

# A CoMFA study of COX-2 inhibitors with receptor based alignment

Prasanna A. Datar, Evans C. Coutinho\*

*Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India*

Received 14 May 2004; received in revised form 23 July 2004; accepted 23 July 2004

Available online 27 August 2004

## Abstract

A diverse set of 53 cyclooxygenase-2 (COX-2) inhibitors which were aligned in two different ways were subjected to CoMFA analysis. The first method of alignment of the molecules was based on the binding information sourced from the crystallographic study, from which CoMFA Model 1 was derived. The second mode of alignment was generated by docking the inhibitors in the binding pocket using the *DOCK* and *AFFINITY* suite of programs; this gave a second model. The CoMFA Model 2 was slightly better than Model 1 in terms of the statistical parameters  $r^2$  and  $q^2$ . The two models could predict very well the activity of a test set of diverse molecules, with a predictive  $r^2$  of 0.593 and 0.768, respectively. Besides the QSAR results, the docking studies give a deep insight into the H-bonding interactions between the inhibitors and residues in the active site of the enzyme, which can be exploited in designing better inhibitors. Useful ideas on activity improvement could be gleaned from these models.

© 2004 Elsevier Inc. All rights reserved.

**Keywords:** COX-2; 3D-QSAR; CoMFA; Docking; Ligand–receptor interaction

## 1. Introduction

Cyclooxygenase-2 (COX-2) inhibition has been one of the most widely investigated areas of research in the last decade. Pain is a common symptom in many disorders, and relief of pain has become one of the prime objectives. The discovery that the key enzyme cyclooxygenase in the arachidonic acid metabolism, exists in two isoforms, namely, the constitutive cyclooxygenase-1 (COX-1) and the inducible cyclooxygenase-2 (COX-2), has generated new avenues for drug design [1]. Both forms differ in their regulation and expression. COX-1 is responsible for the biosynthesis of prostaglandins (PGs), which are involved in the cytoprotection of the gastrointestinal tract and platelet aggregation. COX-2 is induced by proinflammatory molecules such as interleukin-1 (IL-1), tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), lipopolysaccharide (LPS), and carrageenan, and plays a major role in the biosynthesis of PGs in inflammatory cells (monocytes and macrophages). The interruption of COX-1 activity may lead to gastrointestinal toxicity such as ulcera-

tion, bleeding, and perforation [2]. The traditional non-steroidal anti-inflammatory drugs (NSAIDs) inhibit both COX-1 and COX-2 and hence downregulate the prostaglandins in almost all cells and tissues. This accounts for their anti-inflammatory activity as well as side effects. However, selective inhibition of COX-2 over COX-1 is beneficial for treatment of inflammatory diseases with reduced ulcerogenic side effects.

The two COX isoforms are approximately 60% homologous. The ability to inhibit one isoform selectively, is attributed to the different amino acids at position 523 in the COX-1 and COX-2 enzymes [3]. The COX-1 enzyme has an isoleucine residue at this position, while the corresponding residue in the COX-2 enzyme is valine. This seemingly minute difference in structure, results in a larger central channel in the COX-2 isoform. As a result, molecules, which are too large to enter the COX-1 channel are able to enter the COX-2 channel, based solely on size discrimination at the active site. Exploitation of this phenomenon, has been the key in developing selective COX-2 inhibitors.

Rational drug design has led to a wide range of diaryl-heterocycles that have been developed as selective COX-2 inhibitors. Celecoxib [4] and rofecoxib [5] were among the

\* Corresponding author. Tel.: +91 22 26670871; fax: +91 22 26670816.  
E-mail address: [evans-im@eth.net](mailto:evans-im@eth.net) (E.C. Coutinho).

earliest to be marketed for the treatment of acute pain, rheumatoid arthritis, and osteoarthritis.

Various approaches for development of COX-2 inhibitors have appeared in the literature over the past several years. The methods employed include 3D-QSAR studies such as comparative molecular field analysis (CoMFA) [6–9] and receptor surface analysis (RSA) [10]. All models reported are of sufficiently high quality to be used as an estimator of biological activity of unknown compounds; however, each model has some limitations. Most of the data sets examined were small and limited in structural diversity. Chavatte et al., had initially derived a 3D-QSAR model on a set of selective COX-2 inhibitors [8a] and later revised it to include a larger and more diverse set of compounds [8b]. However, some inhibitors belonging to the pyridine, thiazolone, thiadiazole and oxadiazole families were not considered. On the other hand, Desiraju et al., [6] used only the imidazole class for their studies. These are some of the shortcomings of the previously reported CoMFA models. Our intention in this work is to develop a 3D-QSAR model using as input diverse compounds that will ultimately be able to predict the activity of a diverse set of test molecules. Also in this paper we report use of the docking methodology as implemented in *AFFINITY* [11] which is based on both Monte Carlo and simulated annealing strategies, to generate simultaneously the conformations and alignment needed for the CoMFA study.

## 2. Materials and methods

A training set of 53 molecules and 14 test molecules were selected from the literature. The activity of all the molecules was measured by the same assay on human recombinant enzyme [12]. The enzyme inhibitory activity of all these molecules covered 7 log orders. The molecules were from nine different structural classes, which were distinguished on the basis of the central cyclic ring (marked B in Fig. 1) and included pyrrole [13], imidazole [14,15], cyclopentene [16,17], benzene, [18] pyrazole, [3] spiroheptene, [19] isoxazole, [20] pyridine, [21] thiazolones, [22] thiadiazoles and oxadiazoles [23]. The chemical structures of the 53 training and 14 test set molecules are given in Table 1(A–I).

CoMFA studies were carried out using *SYBYL* [24] (Tripos Inc., USA) software, while the docking studies were carried out using *DOCK* 4.0 [25] and *AFFINITY* [11] (Accelrys Inc. USA) programs on a Silicon Graphics O2 workstation.

## 3. Molecular modeling and QSAR methods

Structures were built with the *BUILDER* module in *INSIGHT II* v.98 (Accelrys Inc., USA). The structures were optimized with the CVFF force field [26] to an energy gradient of 0.001 kcal/(mol Å).

Table 1

Training (1–53) and test set (54–67) molecules used in CoMFA studies of COX-2 inhibitors

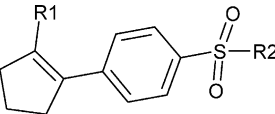
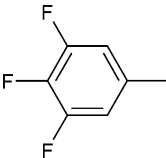
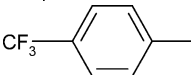
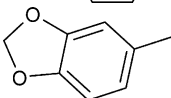
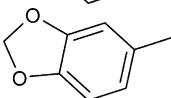
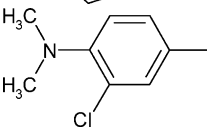
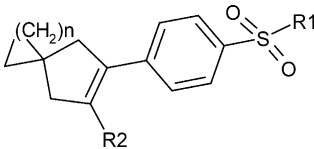
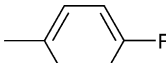


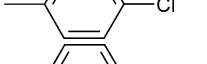
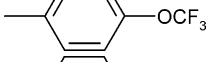
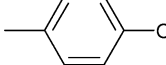
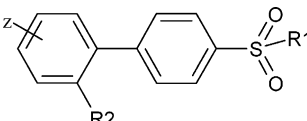
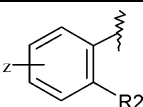
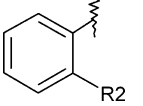
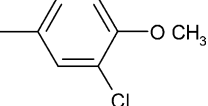
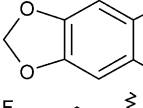
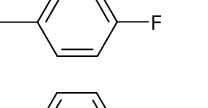
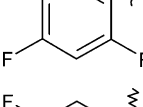
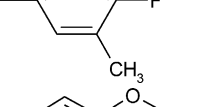
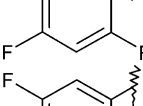
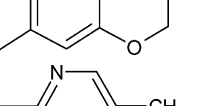
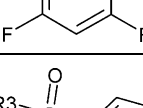
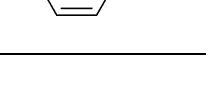
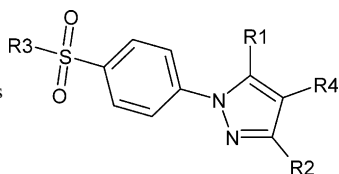
A. Cyclopentene class			
			
ID	pIC <sub>50</sub> (mol)	R1	R2
1	5.53		NH <sub>2</sub>
2	6.82		NH <sub>2</sub>
3	7.67		CH <sub>3</sub>
4	8.69		NH <sub>2</sub>
54	8.69		NH <sub>2</sub>
B. Spiroheptene class			
			
ID	pIC <sub>50</sub> (mol)	R1	R2
5	8.39	CH <sub>3</sub>	
6	7.20	CH <sub>3</sub>	
7	9.00	NH <sub>2</sub>	
8	6.86	CH <sub>3</sub>	
9	8.00	CH <sub>3</sub>	
55	8.09	CH <sub>3</sub>	
C. Benzene class			
			

Table 1 (Continued)

ID	pIC <sub>50</sub> (mol)		R1	R2
10	4.94		CH <sub>3</sub>	
11	7.08		CH <sub>3</sub>	
12	8.69		NH <sub>2</sub>	
13	6.46		CH <sub>3</sub>	
14	4.28		CH <sub>3</sub>	

## D. Pyrazole class



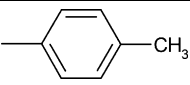
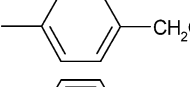
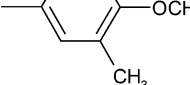

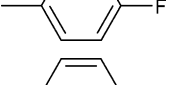
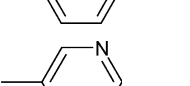
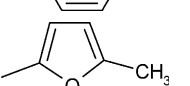
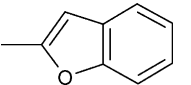

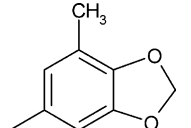
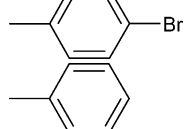
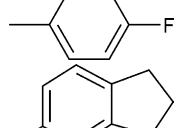
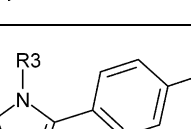
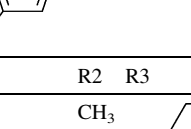
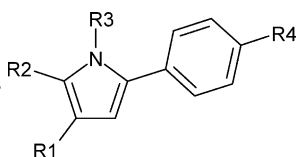
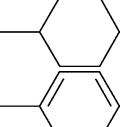
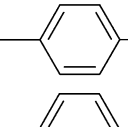
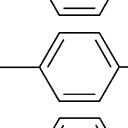
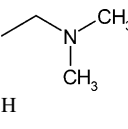
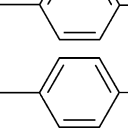
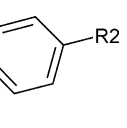
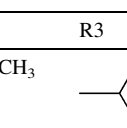
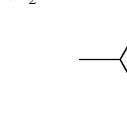
ID	pIC <sub>50</sub> (mol)	R1	R2	R3	R4
15	7.39		CF <sub>3</sub>	NH <sub>2</sub>	H
16	4.03		CF <sub>3</sub>	NH <sub>2</sub>	H
17	8.03		CF <sub>3</sub>	NH <sub>2</sub>	H
18	5.44		CF <sub>3</sub>	NH <sub>2</sub>	OH
19	7.00		CF <sub>3</sub>	CH <sub>3</sub>	H
20	7.38		CF <sub>3</sub>	NH <sub>2</sub>	H
21	4.34		CF <sub>3</sub>	NH <sub>2</sub>	H
22	5.48		CHF <sub>2</sub>	NH <sub>2</sub>	H
23	6.05		CF <sub>3</sub>	NH <sub>2</sub>	H

Table 1 (Continued)

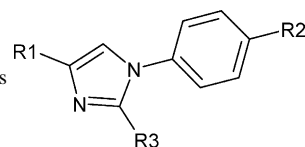
ID	pIC <sub>50</sub> (mol)	R1	R2	R3	R4
24	7.28		CF <sub>3</sub>	NH <sub>2</sub>	H
53	8.03		CF <sub>3</sub>	NH <sub>2</sub>	H
56	4.32		H	NH <sub>2</sub>	CH <sub>3</sub>
57	6.46		CN	NH <sub>2</sub>	H
58	7.50		CF <sub>3</sub>	NH <sub>2</sub>	H

## E. Pyrrole class



ID	pIC <sub>50</sub> (mol)	R1	R2	R3	R4
25	6.28	H	CH <sub>3</sub>		SO <sub>2</sub> CH <sub>3</sub>
26	7.22	H	CH <sub>3</sub>		SO <sub>2</sub> CH <sub>3</sub>
27	5.54	H	CH <sub>3</sub>		SO <sub>2</sub> CH <sub>3</sub>
28	4.00		CH <sub>3</sub>		SO <sub>2</sub> CH <sub>3</sub>
29	4.99	H	H		SO <sub>2</sub> CH <sub>3</sub>
59	6.29	H	H		F
60	5.79	-COCH <sub>3</sub>	CH <sub>3</sub>		SO <sub>2</sub> CH <sub>3</sub>

## F. Imidazole class



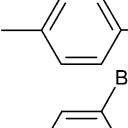
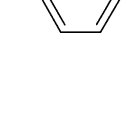
ID	pIC <sub>50</sub> (mol)	R1	R2	R3
30	5.24	CF <sub>3</sub>	-SO <sub>2</sub> CH <sub>3</sub>	
31	8.15	CF <sub>3</sub>	-SO <sub>2</sub> NH <sub>2</sub>	

Table 1 (Continued)

ID	pIC <sub>50</sub> (mol)	R1	R2	R3
32	6.04	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
33	4.00	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
34	7.09	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
35	4.91	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
36	6.29	CF <sub>3</sub>	–SO <sub>2</sub> NH <sub>2</sub>	
37	4.10	CH <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
38	3.02	CH <sub>2</sub> OH	–SO <sub>2</sub> CH <sub>3</sub>	
39	5.23	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
40	5.76	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
41	7.58	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
42	6.36	CF <sub>3</sub>	–SO <sub>2</sub> NH <sub>2</sub>	
61	5.77	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	
62	6.32	CF <sub>3</sub>	–SO <sub>2</sub> CH <sub>3</sub>	

G. Pyridine class

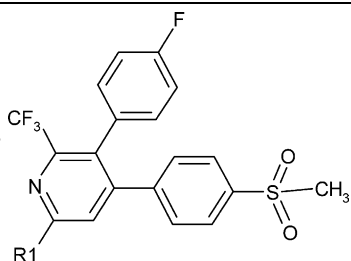
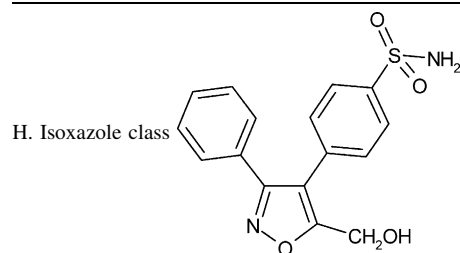


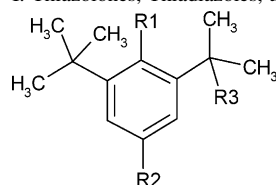
Table 1 (Continued)

ID	pIC <sub>50</sub> (mol)	R1
43	6.22	Br
44	6.69	–OCH <sub>2</sub> C≡CH
64	6.52	–OCH <sub>3</sub>
65	6.52	–OCH <sub>2</sub> CH <sub>3</sub>



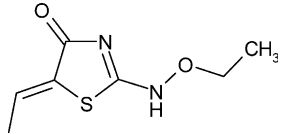
ID	pIC <sub>50</sub> (mol)
63	6.74

I. Thiazolones, Thiadiazoles, and Oxadiazoles class



ID	pIC <sub>50</sub> (mol)	R1	R2	R3
45	6.85	OH		CH <sub>3</sub>
46	4.00	OH		CH <sub>3</sub>
47	5.18	OH		CH <sub>3</sub>
48	7.32	OH		CH <sub>3</sub>
49	7.19	OH		CH <sub>3</sub>
50	5.76	OH		CH <sub>3</sub>
51	4.25	OH		CH <sub>3</sub>
52	6.24	OH		H
66	5.82	OCH <sub>3</sub>		CH <sub>3</sub>

Table 1 (Continued)

ID	pIC <sub>50</sub> (mol)	R1	R2	R3
67	5.49	OH		CH <sub>3</sub>

### 3.1. Alignment of molecules

One of the prime requirements in CoMFA studies, is the alignment of all compounds relative to one another, so that they have a comparable conformation and orientation in space. Two different methods were employed for generation of conformation and alignment. In the first method, the X-ray crystal structure of SC-558 (PDB code:1CX2), [27] a selective COX-2 inhibitor complexed with the enzyme, was used as a template for superimposition, assuming that this conformation represents the most probable bioactive conformation. A nine point feature (Fig. 1) was selected for the alignment of all compounds over the SC-558 template. The method called ‘template force’ in *DISCOVER* (Accelrys Inc., USA) was used to force the common features of the ligands over the template structure, with a force constant of 50 kcal/(mol Å<sup>2</sup>). Further refinement of structures was carried out to reduce the strain, by minimization without constraints. The structures were refined using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, to an energy gradient of 0.001 kcal/(mol Å). The molecules aligned in this manner are exhibited in Fig. 2a and were then subjected to CoMFA analysis.

### 3.2. Alignment by docking

A second mode of alignment was also carried out using docking strategies. Again the X-ray crystal structure of COX-2 complexed with SC-558 was selected for this docking strategy. Hydrogens were added to the enzyme and the heme part was removed from the complex. Residues such as histidine, lysine, aspartate and glutamate were ionized by setting the pH of the system to 8.0. This resulted in an overall charge of −3.0 for the system. The enzyme along with the inhibitor was energy minimized using conjugate gradients, with the backbone atoms of the enzyme tethered by a force constant of 100 kcal/(mol Å<sup>2</sup>). The ligand (SC-558) was allowed to move freely during the minimization. Electrostatic energy of the system was calculated with a distance dependent dielectric constant ( $\epsilon$ ). The minimization was carried out till a gradient of 0.01 kcal/(mol Å) was reached.

All the training and test set molecules were then docked into the COX-2 enzyme as prepared above, initially with the *DOCK* program. Selected configurations were then refined with the *AFFINITY* program.

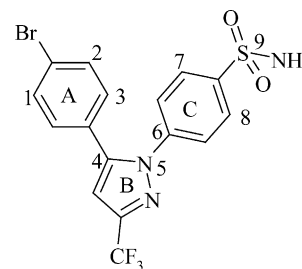


Fig. 1. Nine point feature defined for superimposition of the training and test set molecules on to SC-558.

### 3.3. DOCK protocol

All residues within a 15 Å radius from the centroid of SC-558 defined the active site. A standard protocol for docking as described in the *DOCK* manual was followed. The various binding modes were scored based on interaction energy. The ligand orientation with ‘best *DOCK* score’ (lowest interaction energy) was subjected to refinement with the docking program *AFFINITY*.

In *AFFINITY*, a combination of Monte Carlo and simulated annealing procedures were used to dock the ligand. *AFFINITY* requires a reasonably good starting conformation and position of the ligand inside the pocket, since it is not efficient in sampling a large conformational space. For docking by *AFFINITY*, residues not in the binding site were held rigid, while both the binding site atoms (residues within a 15 Å radius from the ligand) and the ligand atoms were free to move during the docking procedure. Various empirical penalty terms such as ‘ligand confinement’ (to prevent drift of the ligand away from the pocket), ‘confine radius’ (width: 10 Å of the flat bottomed potential restraining-force: 100 kcal/(mol Å<sup>2</sup>)—the relative positions of the centers of mass of the binding site and the ligand; and tethering-force: 100 kcal/(mol Å<sup>2</sup>)—residues other than those in the binding site) were used in the docking strategy. Initial random movements of the ligand were executed with the Monte Carlo method. The energy of the system after each Monte Carlo step was calculated with a quartic potential for the VDW interactions while neglecting the Coulombic interactions. The maximum movement of the ligand was set to 1.0 Å, and the maximum angle of rotation was set to 180°. Only those configurations generated by the Monte Carlo procedure were selected whose energy was within 200 kcal/mol of the lowest energy configuration. The best structures obtained from the Monte Carlo procedure were then minimized and ranked on the basis of energy, with the electrostatic component calculated by the more accurate Cell multipole method [28]. The 10 lowest energy conformations were then refined with simulated annealing, where a linear decrease in temperature from 500 K to 300 K in steps of 100 K for 100 fs was adopted. The annealed structures were minimized to a gradient of 0.001 kcal/(mol Å). The lowest energy struc-

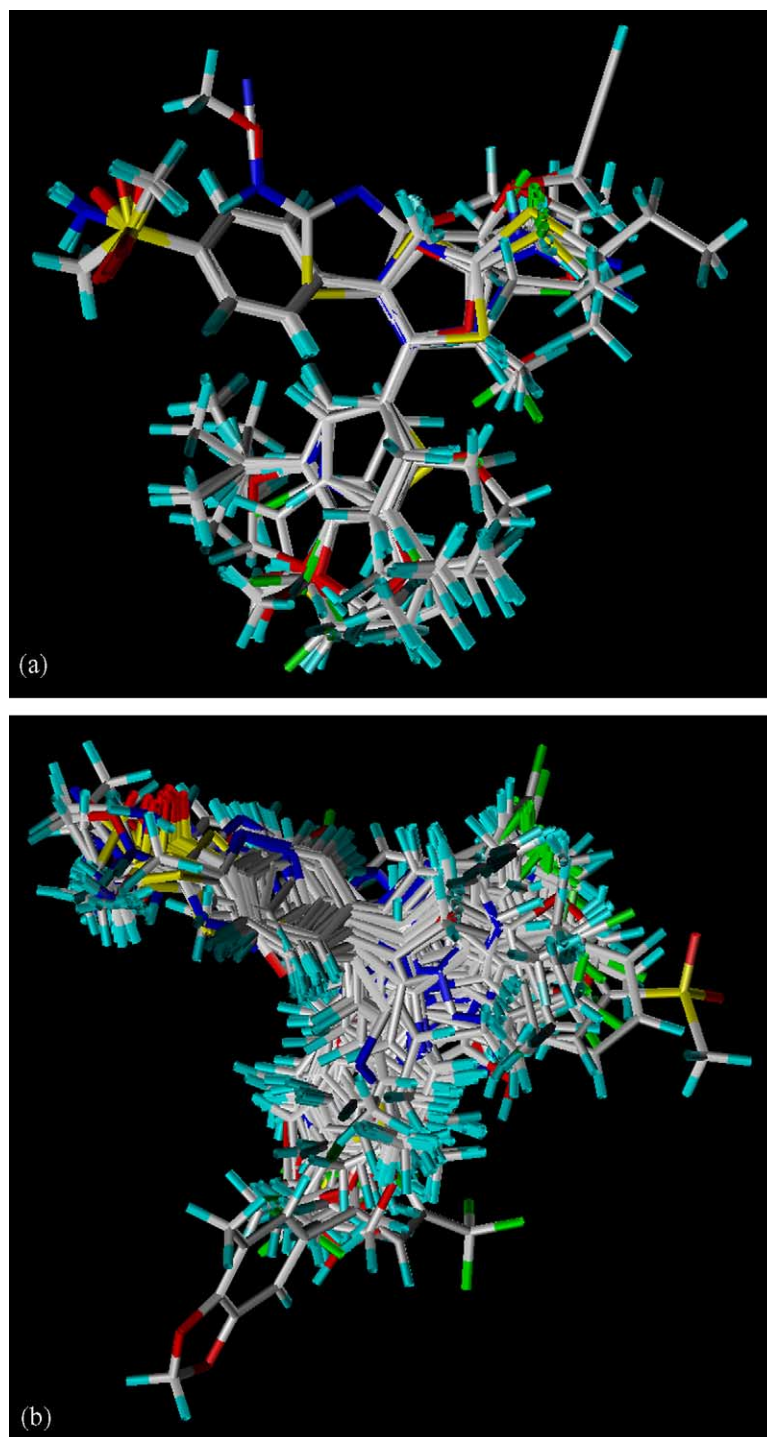


Fig. 2. (a) Alignment based on template forcing molecules on to the crystal structure of SC-558. (b) Alignment based on docking studies using *AFFINITY*.

ture of each ligand was then used to construct the alignment of the *ensemble*. The protein–ligand complexes generated by *AFFINITY* were then superimposed using only the protein backbone atoms, letting the ligands take up their positions without regard to each other. The molecules aligned in this manner are shown in Fig. 2b.

### 3.4. CoMFA studies

CoMFA [29] was carried out with the QSAR module in the *SYBYL* (Tripos Inc., USA) [24] software. The CoMFA model development was done with the grid, grid spacing, probe, non-bonded cutoff, scaling of variables, column

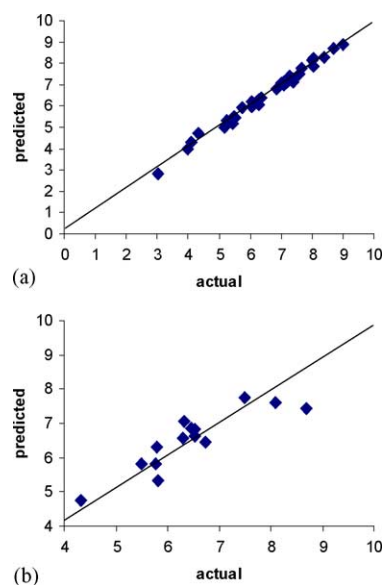


Fig. 3. Plot of actual vs. predicted activity ( $\text{pIC}_{50}$ ) of the training (a) and test (b) sets by CoMFA Model 2.

filtering and PLS [30] analysis as implemented in our earlier paper [31]. Orientation [31,32] of the assembly of molecules in the grid was also explored to improve the cross-validated  $r^2$  ( $q^2$ ) of the model. A standard leave-one-out (LOO) cross-validation and final fitting of equations with optimal number of components was followed till models with the highest  $q^2$ , lowest PRESS and lowest SDEP were obtained.

Final models were selected primarily based on the values of better cross-validated  $r^2$  as well as better predictive  $r^2$  for the test set molecules (Table 2). To gauge the robustness of

the models, bootstrapping [33] with 100 groups and randomization (scrambling of activity values) with 100 runs was carried out [34]. The best CoMFA model was improved using the  $q^2$ -guided region selection [35] ( $q^2$ -GRS).

### 3.5. Graphical analysis

CoMFA contour maps were generated using scalar products of coefficients and standard deviation (STDEV\*-COEFF) set at 80% and 20% levels for favored and disfavored levels, respectively for steric and electrostatic fields.

## 4. Results and discussion

CoMFA models were generated for each of the two alignment methods. Table 2 displays all the statistical data for the two models. Cross-validated  $r^2$  ( $q^2$ ) of 0.6–0.7 reflects a good predictive power of the models. A plot of actual  $\text{pIC}_{50}$  versus predicted  $\text{pIC}_{50}$  for the training set (a) and the test set (b) for CoMFA Model 2 are shown in Fig. 3. Reasonably low standard error of estimate, standard deviation of error of prediction and high values of the correlation coefficient, speak of a good correlation between the training set molecules. The high bootstrap  $r^2$  values with its low standard deviation and negative/low  $q^2$  values in the randomization tests (data not given) reflect robustness of the models, devoid of any chance factors.

Region focusing ( $q^2$ -GRS) improved the predictive power of Model 1. Predictive  $r^2$  of both models for the diverse test set molecules (Table 2) indicates that the models

Table 2  
Statistical results of conventional CoMFA and  $q^2$ -GRS models

	CoMFA Model 1 (template forcing <sup>a</sup> )	$q^2$ -GRS of Model 1 (template forcing <sup>a</sup> )	CoMFA Model 2 (docking <sup>a</sup> )	$q^2$ -GRS of Model 2 (docking <sup>a</sup> )
No. of molecules	33	33	34	34
No. of components	6	6	6	6
$q^{2b}$	0.624	0.658	0.733	0.763
PRESS <sup>c</sup>	0.985	0.940	0.804	0.756
$r^{2d}$	0.971	0.964	0.989	0.993
$r^{2e}_{\text{bootstrap}}$	0.984	0.986	0.992	0.997
$s^f$	0.274	0.305	0.160	0.127
$F^g$	145.1	116.2	418.6	666.9
$r^2_{\text{pred}}^h$	0.593	0.634	0.768	0.647
SDEP <sup>i</sup> <sub>test</sub>	0.455	0.641	0.509	0.628
Steric	0.327	0.372	0.297	0.355
Electrostatic	0.673	0.628	0.703	0.645

<sup>a</sup> Alignment method.

<sup>b</sup> Cross-validation correlation coefficient.

<sup>c</sup> Predictive sum square error.

<sup>d</sup> Correlation coefficient.

<sup>e</sup> Bootstrap  $r^2$ .

<sup>f</sup> Standard error of estimate.

<sup>g</sup>  $F$ -ratio.

<sup>h</sup> Predictive  $r^2$ .

<sup>i</sup> Standard deviation of error of prediction.



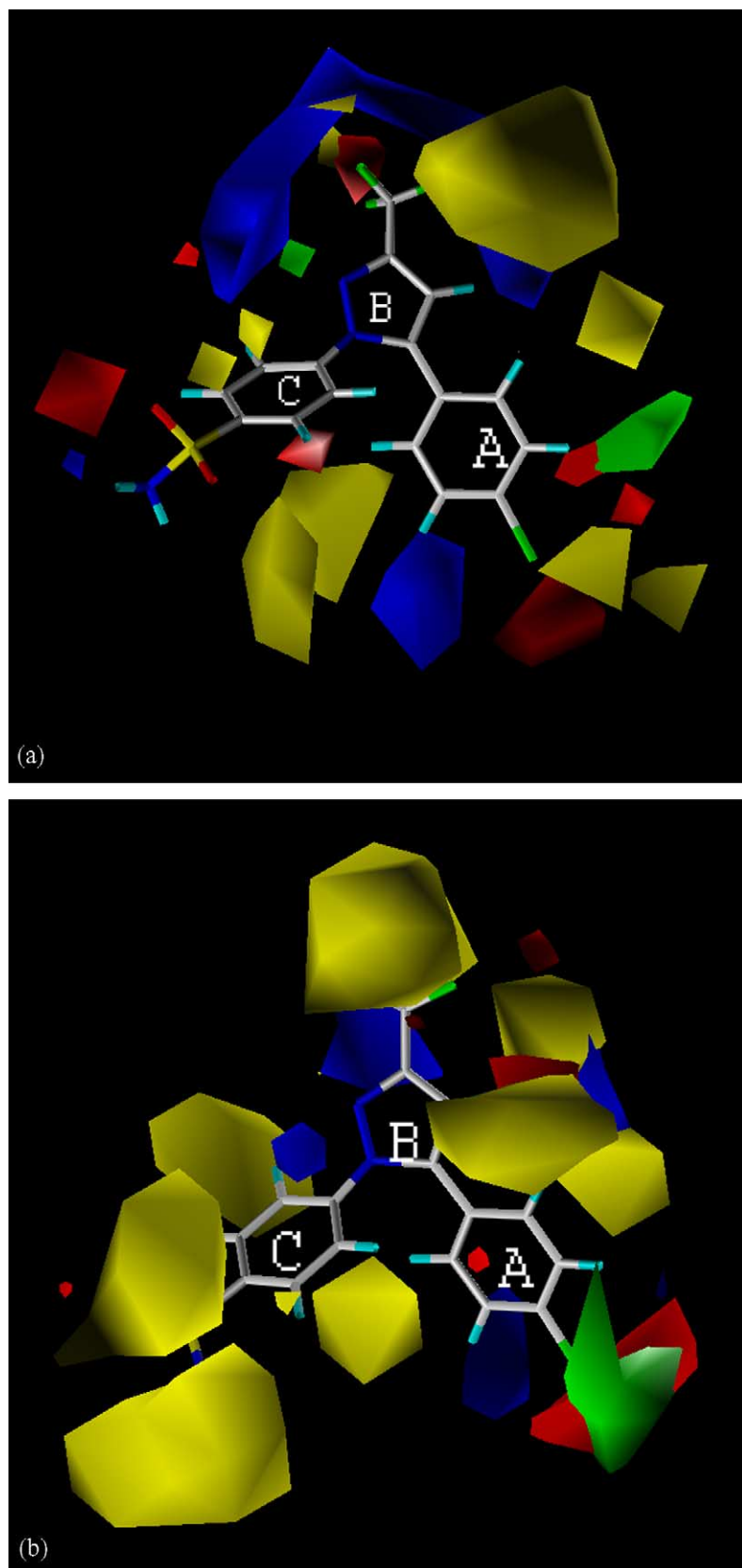


Fig. 4. CoMFA contour maps depicted around SC-558 generated from (a) Model 1 and (b) Model 2; green and yellow regions show, respectively, where steric fields are favorable or unfavorable, while red and blue regions depict where electronegative or electropositive groups are favored.



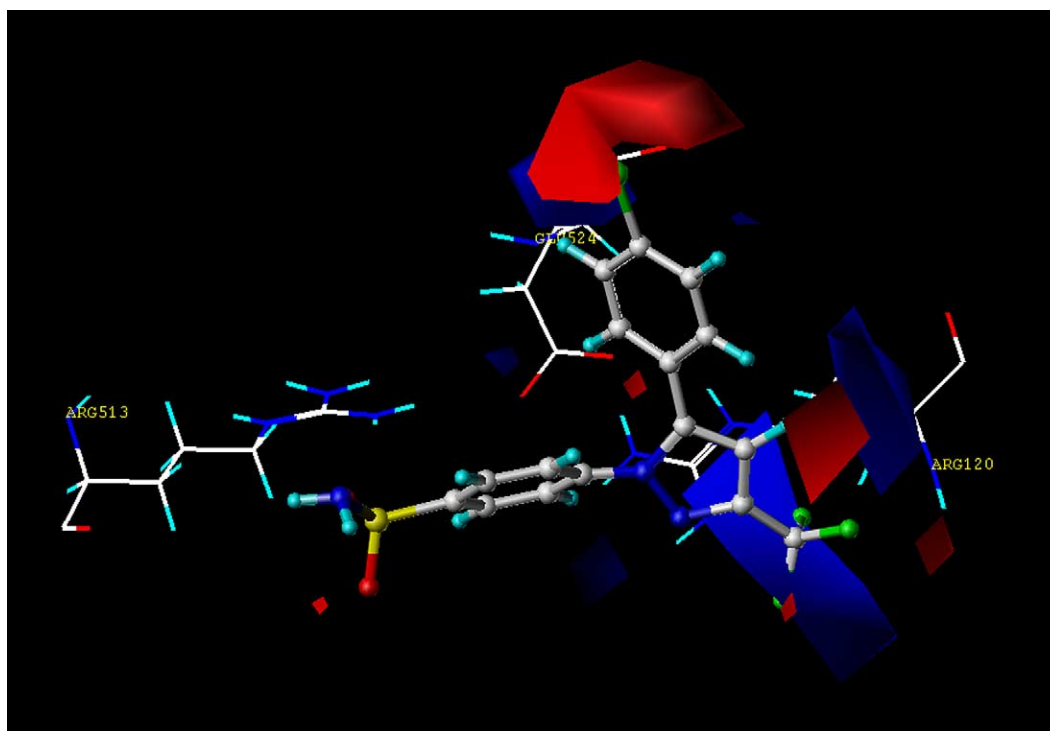


Fig. 5. The CoMFA electrostatic fields around SC-558 represented by red and blue regions correlate with the presence of Arg120, Arg513 and Glu524 residues near the binding pocket.

can predict well the activity of molecules not considered in the training set. Model 2 was unable to predict accurately the activity of only one test molecule ID **54** (Table 1A). The failure to predict the activity of this molecule, may be explained by its odd orientation in the active site. This

was the lowest energy configuration generated by the docking program *AFFINITY*. When this conformation of molecule ID **54** was replaced by its next lower energy conformation, the predictive  $r^2$  improved from 0.275 to 0.768. For Model 2,  $q^2$ -GRS actually had a detrimental

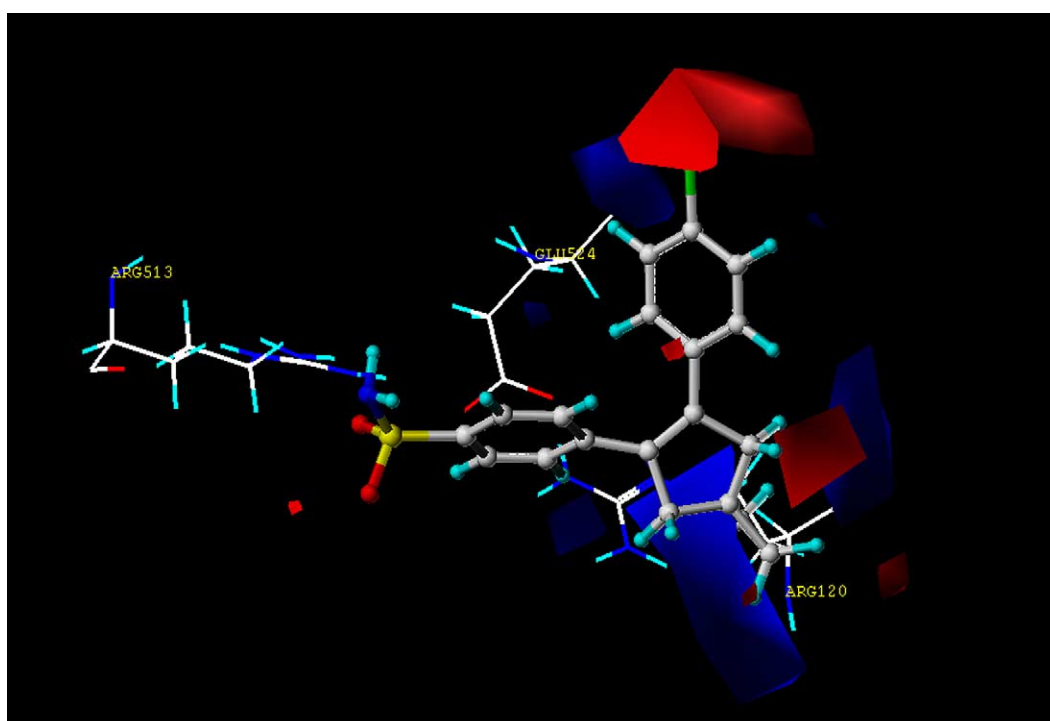


Fig. 6. Contour map of molecule 7 (the highest active molecule in the series) superimposed on the topology of the COX-2 enzyme.

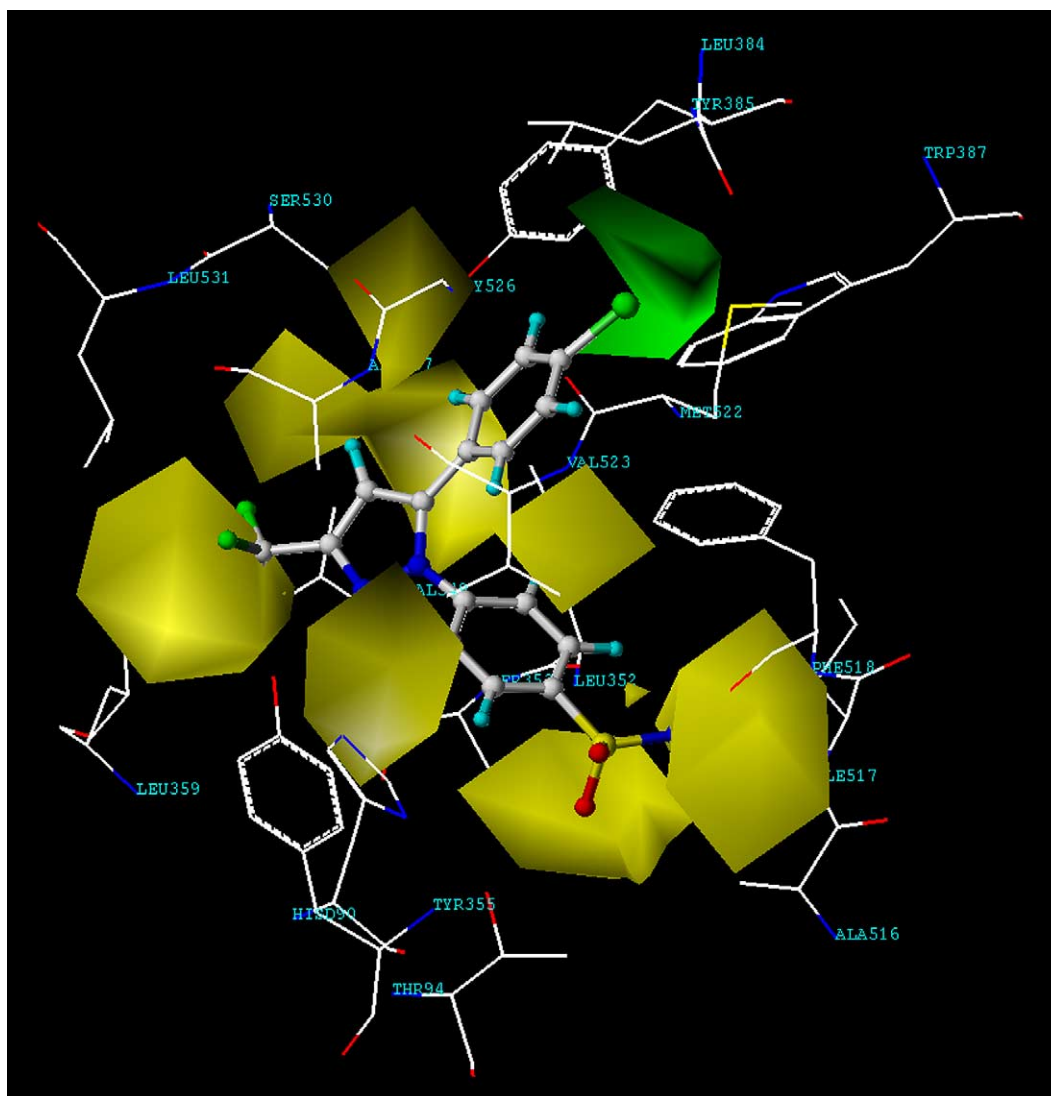


Fig. 7. Contour map of SC-558 correlated with the topology of the COX-2 enzyme. The steric fields mapped in CoMFA Model 2 nicely match with the presence of bulky residues in the active site.

effect on the results. The steric and electrostatic contributions of all models as given in Table 2 show that the electrostatic field is more important in correlating the biological activity.

#### 4.1. CoMFA contour maps

The CoMFA steric and electrostatic fields for Models 1 and 2 are presented as contour plots around SC-558 as shown in Fig. 4a and b, respectively. In the CoMFA contour plots, the green colored regions indicate space where increased steric bulk will enhance activity and yellow-colored regions are associated with decreased activity. A small yellow region present near the sulfonamide moiety of ring C (Fig. 1), suggests a cause for the lower activity of molecules containing  $-\text{SO}_2\text{CH}_3$  group. At the *ortho* position of ring A, the

presence of yellow contours suggests this is an unfavorable steric region. Another third yellow contour near the trifluoromethyl moiety suggests that bulky groups here are adverse to the inhibitory activity. A green contour near the *meta* position on ring A suggests that addition of bulk here will increase activity. It is observed in compounds with ID's **12** (Table 1C) and **17** (Table 1D); where the *m*- $\text{CH}_3$  on ring A occupies the region enclosed in the green contour.

Regions where increased positive charge is favorable for inhibitory activity are indicated in blue, while regions where increased negative charge is favorable for activity are indicated in red. A small red polyhedral near ring B indicates that an electronegative group like fluorine in molecule ID **12** (Table 1C) is beneficial for inhibitory activity. A large blue region very prominently surrounds ring B, suggesting that the inclusion of an electropositive group in this extended

area will enhance the activity. It is interesting to note that the  $\text{CF}_3$  group in particular enhances the activity, while other groups such as CN (in molecule ID 57, Table 1D) and  $\text{CH}_2\text{OH}$  (in molecule ID 38, Table 1F) have a negative effect on the activity. The red contour near the *para* position of ring A and near the  $\text{SO}_2\text{NH}_2$  group of ring C suggest placement of an electronegative group would be favorable for the activity. This can be seen for the Br substituent in molecule ID 53 (Table 1D), F substituent in molecule ID 5 (Table 1B) and Cl substituent in molecule ID 7 (Table 1B). A blue region is observed near the *meta* position of ring A where an electropositive group would enhance the activity, thus an electronegative substituent such as Cl in molecules ID 10 (Table 1C) and 34 (Table 1F) may be responsible for a decrease in the activity. A small red region is also observed near the *meta* position of ring A, in conjunction with the steric green region. Positioning an electronegative substitution here will increase the activity. This is true for molecules ID 9 (Table 1B) and 31 (Table 1F).

There is a slight difference in the contour plots generated from Models 1 and 2. Yellow contours dominate Model 2, except at the *para* position of ring A. It also shows the  $\text{CH}_3$  group of  $\text{SO}_2\text{CH}_3$  to be unfavorable as discussed above. Molecules that have F, Br and Cl at the *para* position of ring A are positioned within a red contour and have good activity. A small red contour was observed around the *ortho* position of ring A and another near the  $\text{SO}_2\text{NH}_2$  group of ring C in Model 2. The blue contour near the  $\text{CF}_3$  group of ring B is less prominent in Model 2. All in all, the electrostatic field is less dense in the contours generated from Model 2.

Receptor topology and contour map correlation. As the CoMFA Model 2 was derived by docking molecular structures into the active site, it seems interesting if one can describe the contour maps in relation to the enzyme site structure, in order to correlate [29] ligand-receptor binding and biological activity. Superimposition of the docked conformation of SC-558 over the crystal structure and subsequent positioning of contour plots over the enzyme topology is depicted in Fig. 5. The contours around the highest active molecule of the series, ID 7 (Table 1B) placed within the active site of COX-2 is depicted in Fig. 6. In both Figs. 5 and 6, the presence of charged residues such as Arg513, Glu524 and Arg120 in the vicinity of the blue and red contours is clearly visible, which resemble earlier reports [36] on the importance of electrostatic interactions in the binding to the COX-2 enzyme. Fig. 7 depicts how bulky residues in the active site can be correlated to the CoMFA steric contours. Bulky residues can be correlated with the yellow contours and are seen to fill most of the space around the molecule except at the *para* position of ring A, where a green contour is visible. This green contour correlates with the extra space available at Leu384 [37] in the COX-2 enzyme.

It is interesting to see a green region at the *para* position of ring A in Model 2, where substitution by bulky groups may enhance activity by interacting with hydrophobic residues such as Tyr385, Trp387, Ala527, Phe381 and Leu384.

Residues that participate in van der Waals contact are Val349, Ser353, Tyr355 with ring B; Phe518 with ring C; and Val523 and Ala527 with ring A. It seems that rings C and B are less amenable to additions, while around ring A there is more space for modifications.

Desiraju et al. [6] have also generated 3D-QSAR models where the 1,2-diarylimidazole structures were constrained to the conformation of SC-558 to yield a model with a  $q^2$  of 0.56 and predictive  $r^2$  of 0.79 for the test set. Similarly, taking the crystal structure of SC-558 as the template, Chavatte and coworkers [8] have developed a CoMFA model for each class of compounds with  $q^2$  ranging from 0.7 to 0.74 and with predictive  $r^2$  ranging from 0.65 to 0.74. Liu et al. [9] have used docking of 1,5-diarylpyrazoles and generated a CoMFA model with  $q^2$  of 0.635 and a predictive  $r^2$  of 0.8. In comparison to these earlier reports, our CoMFA models have comparable  $q^2$  and predictive  $r^2$  values for both the cases where ligands were aligned based on the docking strategy or by template forcing on to the SC-558 crystal structure. It must be emphasized here that our  $r^2_{\text{pred}}$  is for a large diverse test set, which is limited in each of the three studies quoted above. Overall, our model for COX-2 inhibitors is a refinement of the existing CoMFA models and exhibits enhanced predictive capabilities. Our models help to better comprehend the structure–activity relationships existent for this class of compounds and also facilitate the design of new inhibitors with good chemical diversity.

## 5. Conclusions

The main idea of using docking as a strategy for alignment was the fact that the COX-2 enzyme is known to be flexible in accommodating ligands in the binding pocket. Thus each molecule can adopt a binding pattern with some variations from that of SC-558. Improvement in the external predictivity of Model 2 by replacing one conformation with another for molecule ID 54 strongly validates our premise and suggests that the lowest energy conformation does not always satisfactorily describe binding for all molecules.

The different binding configurations generated by AFFINITY help to throw light on the different arrangement of the ligand in the binding pocket. Since the diarylheterocycles are small in size, the alignment on a template of SC-558 severely limits the space where steric and electrostatic fields are defined. In comparison to this, the structures aligned using docking protocols occupy a wider space in the binding pocket that may be more efficient at sampling significant steric and electrostatic interactions by CoMFA.

## Acknowledgement

This work was made possible by a grant from the All India Council of Technical Education (AICTE), New Delhi (F. No. 8020/RID/R&D-60/2001-02) to E. Coutinho. P.A.

Datar thanks the Amrut Mody Research Foundation for support.

## References

- [1] W. Xie, J.G. Chipman, D.L. Roberson, R.L. Erickson, D.L. Simmons, Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mRNA splicing, *Proc. Natl. Acad. Sci. U.S.A.* 88 (1991) 2692–2696.
- [2] (a) M.C. Allison, A.G. Howatson, C.J. Torrance, F.D. Lee, R.I.G. Russell, Gastrointestinal damage associated with the use of nonsteroidal antiinflammatory drugs, *N. Engl. J. Med.* 327 (1992) 749–754;  
(b) J.L. Wallace, Distribution and expression of cyclooxygenase (COX) isoenzymes, their physiological roles, and the categorization of nonsteroidal anti-inflammatory drugs (NSAIDs), *Am. J. Med.* 107 (6A) (1999) 11S–17S.
- [3] N.S. Buttar, K.K. Wang, The “Aspirin” of the new millenium: cyclooxygenase-2 inhibitors, *Mayo Clin. Proc.* 75 (2000) 1027–1038.
- [4] T.D. Penning, J.J. Talley, S.R. Bertenshaw, J.S. Carter, P.W. Collins, S. Doctor, M.J. Grant, L.F. Lee, J.W. Malecha, J.M. Miyashiro, R.S. Rogers, D.J. Rogier, S.S. Yu, G.D. Anderson, E.G. Burton, J.N. Cogburn, S.A. Gregory, C.M. Koboldt, W.E. Perkins, K. Seibert, A.M. Veenhuizen, Y.Y. Zhang, P.C. Isakson, Synthesis and biological evaluation of the 1,5-diarylpyrazole class of cyclooxygenase-2 inhibitors: identification of 4-[5-(4-methylphenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl]benzenesulfonamide (SC-58635, Celecoxib), *J. Med. Chem.* 40 (1997) 1347–1365.
- [5] P. Prasit, Z. Wang, C. Brideau, C.-C. Chan, S. Charleson, W. Cromlish, D. Eithier, J.F. Evans, A.W. Ford-Hutchinson, J.Y. Gauthier, R. Gordon, J. Guay, M. Gresser, S. Kargman, B. Kennedy, Y. Leblanc, S. Leger, J. Mancini, G.P.O. Neil, M. Ouellet, M.D. Percival, H. Perrier, D. Riendeau, I. Rodger, P. Tagari, M. Therien, P. Vickers, E. Wong, L.-J. Xu, R.N. Young, R. Zamboni, S. Boyce, N. Rupniak, M. Forrest, D. Visco, D. Patrick, The discovery of rofecoxib [MK 966, Vioxx, 4-(4'-methylsulfonylphenyl)-3-phenyl-2(5H)-furanone], an orally active cyclooxygenase-2 inhibitor, *Bioorg. Med. Chem. Lett.* 9 (1999) 1773–1778.
- [6] G.R. Desiraju, B. Gopalakrishnan, R.K.R. Jetti, D. Raveendra, J.A.R.P. Sarma, H.S. Subramanya, Three-dimensional quantitative structural activity relationship (3D-QSAR) studies of some 1,5-diarylpyrazoles: analogue-based design of selective cyclooxygenase-2 inhibitors, *Molecules* 7 (2000) 945–955.
- [7] C. Marot, P. Chavette, D. Lesieur, Comparative molecular field analysis of selective cyclooxygenase-2 (COX-2) inhibitors, *Quant. Struct. Act. Relat.* 19 (2000) 127–134.
- [8] (a) C. Marot, P. Chavatte, D. Lesieur, Comparative molecular field analysis of selective cyclooxygenase-2 (COX-2) inhibitors, *Quant. Struct. Act. Relat.* 19 (2000) 127–134;  
(b) P. Chavatte, S. Yous, C. Marot, N. Baurin, D. Lesieur, Three-dimensional quantitative structure-activity relationships of cyclooxygenase-2 (COX-2) inhibitors: a comparative molecular field analysis, *J. Med. Chem.* 44 (2001) 3223–3230.
- [9] H. Liu, X. Huang, J. Shen, X. Luo, M. Li, B. Xiong, G. Chen, J. Shen, Y. Yang, H. Jiang, K. Chen, Inhibitory mode of 1,5-diarylpyrazole derivatives against cyclooxygenase-2 and cyclooxygenase-1: molecular docking and 3D QSAR analyses, *J. Med. Chem.* 45 (2002) 4816–4827.
- [10] P. Singh, R. Kumar, Novel inhibitors of cyclooxygenase-2: the sulfones and sulfonamides of 1,2-diaryl-4,5-difluorobenzene. Analysis of quantitative structure-activity relationship, *J. Enzymol. Inhib.* 13 (1998) 409–417.
- [11] Affinity, Insight 98, Molecular Modelling Program Package, Accelrys Inc., San Diego, CA, 1997.
- [12] J.K. Gierse, S.D. Hauser, D.P. Creely, C. Koboldt, S.H. Rangwala, P.C. Isakson, K. Seibert, Expression and selective inhibition of the constitutive and inducible forms of human cyclo-oxygenase, *Biochem. J.* 305 (1995) 479–484.
- [13] I.K. Khanna, R.M. Weier, Y. Yu, P.W. Collins, J.M. Miyashiro, C.M. Koboldt, A.W. Veenhuizen, J.L. Currie, K. Seibert, P.C. Isakson, 1,2-Diarylpyrroles as potent and selective inhibitors of cyclooxygenase-2, *J. Med. Chem.* 40 (1997) 1619–1633.
- [14] I.K. Khanna, R.M. Weier, Y. Yu, X.D. Xu, F.J. Koszyk, P.W. Collins, C.M. Koboldt, A.W. Veenhuizen, W.E. Perkins, J.J. Casler, J.L. Masferrer, Y.Y. Zhang, S.A. Gregory, K. Seibert, P.C. Isakson, 1,2-Diarylimidazoles as potent, cyclooxygenase-2 selective, and orally active antiinflammatory agents, *J. Med. Chem.* 40 (1997) 1634–1647.
- [15] I.K. Khanna, Y. Yu, R.M. Huff, R.M. Weier, X. Xu, F.J. Koszyk, P.W. Collins, J.N. Cogburn, P.C. Isakson, C.M. Koboldt, J.L. Masferrer, W.E. Perkins, K. Seibert, A.W. Veenhuizen, J. Yuan, D.C. Yang, Y.Y. Zhang, Selective cyclooxygenase-2 inhibitors: heteroaryl modified 1,2-diarylimidazoles are potent, orally active antiinflammatory agents, *J. Med. Chem.* 43 (2000) 3168–3185.
- [16] D.B. Reitz, J.J. Li, M.B. Norton, E.J. Reinhart, J.T. Collins, G.D. Anderson, S.A. Gregory, C.M. Koboldt, W.E. Perkins, K. Seibert, P.C. Isakson, Selective cyclooxygenase inhibitors: novel 1,2-diarylcyclopentenes are potent and orally active COX-2 inhibitors, *J. Med. Chem.* 38 (1994) 3878–3881.
- [17] J.J. Li, G.D. Anderson, E.G. Burton, J.N. Cogburn, J.T. Collins, D.J. Garland, S.A. Gregory, H.C. Huang, P.C. Isakson, C.M. Koboldt, E.W. Logusch, M.B. Norton, W.E. Perkins, E.J. Reinhart, K. Seibert, A.W. Veenhuizen, Y. Zhang, D.B. Reitz, 1,2-Diarylcyclopentenes as selective cyclooxygenase-2 inhibitors and orally active antiinflammatory agents, *J. Med. Chem.* 38 (1995) 4570–4578.
- [18] J.J. Li, M.B. Norton, E.J. Reinhart, G.D. Anderson, S.A. Gregory, P.C. Isakson, C.M. Koboldt, J.L. Masferrer, W.E. Perkins, K. Seibert, Y. Zhang, B.S. Zweifel, D.B. Reitz, Novel terphenyls as selective cyclooxygenase-2 inhibitors and orally active antiinflammatory agents, *J. Med. Chem.* 39 (1996) 1846–1856.
- [19] H.C. Huang, J.J. Li, D.J. Garland, T.S. Chamberlain, E.J. Reinhart, R.E. Manning, K. Seibert, C.M. Koboldt, S.A. Gregory, G.D. Anderson, A.W. Veenhuizen, Y. Zhang, W.E. Perkins, E.G. Burton, J.N. Cogburn, P.C. Isakson, D.B. Reitz, Diarylspiro[2.4]heptenes as orally active, highly selective cyclooxygenase-2 inhibitors: synthesis and structure-activity relationships, *J. Med. Chem.* 39 (1996) 253–266.
- [20] J.J. Talley, D.L. Brown, J.S. Carter, M.J. Graneto, C.M. Koboldt, J.L. Masferrer, W.E. Perkins, R.S. Rogers, A.F. Shaffer, Y.Y. Zhang, B.S. Zweifel, K. Seibert, 4-[5-Methyl-3-phenylisoxazol-4-yl]-benzenesulfonamide, valdecoxib: a potent and selective inhibitor of COX-2, *J. Med. Chem.* 43 (2000) 775–777.
- [21] L.F. Lee, Pat. No. WO 9624585, *Chem. Abstr.* 125 (1996) 247622.
- [22] Y. Song, D.T. Connor, R. Doubleday, R.J. Sorenson, A.D. Sercel, P.C. Unangst, B.D. Roth, R.B. Gilbertsen, K. Chan, D.J. Schrier, A. Guglietta, D.A. Bornemeier, R.D. Dyer, Synthesis, structure-activity relationships, and in vivo evaluations of substituted di-tert-butylphenols as a novel class of potent, selective, and orally active cyclooxygenase-2 inhibitors. Part I. Thiazolone and oxazolone series, *J. Med. Chem.* 42 (1999) 1151–1160.
- [23] Y. Song, D.T. Connor, A.D. Sercel, R.J. Sorenson, R. Doubleday, P.C. Unangst, B.D. Roth, R.B. Gilbertsen, K. Chan, D.J. Schrier, A. Guglietta, D.A. Bornemeier, R.D. Dyer, Synthesis, structure-activity relationships, and in vivo evaluations of substituted di-tert-butylphenols as a novel class of potent, selective, and orally active cyclooxygenase-2 inhibitors. 2.1,3,4- and 1,2,4-thiadiazole series, *J. Med. Chem.* 42 (1999) 1161–1169.
- [24] Sybyl, Version 6.7, 2000, Molecular Modelling Software, Tripos Associates Inc., 1699 South Hanley Road, St. Louis, MO 63144.
- [25] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E. Ferrin, A geometric approach to macromolecule-ligand interactions, *J. Mol. Biol.* 161 (1982) 269–288.

- [26] P. Dauber-Osguthorpe, V.A. Roberts, D.J. Osguthorpe, J. Wolff, M. Genest, A.T. Hagler, Structure and energetics of ligand binding to proteins: *E. coli* dihydrofolate reductase-trimethoprim, a drug-receptor system., *Proteins: Structure Function Genetics* 4 (1988) 31–37.
- [27] R.G. Kurumbail, A.M. Stevens, J.K. Gierse, J.J. McDonald, R.A. Stegeman, J.Y. Pak, D. Gildehaus, J.M. Miyashiro, T.D. Penning, K. Seibert, P.C. Isakson, W.C. Stallings, Structural basis for selective inhibition of cyclooxygenase-2 by antiinflammatory agents, *Nature* 384 (1996) 644–648.
- [28] H.Q. Ding III, N. Karasawa, W.A. Goddard, Atomic level simulations on a million particles: the cell multipole method for Coulomb and London nonbond interactions, *J. Chem. Phys.* 97 (1992) 4309–4315.
- [29] R.D. Cramer III, D.E. Patterson, J.D. Bunce, Comparative molecular fields analysis (CoMFA) 1. effect of shape on binding of steroids to carrier proteins, *J. Am. Chem. Soc.* 110 (1988) 5959–5967.
- [30] S. Wold, E. Johansson, M. Cocchi, in: H. Kubinyi (Ed.), *3D-QSAR in Drug Design: Theory Methods and Applications*, ESCOM Lieden, The Netherlands, 1993, pp. 523–550.
- [31] P. Datar, P. Desai, E. Coutinho, K. Iyer, CoMFA and CoMSIA studies of angiotensin (AT1) receptor antagonists, *J. Mol. Model.* 8 (2002) 290–301.
- [32] S.J. Cho, A. Tropsha, Cross-validated  $R^2$ -guided region selection for comparative molecular field analysis: a simple method to achieve consistent results, *J. Med. Chem.* 38 (1995) 1060–1066.
- [33] C. Leger, D.N. Politis, J.P. Romano, Bootstrap technology and applications, *Technometrics* 34 (1992) 378–399.
- [34] M. Clark, R.D. Cramer III, The probability of chance correlation using partial least squares (PLS), *Quant. Struct. Act. Relat.* 12 (1993) 137–145.
- [35] A. Tropsha, S.J. Cho, in: H. Kubinyi, G. Folkers, Y.C. Martin (Eds.), *Cross-Validated  $R^2$  Guided Region Selection for CoMFA Studies. 3D QSAR Methodology CoMFA and Related Approaches*, vol. III, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 57–69.
- [36] O.-Y. So, E. Scarafia, A.Y. Mak, O.H. Callan, D.C. Swinney, The dynamics of prostaglandin H synthases. Studies with prostaglandin H synthase 2 Y355F unmask mechanisms of time-dependent inhibition and allosteric activation, *J. Biol. Chem.* 273 (1998) 5801–5807.
- [37] M.F. Browner, in: J. Vane, J. Botting (Eds.), *Selective COX-2 Inhibitors. Pharmacology, Clinical Effects and Therapeutic Potential*, Kluwer Academic Publishers, Dordrecht, 1998, pp. 19–26.