Bioorganic & Medicinal Chemistry Letters 21 (2011) 456–462

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

CoMFA and CoMSIA studies on 5-hydroxyindole-3-carboxylate derivatives as 5-lipoxygenase inhibitors: Generation of homology model and docking studies

P. Aparoy, G. K. Suresh, K. Kumar Reddy, P. Reddanna *

School of Life Sciences, University of Hyderabad, Hyderabad 500 046, India

ARTICLE INFO

Article history:
Received 2 May 2010
Revised 2 October 2010
Accepted 23 October 2010
Available online 30 October 2010

Keywords:
5-LOX
3D-QSAR
CoMFA
CoMSIA
Cross validation
Homology modeling
Docking

ABSTRACT

In this study, comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed on a series of 2-substituted 5-hydroxyindole-3-carboxylate derivatives as potent 5-LOX inhibitors with IC_{50} values ranging from 0.031 to 13.4 μM. Two datasets of same molecules were prepared with two different partial atomic charges; one with Gasteiger–Huckel and another with the ESPFIT charges obtained from the GAUSSIAN package. CoMFA and CoMSIA models were generated for both the datasets and the results were analysed. With regard to the non-cross validated r^2 values (r^2_{ncv}) and cross-validated q^2 values (q^2_{cv}) of the resulting QSAR models, the dataset with ESPFIT charges yielded higher values; hence it was further used in the study. The CoMFA and CoMSIA models have been further validated for their stability and robustness using group validation and bootstrapping techniques and for their predictive abilities using an external test set of ten compounds. The predictive power of the CoMSIA model was higher than the CoMFA model, the high predictive r^2 values of the test set reveals that the models prove to be useful tools for activity prediction of newly designed 5-LOX inhibitors. The ESPFIT-derived charges yielded better models than those based on charges calculated from Gasteiger–Huckel charges. We generated a homology model for human 5-LOX and identified the key residues at the binding site. The 3D-QSAR models were compared with the interactions at the active site to further elucidate the accuracy of the models. The data generated from 3D-QSAR study was used to design potential 5-LOX inhibitors.

Lipoxygenases (LOXs) (linoleate: oxygen oxido reductase, EC 1.13.11.12) are a group of closely related non-heme iron containing dioxygenases. These enzymes catalyze the addition of molecular oxygen into Poly Unsaturated Fatty Acids (PUFAs) containing cis, cis 1–4 pentadiene structures to give their hydroperoxy derivatives.[1] LOXs are classified according to their positional specificity of arachidonate oxygenation into 5-, 8-, 9-, 11-, 12- and 15-LOXs.[2] One of the LOX pathways of arachidonic acid metabolism, the 5-LOX pathway, is the source of potent pro-inflammatory mediators.[3] LOX metabolites are potent physiological effectors in a variety of cellular responses, associated with normal host defense and inflammation. In particular, leukotrienes (LTs), the mediators of allergy and asthma, are produced through the 5-LOX pathway. Products of the 5-LOX pathway are thus important mediators of inflammation. Inhibitors of the 5-LOX pathway, therefore, have therapeutic potential in a variety of inflammatory and allergic diseases. 5-LOX plays a key role in gastroesophageal reflux disease (GERD), rheumatoid arthritis and Crohn's disease.[4] High expression of 5-LOX was found in prostate, lung and other cancer cell lines.[5–8] Recently it has been shown that 5-LOX (ALOX5) is critical regulator for leukemia cancer stem cells (LSCS) in chronic myeloid leukemia (CML).[9] Currently an emerging strategy of therapeutic value consists of creating molecules with specific 5-LOX inhibition activity.

In our earlier studies, a theoretical 3D model of potato 5-LOX was elaborated by homology modeling.[10] The 5-LOX active site was then characterized from a structural point of view and used to study the docking of selected inhibitors. This shed new light on the binding features of the enzyme. In a more recent study, chemical feature based pharmacophore modeling of inhibitors of 5-LOX have been carried out by using HypoGen module within Catalyst program package.[11] The fact that 5-LOX inhibitors can be successfully identified by employing pharmacophore based virtual screening explains its usefulness in predicting activities of large datasets of molecules. Thus, our earlier studies provided homology and pharmacophore models which help in designing the novel 5-LOX inhibitors.

In this study, we have performed three dimensional quantitative structure–activity relationship (3D-QSAR), using the comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) techniques.[12,13] The study

* Corresponding author. Tel.: +91 40 23134542; fax: +91 40 23010745.
E-mail addresses: prsl@uohyd.ernet.in, preddanna@yahoo.com (P. Reddanna).

0960-894X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2010.10.119

has been carried out on 2-substituted 5-hydroxyindole-3-carboxylate derivatives to determine the influence of steric, electrostatic and hydrophobic fields of these compounds on their 5-LOX inhibitory activity. Furthermore, these fields have been mapped onto binding pocket of human 5-LOX model that has also been generated in this study to better understand the interactions.

Fourty compounds (Table 1) with IC_{50}s ranging from 0.031 to 13.4 μM for 5-LOX were selected from the literature.[14] The IC_{50} values were converted into corresponding pIC_{50} values by the formula in Eq. (1). The calculated pIC_{50} values ranged from 4.87 to 7.50.

$$pIC_{50} = -\log IC_{50} \quad (1)$$

The initial set of compounds has been randomly divided into training set (30 compounds) and test set (10 compounds). The structures of the molecules in training set and test set are included in Supplementary data. All molecular modeling calculations were performed using SYBYL program package version 8.0 (Tripos Associates Inc.) on a Linux operating system.[15] Molecular building was done with a molecular sketch program.

Two datasets of the compounds with different partial charges were prepared; first dataset with charges calculated by the Gasteiger-Hückel method and the second dataset with ESPFIT charges.

For the first dataset the molecular geometry of each compound was first minimized using a standard Tripos molecular mechanics force field with a 0.001 kcal/mol Å energy gradient convergence criterion, and their charges were calculated by the Gasteiger-Hückel method. Partial atomic charges were assigned to each atom and then energy minimization of each molecule was performed by using steepest descent method 500 steps and Conjugate gradient method until molecule converged using Tripos standard force field with a distance-dependent dielectric function.

For the generation of second dataset GAUSSIAN 03 package was used. All the molecules were optimized using HF/6-31G* basis set and the partial atomic charges for each of them were obtained with ESP fitting (HF/6-31G* OPT ESP).

A proper alignment of the structures is critical for obtaining valid 3D-QSAR models.[16,17] Further, to obtain a reliable and consistent alignment, the lowest energy conformation of the most active molecule, compound **24** (Fig. 1) was used as the template. The atoms used for alignment were the common substructure of each molecule as shown in the Figure 1B. An automatic alignment method was carried out by using Database alignment. The conformations of all aligned molecules of the training set are shown in Figure 1C.

The standard CoMFA procedure implemented in SYBYL8.0 was used in the present study. A 3D cubic lattice with a grid spacing of 2 Å was created automatically by the program to encompass all the aligned ligands. CoMFA[12] steric and electrostatic fields were calculated using Tripos force field taking a sp3 carbon probe atom with a Van der Waals radius of 1.52 Å and a charge of +1.0 to generate steric (Lennard–Jones 6–12 potential) field energies and electrostatic (Coulombic potential) fields with a distance-dependent dielectric constant at each lattice point. Steric and electrostatic fields generated were scaled by the CoMFA-Standard method in SYBYL with default cut-off energy of 30 kcal/mol. The minimum column filtering was set to 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points whose energy variation was below this threshold. CoMSIA similarity indices were derived according to Klebe et al.[13]

The 3D-QSAR models of CoMFA and CoMSIA descriptors were derived using PLS regression method[18] as implemented in the SYBYL package. The cross-validated coefficient, q^2, was calculated using Eq. (2).

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

Where $Y_{predicted}$, $Y_{observed}$, and Y_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively. $\sum(Y_{predicted} - Y_{observed})^2$ is the predictive sum of squares (PRESS).

Table 1
Observed/experimental and CoMFA/CoMSIA predicted 5-LOX inhibitory activity (IC_{50}) of 2-substituted 5-hydroxyindole-3-carboxylate derivatives in the training set

S. No.	Exptl. IC_{50} (μM)	pIC_{50}	CoMFA predicted activity (error)		CoMSIA predicted activity (error)	
			Gasteiger–Huckel	ESPFIT	Gasteiger–Huckel	ESPFIT
1	8.1	5.0915	5.075 (0.1158)	5.147 (0.0555)	5.074 (0.0175)	5.089 (0.0025)
2	9.8	5.0088	4.893 (0.1158)	4.891 (0.1178)	4.953 (0.0558)	4.941 (0.0678)
3	13.4	4.8729	4.957 (−0.0841)	4.878 (0.0051)	4.976 (0.1031)	5.026 (0.1531)
4	8.3	5.0809	5.142 (−0.0611)	5.144 (0.0631)	5.142 (0.0611)	5.109 (0.0281)
5	6.9	5.1612	5.138 (0.0232)	5.095 (0.0662)	5.121 (0.0402)	5.158 (0.0032)
6	3.5	5.4559	5.367 (0.0889)	5.53 (−0.0741)	5.28 (0.1759)	5.416 (0.0399)
7	0.3	6.5229	6.061 (0.4619)	6.112 (0.4109)	6.054 (0.4689)	5.95 (0.5729)
8	2	5.699	5.789 (−0.09)	5.873 (−0.174)	5.778 (−0.079)	5.833 (−0.134)
9	1.6	5.7959	5.893 (−0.0971)	5.879 (0.0831)	5.871 (0.0751)	5.785 (0.0109)
10	1.2	5.9208	5.847 (0.0738)	5.81 (0.1108)	5.946 (0.0252)	5.862 (0.0588)
11	4.8	5.3188	5.769 (−0.4502)	5.734 (0.4152)	5.819 (0.5002)	5.724 (0.4052)
12	0.7	6.1549	6.178 (−0.0231)	6.164 (0.0091)	6.155 (0.0001)	6.226 (0.0711)
13	7.3	5.1367	5.11 (0.0267)	4.975 (0.1617)	5.143 (0.0063)	5.075 (0.0617)
14	0.24	6.6198	7.093 (−0.4732)	7.025 (0.4052)	7.171 (0.5512)	7.118 (0.4982)
15	0.086	7.0655	6.993 (0.0725)	7.033 (0.0325)	6.872 (0.1935)	7.029 (0.0365)
16	0.097	7.0132	7.014 (−0.0008)	7.105 (0.0918)	6.985 (0.0282)	6.978 (0.0352)
17	0.084	7.0757	7.049 (0.0267)	6.975 (0.1007)	7.05 (0.0257)	7.042 (0.0337)
18	0.096	7.0177	7.052 (−0.0343)	7.017 (0.0007)	7.019 (0.0013)	7.013 (0.0047)
19	0.098	7.0088	7.039 (−0.0302)	7.004 (0.0048)	7.081 (0.0722)	7.058 (0.0492)
20	0.13	6.8861	6.9 (−0.0139)	6.887 (0.0009)	6.893 (0.0069)	6.931 (0.0449)
21	0.15	6.8239	6.796 (0.0279)	6.771 (0.0529)	6.822 (0.0019)	6.845 (0.0211)
22	0.045	7.3468	6.97 (0.3768)	7.04 (0.3068)	6.913 (0.4338)	6.924 (0.4228)
23	0.067	7.1739	7.136 (0.0379)	7.184 (0.0101)	7.149 (0.0249)	7.137 (0.0369)
24	0.031	7.5086	7.5 (0.0086)	7.451 (0.0576)	7.502 (0.0066)	7.551 (0.0424)
25	0.049	7.3098	7.347 (−0.0372)	7.336 (0.0262)	7.313 (0.0032)	7.257 (0.0528)
26	0.13	6.8861	6.816 (0.0701)	6.843 (0.0431)	6.869 (0.0171)	6.849 (0.0371)
27	0.65	6.1871	6.242 (−0.0549)	6.212 (0.0249)	6.188 (0.0009)	6.179 (0.0081)
28	0.11	6.9586	6.979 (−0.0204)	6.923 (0.0356)	6.954 (0.0046)	6.853 (0.1056)
29	0.17	6.7696	6.719 (0.0506)	6.812 (0.0424)	6.786 (0.0164)	6.877 (0.1074)
30	1.2	5.9208	5.928 (−0.0072)	5.943 (0.0222)	5.911 (0.0098)	5.957 (0.0362)

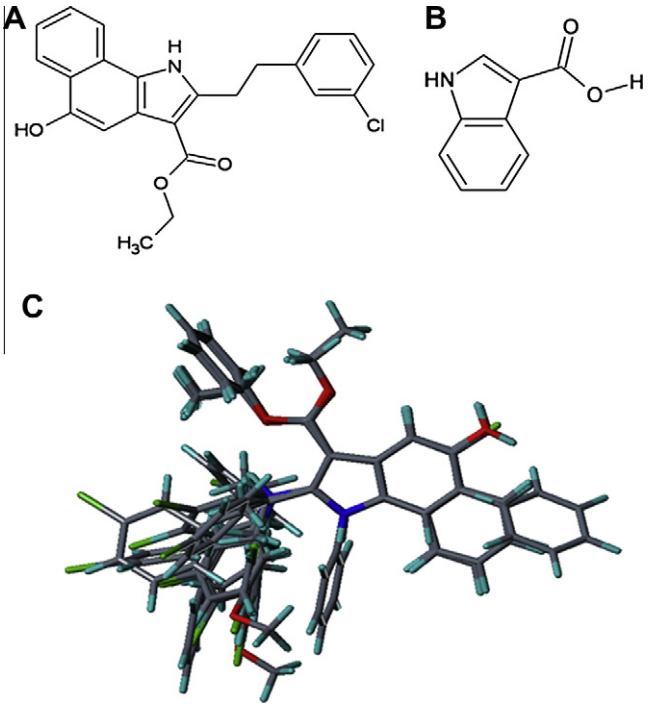

Figure 1. (A) The most active molecule in training set (compound **24**, IC_{50} 0.031 μM). (B) Core of the dataset molecules. (C) Alignment of the compounds in training set.

PLS analysis was carried out for both the datasets and the results were analyzed. The LOO cross-validated correlation coefficient q^2 was obtained with an optimal number of components (N) of six (Table 2). A value of 1.0 for q^2 or r^2 corresponds to perfect predictions, while a value of 0.0 implies that there is no model at all. The q^2 and r^2 are good indicators of the accuracy of the model and value greater than 0.5 is said to be reasonable for predicting. The non-cross validated correlation coefficient r^2 obtained for both the datasets was very close to 1.0 as shown in Table 2. The optimum number of components (ONC) which corresponds to the lowest standard error of prediction (SEP) were used to generate the final PLS regression models. To test the robustness and stability of the model the boot strapping analysis for 100 runs and group cross-validation for five groups were performed. Cross-validation,[19] leave-one-out were carried out and confirmed by the average value for 50 runs from each cross validation. CoMFA and CoMSIA contour maps were generated by interpolation of the pair wise products between the PLS coefficients and the standard deviations of the corresponding CoMFA or CoMSIA descriptor values.

Among the two datasets with different partial atomic charges, the dataset with ESPFIT charges yielded higher non-cross validated r^2 values (r^2_{ncv}) and cross-validated q^2 values (q^2_{cv}). The PLS analysis clearly shows that the partial atomic charges derived by ESPFIT charges produced better CoMFA and CoMSIA models than the Gasteiger–Huckel charges. Hence, the 3D-QSAR models developed using the ESPFIT charges has been further used in the study.

The statistical parameters associated with the CoMFA model are listed in Table 2. The steric field descriptors explain 53% of variance while the electrostatic descriptors explain 47% of the variance. To further evaluate the robustness and statistical confidence of the obtained models, bootstrap analysis for 100 runs was performed and cross validation analysis was applied to the training set compounds to investigate the stability of the CoMFA model. This yielded r^2 bootstrap of 0.985 with standard error value of 0.113 and q^2 cross validated 0.738.

Table 2
Statistical parameters of the CoMFA and CoMSIA models developed using the two datasets; one dataset with Gasteiger–Huckel charges and other with HF/6-31G* ESPFIT charges

PLS statistics	Gasteiger–Huckel CoMFA Model	HF/6-31G* CoMFA model	Gasteiger–Huckel CoMSIA model	HF/6-31G* CoMSIA model
Training set (30)				
R^2_{ncv} [a]	0.958	0.964	0.948	0.951
SEE[b]	0.194	0.179	0.216	0.211
F_{test} [c]	88.42	103.91	70.32	74.08
R^2_{cv} [d]	0.679	0.738	0.673	0.711
R^2_{pred} [e]	0.639	0.661	0.726	0.713
PLS components[f]	6	6	6	6
Contribution				
Steric	0.678	0.530	0.110	0.099
Electrostatic	0.322	0.470	0.291	0.449
Hydrophobic			0.324	0.257
Acceptor			0.134	0.139
Donor			0.141	0.055
R^2_{boot} [g]	0.980 (0.015)	0.985 (0.013)	0.978 (0.131)	0.977 (0.023)
SEE_{boot} [h]	0.131 (0.117)	0.113 (0.098)	0.133 (0.023)	0.140 (0.137)
R^2_{5cv} [i]	0.681	0.730	0.649	0.728

Values in parentheses are standard deviations.
[a] Non-cross validated correlation coefficient.
[b] Standard error of estimate.
[c] Ratio of r^2_{ncv} explained to unexplained = $r^2_{ncv}/(1 - r^2_{ncv})$.
[d] Cross validated correlation coefficient after the leave-one-out procedure.
[e] Predicted correlation coefficient for the test set of compounds.
[f] Optimal number of principal components.
[g] Average of correlation coefficient for 100 samplings using the bootstrapped procedure.
[h] Average Standard error of estimate for 100 samplings using the bootstrapped procedure.
[i] Cross-validated correlation coefficient using the five cross-validation groups.

The predicted activities ($-\log IC_{50}$) for the inhibitors verses their experimental values ($-\log IC_{50}$) and their residuals are listed in Table 1. The activity of the inhibitors was predicted better by the CoMFA model from ESPFIT charges than that of Gasteiger–Huckel charges. The CoMSIA models of both the datasets predicted the activity of inhibitors quite similarly. The ESPFIT-derived charges yielded better models than those based on charges calculated from Gasteiger–Huckel charges. However the simpler Gasteiger–Huckel charges also gave a reasonable good model. The results suggest that it is more important to have refined charge sets for better model generation.

The CoMSIA analysis revealed a cross validated q^2 of 0.711 for six components and a non cross validated r^2 of 0.951, F value of 74.08 and standard error of estimate SEE of 0.216 (Table 2). Bootstrap analysis for 100 runs was then carried out for further validation of the model and obtained r^2 bootstrap of 0.978 and standard error value SEE of 0.140. Cross validation analysis was applied to investigate the stability of the model and obtained r^2 cross validated at 0.738. All the statistical parameters of CoMFA and CoMSIA obtained show that the models generated are reasonable. Further, the accuracy of these models was elucidated using an external test set of ten compounds. The activity of these compounds (compounds **31**–**40**) was predicted by the CoMFA and CoMSIA models. The predicted results of the test set were listed in the Table 3. The plot showing correlation between the experimental and the predicted activities of CoMFA and CoMSIA are shown in the Figure 2. The predictive correlation coefficient r^2_{pred}[20] was based only on molecules not included in the training set and is computed using the following Eq. (3).

Table 3
Observed/experimental and CoMFA/CoMSIA predicted 5-LOX inhibitory activity (IC_{50}) of 2-substituted 5-hydroxyindole-3-carboxylate derivatives in the test set

S. No.	Exptl. IC_{50} (μM)	pIC_{50}	CoMFA predicted activity(error)		CoMSIA predicted activity (error)	
			Gasteiger–Huckel	ESPFIT	Gasteiger–Huckel	ESPFIT
1	0.3	6.5229	5.515 (1.0079)	5.493 (1.029)	5.979 (0.5439)	5.813 (0.709)
2	10.2	4.991	5.003 (−0.0116)	5.152 (−0.161)	4.926 (0.0654)	5.129 (−0.138)
3	3.4	5.4685	5.257 (0.2115)	5.082 (0.386)	4.888 (0.5805)	4.923 (0.545)
4	30	4.5229	5.309 (−0.7861)	5.092 (−0.57)	5.076 (−0.5531)	5.057 (−0.535)
5	1.7	5.7696	6.06 (−0.2904)	5.965 (−0.196)	6.131 (−0.3614)	6.037 (−0.268)
6	2.1	5.6778	5.939 (−0.2612)	5.89 (−0.213)	6.052 (−0.3742)	5.902 (−0.225)
7	0.14	6.8539	6.987 (−0.1331)	7.051 (−0.198)	6.878 (−0.0241)	6.994 (−0.141)
8	0.15	6.8239	6.965 (−0.1411)	6.293 (0.530)	6.861 (−0.0371)	6.292 (0.531)
9	0.25	6.6021	7.195 (−0.5929)	6.926 (−0.324)	7.149 (−0.5469)	7.065 (−0.463)
10	0.33	6.4815	5.952 (0.5295)	5.994 (0.487)	5.861 (0.6205)	5.993 (0.488)

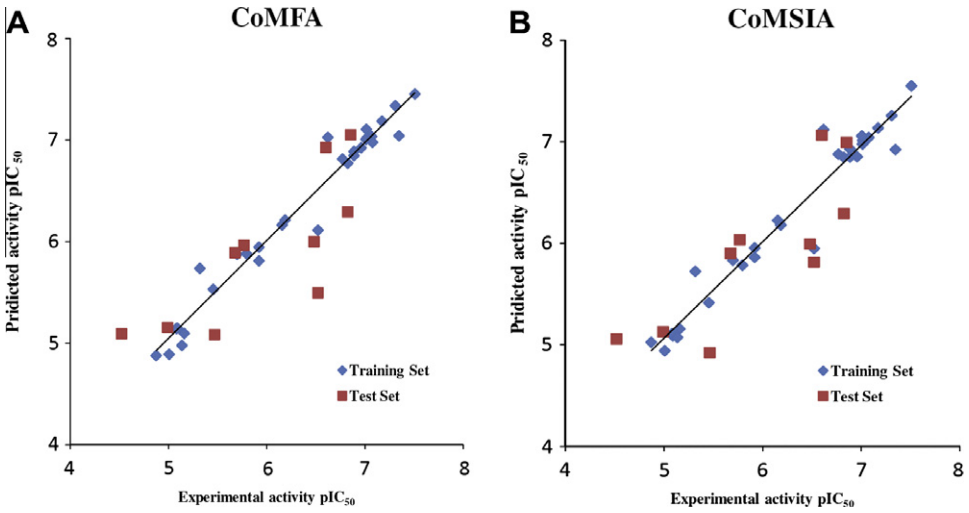

Figure 2. Plot showing experimental activities versus predicted activities (ESPFIT dataset) for the (A) CoMFA model and (B) CoMSIA model. The molecules in training set are shown in blue and molecules in test set in red.

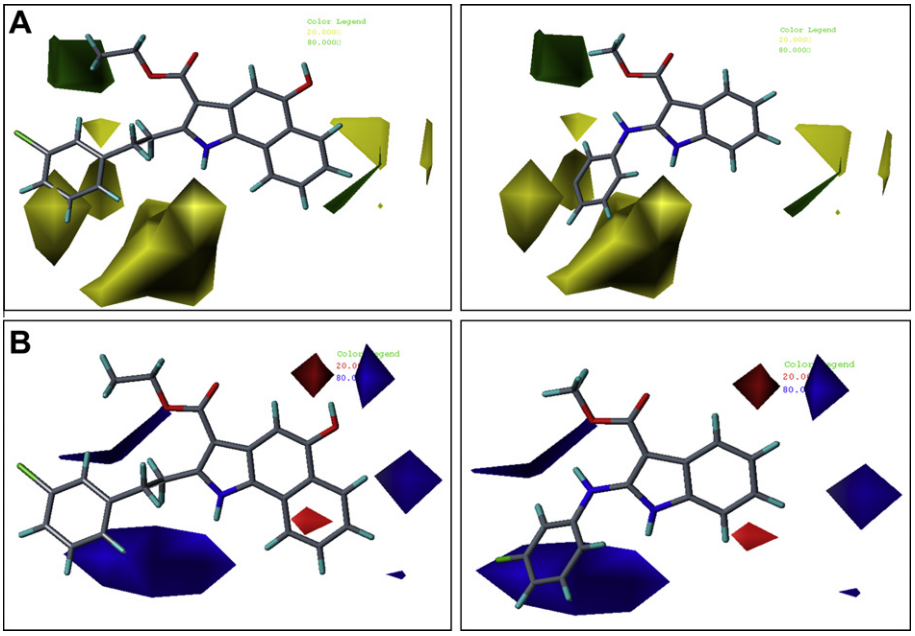

Figure 3. PLS contours from 3D-QSAR models for 5-LOX inhibitors. (A) CoMFA steric contour map based on the 30 compound training set: the green (yellow) regions show where steric groups increase (decrease) the 5-LOX inhibitory activity. (B) CoMFA electrostatic contour map based on the 30 compound training set: the red (blue) regions show where electronegative groups increase (decrease) the 5-LOX inhibitory activity. The most active member of the series, compound **24** is shown in the left and most inactive, compound **3** is shown in the right.

$$r^2_{pred} = \frac{SD - PRESS}{SD} \quad (3)$$

Where SD is the sum of the squared deviations between the biological activity of molecules in the test set and the mean biological activity of the training set molecules and PRESS is the sum of the squared deviations between predicted and actual activity values for every molecule in the test set.

The Figure 2 shows that the predicted activity values of the test set compounds are in good agreement with the experimental values in a tolerable error range, with r^2_{pred} of 0.661 and 0.713 for CoMFA and CoMSIA models, respectively. The correlation between CoMFA and CoMSIA predicted and the experimental activities of the test set compounds are shown in Table 3. Testing results indicate that the CoMFA and CoMSIA models of both the datasets could be reliably used to design novel and more potent 5-LOX inhibitors. As the dataset with ESPFIT charges yielded higher statistical parameters and little better prediction than the simple Gasteiger–Huckel charges it has been further used in the study. The contours obtained from 3D-QSAR models have been compared with the interactions at the active site of 5-LOX and used for design of more potential inhibitors.

The contour maps of 3D-QSAR CoMFA and CoMSIA provides the CoMFA steric, electrostatic fields and CoMSIA steric, electrostatic, hydrophobic, hydrogen acceptor, hydrogen donor field maps. Contour maps provide information on factors affecting the activity of the study compounds. This is particularly important when increasing or reducing the activity of a compound by changing its molecular structural features contributing to the interaction between the ligand and the active site region of a receptor. The field energies at each lattice point were calculated as the scalar results of the coefficient, and the standard deviation associated with a particular column of the data table ('S.D.*coeff') was plotted as the percentage of the contribution to the CoMFA and CoMSIA equations. In the present study of 5-LOX inhibitors the CoMFA steric field effect is 0.53 and that of electrostatic effect is 0.47. Contour maps of CoMFA and CoMSIA are described in Figures 3 and 4, respectively.

Further to understand and expound the interactions, 3D structure of 5-LOX was generated and the contours (Figs. 3 and 4) obtained were compared with the active site of 5-LOX. The amino acid sequence of human 5-LOX (Accession id: NP_000689.1) was obtained from the databank in the National center for Biotechnology Information, NCBI (www.ncbi.nlm.nih.gov). The BLAST program against Protein Data Bank (PDB)[21] available at NCBI was used to select template structures for homology modeling of 5-LOX. The crystal structure of rabbit reticulocyte 15-LOX (PDB code: 1LOX[22]) was obtained from Protein Data Bank database and it shares 38% identity with 5-LOX. Human 5-LOX was modeled against rabbit reticulocyte 15-LOX using the Discovery Studio 2.1.[23] 10 models were generated using Homology model build option in Discovery studio 2.1. Of these, the model with the least root mean square deviation (RMSD) with the template was selected. Their RMSD value was 0.39 Å. The homology model was very similar to the structure of 15-LOX. Detailed methodology used in homology modeling is provided in Supplementary data. The results of Profile-3D[24] and Procheck[25] proved that the generated model was good. The iron binding site was verified using ANAMBS.[26] Metal binding site cut-off distance was defined as 4 Å and the results showed that the iron binding site is properly built. The results show that the iron binding site in the model built is in accordance with the crystal structures (Table 4).

After the final model is generated the binding site was characterized as the amino acids within 10 Å of the ligand in the model. The binding site of 5-LOX was analyzed and was in agreement with the recent model of 5-LOX reported by Charlier et al. (2006)[27] The amino acids surrounding the ligand in the model generated are tabulated (Table 1 of Supplementary data). The knowledge on the orientation of these active site residues is extremely useful in designing a potential inhibitor of 5-LOX.

GOLD (Genetic Optimization of Ligand Docking), a docking program based on genetic algorithm[28] was used to dock the best and the least active molecules in the dataset. The active site was

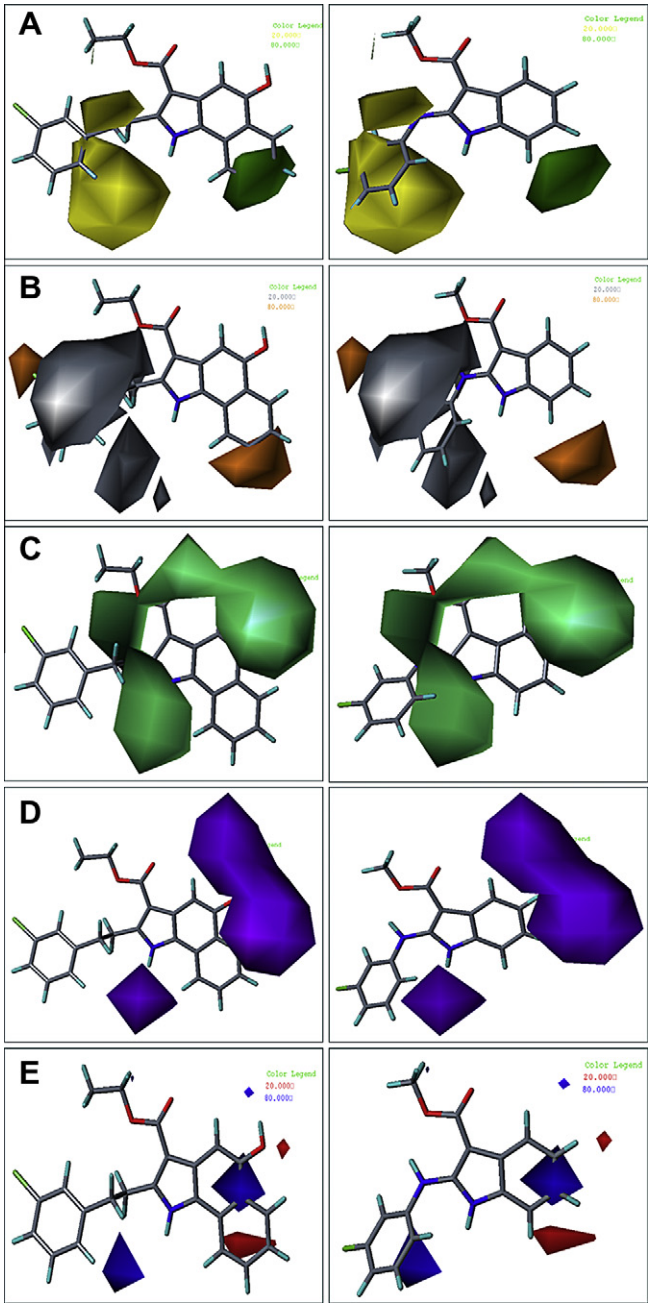

Figure 4. PLS contours from 3D-QSAR models for 5-LOX inhibitors. (A) CoMSIA steric contour: the green (yellow) regions show where steric groups increase (decrease) the 5-LOX inhibitory activity. (B) CoMSIA hydrophobic contour map: the orange (white) regions show where hydrophobic groups increase (decrease) the 5-LOX inhibitory activity. (C) CoMSIA hydrogen bond acceptor contour map: the cyan (blue-green) regions show where hydrogen bond acceptor groups increase (decrease) the 5-LOX inhibitory activity. (D) CoMSIA hydrogen bond donor contour map: the purple regions show where hydrogen bond donor groups decrease the 5-LOX inhibitory activity. (E) CoMSIA electrostatic contour map based on the 30 compound training set: the red (blue) regions show where electronegative groups increase (decrease) the 5-LOX inhibitory activity. The most active member of the series, compound **24** is shown in the left and most inactive, compound **3** is shown in the right.

Table 4
Comparison of iron binding site in 5-LOX model and the template, 1LOX

Iron binding site property calculated using ANAMBS	5-LOX model			15-LOX crystal structure		
No. of atoms	16			16		
Coordinating residues	NE2	HIS	367	NE2	HIS	361
	NE2	HIS	372	NE2	HIS	366
	NE2	HIS	550	NE2	HIS	541
	ND2	ASN	554	ND1	HIS	545
	OXT	ILE	673	OXT	ILE	663
Hydrophobicity within the region	−8.6			−8.3		

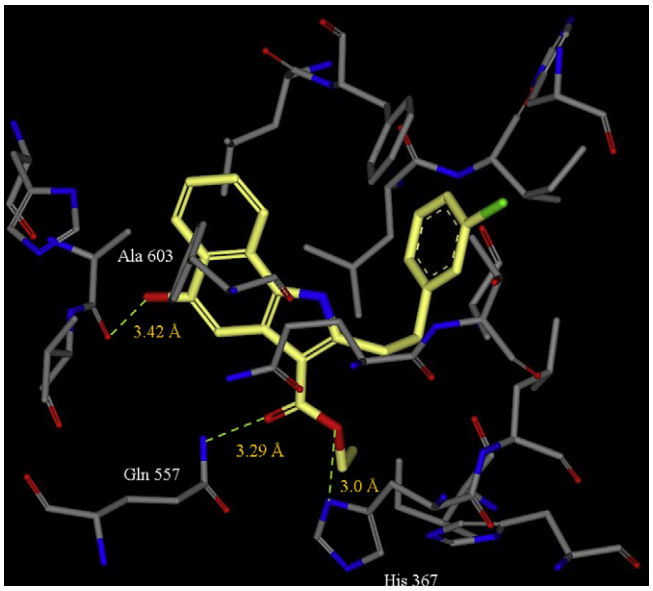

Figure 5. Interactions of the most active molecule with the amino acids at the active site of 5-LOX. Hydrogen bonds are shown in green dotted lines.

defined as a sphere of 10 Å centered on the ligand in the model. During docking, the default algorithm was selected. Input Parameters of the GOLD were set to allow octahedral coordination geometry to iron. After docking, the best and the most energetically favorable conformation of each ligand were selected. A four-stage protocol was set up for energy minimizations of the protein-inhibitor complex[29] (details enclosed in Supplementary data). After the optimization, the protein–ligand complexes were visualized and studied using InsightII. The interactions were further studied in Accelrys DS visualizer 2.5. The most active molecule, compound 24 formed strong hydrogen bond and hydrophobic interactions and is firmly bound in the open cavity that is in the sixth coordination of the iron atom and thus prevents the access of the substrate to the catalytic site of the enzyme efficiently. Strong hydrogen bond interactions are formed with His 367 (ND1···OMe, 3.003 Å), Gln 557 (Nε2···O, 3.294 Å), and Ala 603 (O···OH, 3.420 Å) (Fig. 5). The least active molecule, compound 3 formed only one hydrogen bond interaction with Gln 557 (Nε2···O, 3.2 Å). The difference in the number of hydrogen bonds explains the difference in the activity of these molecules against 5-LOX.

In the CoMFA electrostatic contour of compound **24**, the red color region favored for hydrogen bond acceptor are present corresponding to the amino acid Gln 557. This corresponds well to the docking interactions shown in Figure 5. The blue colored region obtained near the benzyl moiety indicates that electropositive groups are favored in that region. In the steric contour the green coloured regions at the 3́ and 4́ of the benzyl specifies that steric bulk is favoured there. In the docking study it has been seen that hydrophobic amino acids Leu 420, Phe 359, Leu 368 and Ile 415 are present close to the vicinity of the molecule and can be exploited for better inhibitor design. The yellow color region suggests that bulky group would decrease the activity. In the case of compound **27**, the benzyl group present on the indole ring reduced its activity in comparison to other molecules. It has been observed that bulky amino acids Phe 421 is present very close, which may be the cause why bulky groups are not favored in that region. The Steric contours complemented very well with the docking results.

In the CoMSIA contours, hydrogen bond donor unfavorable regions are seen. No favorable regions were seen. This correlated well with the pharmacophore model of 5-LOX reported recently.[11] The pharmacophore features required for 5-LOX inhibition were identified in that study as two hydrogen bond acceptors, one ring aromatic and one hydrophobic feature. Even in a previous pharmacophore model report[27] it has been reported that hydrogen bond donor feature is not required for 5-LOX inhibition. Thus the CoMSIA contour map correlated well with the existing pharmacophore models of 5-LOX. The hydrogen bond acceptor contours suggest that hydrogen bond acceptor groups in the regions corresponding to magenta color will increase the activity, whereas, regions corresponding to blue-green color will decrease the activity. This contour did not correspond well to the interactions visualized. In the CoMSIA hydrophobic contour favorable regions shown in orange corresponded well to the site points in 5-LOX corresponding to hydrophobic amino acids Ile 415, Leu 420 and Phe 359.

All the CoMFA and CoMSIA contours were well explained by the interactions of the molecules within 5-LOX active site except the CoMSIA hydrogen bond acceptor contour. The accuracy of the 3D-QSAR models is reflected by its good correlation with the docking and interaction studies. Further, the QSAR and homology models were used to design 5-hydroxyindole-3-carboxylate derivatives.

The structures of the designed molecules are provided in Supplementary data. Molecules were designed and their activities were estimated by calculation of binding energies using homology model and by activity prediction using the more accurate CoMSIA model. The results show that the molecules may be more effective towards 5-LOX (Table 2 of Supplementary data). The relative binding energies of the designed molecules was calculated using the methodology described in our previous reports[10,29–31] (details enclosed in Supplementary data).

In this study, 3-D CoMFA and CoMSIA QSAR analyses were used to predict the 5-LOX inhibitory activity of a set of compounds. Two datasets of same molecules were prepared with two different partial atomic charges; one with the simple Gasteiger–Huckel charges and another with the ESPFIT charges obtained from the GAUSSIAN package. The QSAR models for both datasets gave good statistical results in terms of q^2 and r^2 values but the one with ESPFIT charges yielded higher values; hence it was further used in the study. The effects of steric, electrostatic, hydrophobic and hydrogen bond donor fields around the aligned molecules on their 5-LOX inhibitory activity were clarified by analyzing the CoMSIA contour maps. The statistical significance of the generated 3D-QSAR models were confirmed using an external set of 10 test set molecules. In

addition to steric and electrostatic fields, hydrophobic and hydrogen bond acceptor fields were also found to be important for 5-LOX inhibitory activity as exemplified by the higher predictive power of the CoMSIA model relative to the CoMFA model. However, to consider structural information about the target as well, a theoretical 3D model of 5-LOX was elaborated by homology with rabbit 15-LOX based on consensus alignment of the sequences. The 5-LOX active site was then characterized from a structural point of view and used for docking of the most active and least active molecules in the dataset. This has shed light on the binding features of the enzyme. The comparison of the interactions of the most active molecule and the contour maps of the 3D-QSAR models provided a better understanding of the 5-LOX-inhibitor interactions.

Acknowledgments

We thank Centre for Modelling, Simulation and Design (CMSD), University of Hyderabad and Bioinformatics Infrastructure Facility, Dept. of Biotechnology, University of Hyderabad for permitting us to use the computational facilities. We duly acknowledge UGC for providing Dr. D. S. Kothari Postdoctoral fellowship to P. Aparoy.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2010.10.119.

References and notes

1. Funk, C. D. *Science* **2001**, *294*, 1871.
2. Rapoport, S. M.; Schewe, T.; Wiesner, R.; Halangk, W.; Ludwig, P.; Janicke-Hohne, M.; Tannert, C.; Hiebsch, C.; Klatt, D. *Eur. J. Biochem.* **1979**, *96*, 545.
3. Brash, A. R. *J. Biol. Chem.* **1999**, *274*, 23679.
4. Davidson, E. M.; Rae, S. A.; Smith, M. J. *Ann. Rheum. Dis.* **1983**, *42*, 677.
5. Nie, D.; Che, M.; Grignon, D.; Tang, K.; Honn, K. V. *Cancer Metastasis Rev.* **2001**, *20*, 195.
6. Anderson, K. M.; Seed, T.; Vos, M.; Mulshine, J.; Meng, J.; Alrefai, W.; Ou, D.; Harris, J. E. *Prostate* **1998**, *37*, 161.
7. Avis, I. M.; Jett, M.; Boyle, T.; Vos, M. D.; Moody, T.; Treston, A. M.; Martinez, A.; Mulshine, J. L. *J. Clin. Invest.* **1996**, *97*, 806.
8. Wang, D.; Dubois, R. N. *Nat. Rev. Cancer* **2010**, *10*, 181.
9. Chen, Y.; Hu, Y.; Zhang, H.; Peng, C.; Li, S. *Nat. Genet.* **2009**, *41*, 783.
10. Aparoy, P.; Reddy, R. N.; Guruprasad, L.; Reddy, M. R.; Reddanna, P. *J. Comput. Aided Mol. Des.* **2008**, *22*, 611.
11. Aparoy, P.; Kumar Reddy, K.; Kalangi, S. K.; Chandramohan Reddy, T.; Reddanna, P. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1013.
12. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
13. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
14. Karg, E. M.; Luderer, S.; Pergola, C.; Buhring, U.; Rossi, A.; Northoff, H.; Sautebin, L.; Troschutz, R.; Werz, O. *J. Med. Chem.* **2009**, *52*, 3474.
15. SYBYL Molecular Modeling Software, Version 8.0, Tripos Inc., St. Louis, MO.
16. Wellsow, J.; Machullab, H. J.; Kovar, K. A. *Quant. Struct.-Act. Relat.* **2002**, *21*, 577.
17. Clark, R. D.; Cramer, R. D.; Van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982.
18. Stahle, L.; Wold, S. *Prog. Med. Chem.* **1988**, *25*, 291.
19. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
20. Golbraikh, A.; Tropsha, A. *J. Mol. Graphics Modell* **2002**, *20*, 269.
21. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
22. Gillmor, S. A.; Villasenor, A.; Fletterick, R.; Sigal, E.; Browner, M. F. *Nat. Struct. Biol.* **1997**, *4*, 1003.
23. Discovery Studio 2.1, Molecular Modeling Software, Accelrys, Inc., San Diego, USA.
24. Profile-3D user guide, Accelrys, Inc., San Diego, USA, 1999.
25. Laskowski, R. A.; Moss, D. S.; Thornton, J. M. *J. Mol. Biol.* **1993**, *231*, 1049.
26. Kuntal, B. K.; Aparoy, P.; Reddanna, P. *Protein Pept. Lett.* **2010**, *17*, 765.
27. Charlier, C.; Henichart, J. P.; Durant, F.; Wouters, J. *J. Med. Chem.* **2006**, *49*, 186.
28. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.
29. Reddy, R. N.; Mutyala, R.; Aparoy, P.; Reddanna, P.; Reddy, M. R. *Curr. Pharm. Des.* **2007**, *13*, 3505.
30. Aparoy, P.; Reddy, R. N.; Guruprasad, L.; Reddanna, P. *Lett. Drug Des. Discov.* **2010**, *7*, 324.
31. Aparoy, P.; Leela, T.; Reddy, R. N.; Reddanna, P. *J. Mol. Graphics Modell* **2009**, *27*, 744.