus* fitted/predicted activities for training (a) and test (b) set molecules

The 3D contour maps generated by this model were used for further analysis to identify the structural features governing COX-II inhibitory activity.

Table 4.4.3 Comparison of actual and predicted activities for coxibs

Compound	pIC_{50}	
	Actual	Predicted
Celecoxib	6.06	6.60
Rofecoxib	6.30	6.75
Valdecoxib	6.06	6.19
Etoricoxib	6.00	6.93

Graphical interpretation of CoMFA model

The steric contour maps indicate green and yellow contours as sterically favorable and unfavorable areas. Blue and red contours in the electrostatic maps indicate areas where positive and negative charge substituent's favor COX-II inhibitory activity. To aid in visualization, the template molecule **117** is displayed in ball and stick. The contour maps are discussed with reference to the template molecule.

Figure 4.4.4 corresponds to the CoMFA steric and electrostatic contour maps generated using the best CoMFA model. Significant green contour surrounding spiroheptene ring represents favorable steric area to increase COX-II inhibitory activity. This is the reason why compound **77** having bulkier substituent [(CH_2OPh(*p*-Cl)] showed high COX-II inhibitory activity (pIC_{50}=7.52). Compound **94** having bulkier 'CH_2OCH_2Ph' group attached to central pyrazole ring showed COX-II inhibitory activity on a higher side (pIC_{50}=7.53). Similar is the case with compound **42**, where the *n*-butyl group provides bulkiness necessary for COX-II inhibition (pIC_{50}=7.69). Another small favorable green contour (not so prominent) is observed near to the trifluromethyl group indicating incorporating bulkier groups here may increase the COX-II inhibitory activity. This is the reason why compound **92** showed high inhibitory activity (pIC_{50}=7.61).

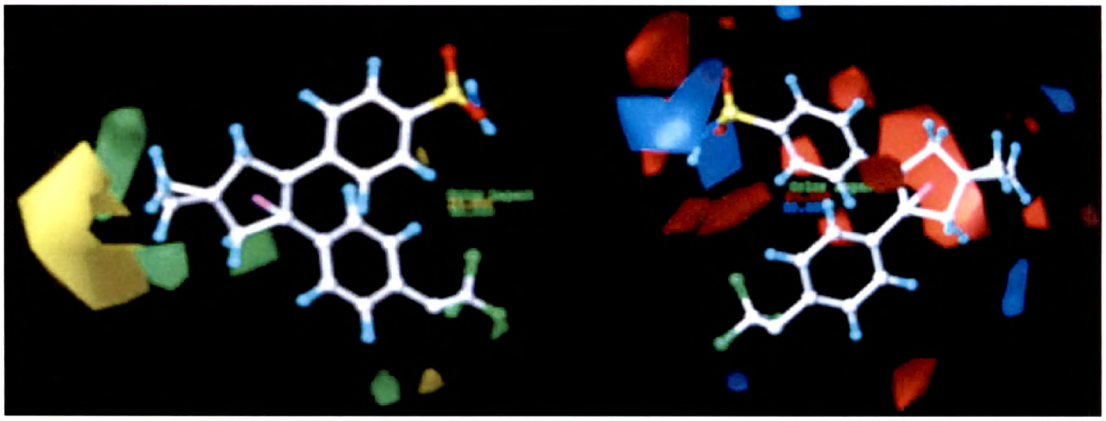


Figure 4.4.4 CoMFA standard deviation coefficient steric and electrostatic contour plots with compound **117**; green contours indicate regions where bulky group increases activity, whereas yellow contour indicates regions where bulky group decreases activity; blue contours indicate regions where positively charged groups increase activity, where red contours indicate regions where negative charge increases activity.

Large and prominent unfavorable contour is observed close to spiroheptene ring. This indicates that bulkier substituents orienting towards this contour may be detrimental for COX-II inhibitory activity. Hence, all the compounds of phenylpyran-2 one series show poor COX-II inhibitory activity (pIC_{50} in the range of 4.22 to 5.82).

Significant red contours observed surrounding the spiroheptene ring represents regions where electronegative groups are favorable for COX-II inhibitory activity. This is in accordance with the compounds belonging to pyrazole class where electronegative trifluromethyl group is kept constant throughout the series. In case with furanone derivatives, electronegative (carbonyl) group is attached to the furan ring. This may be one of the reasons why furanone derivatives displayed good COX-II inhibitory activity.

Both red and blue contours are observed in the vicinity of benzene-bearing sulfonamide group which indicates that both electronegative and electropositive groups are necessary for COX-II inhibitory activity. Hence, compounds **67-70** displayed good activity (pIC_{50} in the range 7.52 to 8.52).

A small blue contour is observed near to the trifluromethyl group indicating electron donating substituents may increase the COX-II inhibitory activity. This may be the reason why compound **88** with methyl amino substituent displayed good inhibitory activity (pIC_{50}=7.56). Another blue contour is also observed in the vicinity of spiroheptene ring which indicates incorporation of electron donor groups may increase the COX-II inhibitory activity. Compounds **97** (pIC_{50}=7.65) and **98** (pIC_{50}=7.55) having methyl and ethyl groups attached to central spiroheptene ring, respectively show good COX-II inhibitory activity.

In brief, a very good universal 3D CoMFA model has been developed based on 143 compounds having different scaffolds. Compounds were aligned using different alignment rules but the combination of both atom and centroid-based alignment rule yielded best results. A good correlation was observed between actual and predicted activities of training set (r^2 0.988) and test set (r^2 0.794) compounds. The model also predicted the activities of coxibs in market very well confirming its utility in designing more potent and selective COX-II inhibitors. The model could be used safely for predicting the activity of any comound belonging to 1,2-diarylheterocyclic/carbocyclic system, which could be synthesized after assessing the *in silico* COX-II inhibitory activity.

4.4.3 References

1. Khanna, I. K.; Weier, R. M.; Yu, Y.; Xu, X. D.; Koszyk, F. J.; Collins, P. W.; Koboldt, C. M.; Veenhuizen, A. W.; Perkins, W. E.; Casler, J. J.; Masferrer, J. J.; Zhang, Y. Y.; Gregory, S. A.; Seibert, K and Isakson, P. C. *J. Med. Chem.* **1997**, 40, 1634.
2. Shin, S. S.; Byun, Y.; Lim, K. M.; Choi, J. K.; Lee, K. W.; Moh, J. H.; Kim, J. K.; Jeong, Y. S.; Kim, J. Y.; Choi, Y. H.; Koh, H. J.; Park, Y. H.; Noh, M. S and Chung, S. *J. Med. Chem.* **2004**, 47, 792.
3. Khanna, I. K.; Weier, R. M.; Yu, Y.; Collins, P. W.; Miyasharo, J. M.; Koboldt, C. M.; Veenhuizen, W.; Currie, J. L.; Seibert, K and Isakson, P. C. *J. Med. Chem.* 40, **1997**, 1619.
4. Penning, T. D.; Talley, J. J.; Bertenshaw, S. R.; Carter, J. S.; Collins, P. W.; Docter, S.; Graneto, M. J.; Lee, L. F.; Malecha, J. W.; Miyashiro, J. M.; Rogers, R. S.; Yu,, S. S.; Anderson, G. D.; Burton, E. G.; Cogburn, J. N.; Gregory, S. A.; Koboldt, C. M.; Perkins, W. E.; Seibert, K.; Veenhuizen, A. W.; Zhang, Y. Y and Isakson, P. C. *J. Med. Chem.* 40, **1997**, 1347.
5. Huang, H. C.; Li,, J. J.; Garland, D. J.; Chamberlain, T. S.; Reinhard, E. J.; Manning, R. E.; Seibert, K.; Kiboldt, C. M.; Gregory, S. A.; Anderson, G. D.; Veenhuizen, A. W.; Zhang, Y.; Perkins, W. E.; Burton, E. G.; Cogburn, J. N.; Isakson, P. C. and Reitz, D. B. *J. Med. Chem.* 39, **1996**, 253.
6. Black, W. C.; Brideu, C.; Chan, C. C.; Charleson, S.; Chauret, N.; Claveau, D.; Etheir, D.; Gordon, R.; Greig, G.; Guay, J.; Huges, G.; Jolicoeur, P.; Leblanc, Y.; Griffith, D. N.; Ouimet, N.; Reindeau, D.; Visco, D.; Wang, Z.; Xu, L and Prasit, P. *J. Med. Chem.* 42, **1999**, 1274.
7. Singh, S. K.; Saibaba, V.; Ravikumar, V.; Rudrawar, S. W.; Daga, P.; Rao, C. S.; Akhila, V.; Hegde, P and Rao, Y. K. *Bioorg. Med. Chem.* 12, **2004**, 1881.
8. Rao, P. N.; Amini, M.; Li, Habeed, A. G and Knaus, E. E. *J. Med. Chem.* 46, **2003**, 4872.