Food Science and Human Wellness 8 (2019) 53–62

Contents lists available at ScienceDirect

Food Science and Human Wellness

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

QSAR modeling of benzoquinone derivatives as 5-lipoxygenase inhibitors

T.K. Shameera Ahamed, Vijisha K. Rajan, K. Muraleedharan *

Department of Chemistry, University of Calicut, Malappuram 673635, India

ARTICLE INFO

Article history:
Received 21 October 2018
Received in revised form 25 January 2019
Accepted 9 February 2019
Available online 22 February 2019

Keywords:
5-lipoxygenase
Benzoquinone
2D-QSAR
3D-QSAR
Molecular docking

ABSTRACT

The inhibitors of 5-LOX control the overproduction of pro-inflammatory mediators known as leukotrienes (LTs) and thus have therapeutic relevance in the treatment of various diseases like asthma, rheumatoid arthritis, inflammatory bowel disease and certain types of cancers. This has increased the search for efficient therapeutic agents for protein 5-LOX and this process is now primarily based on QSAR. In this study, we have developed four different quantitative structure and 5-LOX inhibition activity relationship models of benzoquinone derivative by exploiting CoMFA, RF, SVM, and MLR chemometric methods. Performance of the QSAR models was measured by using cross-validation technique as well as through the external test set prediction. RF model outperforms all other models. SVM and MLR models failed due to the poor performance of the external test set prediction. CoMFA model, which shows relatively good performance was used to explore the essential structural regions where the modification was necessary to design a novel scaffold with improved activity. Moreover, molecular docking of all the derivatives to the binding site of 5-LOX was done to show their binding mode and to identify critical interacting residues inside the active site of 5-LOX. The docking result confirms the stability and rationality of the CoMFA model.

1. Introduction

The biological properties of human 5-LOX have gained much attention from research because of its involvement in the pathogenesis of several diseases [1]. The 5-LOX pathway is the source of potent pro-inflammatory mediators known as leukotrienes (LTs). 5-LOX catalyzes the first two steps in leukotriene A4 (LTA4) biosynthesis, in which the one is the addition of molecular oxygen into 1, 4-cis–cis-pentadiene containing polyunsaturated fatty acids such as arachidonic acid and linoleic acid to give their hydroperoxy derivatives and the second step is the dehydration of this hydroperoxide to the key intermediate, short-lived epoxide leukotriene LTA4, which then converted to LTs [2]. These LTs are essential mediators of excessive and chronic inflammatory and allergic disorders such as Rheumatoid arthritis, Gastroesophageal reflux disease, Atherosclerosis, Inflammatory bowel disease, and Autoimmune, ulcerative colitis, asthma, psoriasis and allergic rhinitis [3]. Besides its important roles in inflammatory diseases, 5-LOX is also involved in the development and progression of numerous types of cancer such as such as leukemia, pancreas, prostate, and colon cancer [4–8]. Therefore, the pharmacological intervention of the 5-LOX pathway to control the formation of LTs is a promising therapeutic strategy for LT-related diseases.

A wide variety of inhibitors have been reported as active against 5-LOX invitro studies. These compounds include phenols, aromatic compounds containing heteroatoms, carboxamides, hydroxamic acids, flavonoids, chalcones, *etc.* Based on the nature of action, these inhibitors have been classified into four types such as redox inhibitors, iron-chelator agents and non-redox or competitive inhibitors and 5-Lipoxygenase activating protein (FLAP) inhibitors [9,10]. Despite all this intensive effort, Zileuton is the only orally active drug that has been approved as 5-LOX inhibitor used for the treatment of asthma [11]. However, zileuton itself has several side effects including liver toxicity and unfavorable pharmacokinetic profile [12]. So, the development of novel inhibitors having high inhibitory potency along with a favorable pharmacokinetic profile is the major challenge among the scientific community.

Drug discovery has been evolved from 'forward pharmacology' to rational drug design with the advancement of computational

* Corresponding author.
E-mail address: kmuralika@gmail.com (K. Muraleedharan).
Peer review under responsibility of KeAi Communications Co., Ltd

Production and hosting by Elsevier

https://doi.org/10.1016/j.fshw.2019.02.001

methods. Computer Aided Drug Discovery (CADD) revolutionize the area of drug design by identifying compounds with desirable characteristics, speed up the hit-to-lead process and improve the chances of getting your compound past the hurdles of preclinical testing. Now CADD is rapidly gaining in popularity, implementation and appreciation. Several publications have appeared in recent years documenting the rational design of 5-LOX inhibitors using computational methods such as ligand-based techniques like scaffold-hopping, pharmacological analysis, Quantitative structure-activity relationship (QSAR) and structure-based techniques like molecular docking, dynamic modeling [13]. QSAR, a ligand-based CADD technique, is used to build computational or mathematical models which attempts to find a statistically significant correlation between various molecular properties of a set of molecules with their experimentally known biological activity. This method predicts the biological activity of known and unknown compounds by using statistical techniques and optimizing new lead molecules.

Several successful QSAR works were performed in recent years with the aim of formulating an excellent predictive model which composed of common chemical features among a considerable number of known 5-LOX inhibitors using conventional 2D QSAR methods such as Multiple Linear Regression (MLR), Principle component analysis (PCA) and partial least square regression (PLS). In literature, a few 3D-QSAR works for 5-LOX inhibitors also been documented [14,15]. Some of these studies [16,17] have used structure-based Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) as a tool to model the activity of human 5-LOX inhibitors. However, most of these studies have focused only on selective and non-redox type 5-LOX inhibitor and not on redox inhibitors. Non-redox inhibitors generally show diminished potency in a condition with elevated peroxide levels whereas redox inhibitors inhibit lipid peroxidation more effectively by scavenging peroxyl free radicals and suppressing the leukotriene formation [10]. From the literature, it is observed that 5-LOX activity of redox inhibitors was enhanced by the presence of extended hydrophobic alkyl groups. Redox potency of these inhibitors might be directly linked with hydrophobicity.

In our previous study, we reported that the quantitative influence of extended hydrophobic alkyl groups of 3', 4'-dihydroxyflavones derivative over the 5-LOX potency using CoMFA methodology [18]. In this work, we have tried to formulate QSAR models to investigate the interaction of a series of benzoquinone derivatives containing various lipophilic and bulky alkyl substituents reported by Rosanna Filosa et al., [19] with the binding site of 5-LOX and predict their inhibitory activities. Here we reported four types of QSAR model: MLR based linear, RF and SVM based nonlinear 2D-QSAR models and CoMFA based 3D-QSAR model and the best two models were chosen to predict 5-LOX inhibitory activities. Furthermore, molecular docking studies have been carried out to rationalize the CoMFA model by demonstrating the common binding model of benzoquinone derivative with human 5-LOX model.

2. Materials and methods

2.1. Dataset

The dataset used in this study consisted of a series of benzoquinone derivative that has been reported as 5-LOX inhibitors in a cell-free assay using purified human recombinant 5-LOX enzyme by Rosanna Filosa et al. [19]. 2D structure of benzoquinone core is displayed in Fig. 1. The experimental IC_{50} values of all compounds in μM were converted into pIC_{50} by taking -Log $(1/IC_{50})$ and were used as the dependent variable. There were a total of 48 benzoquinone derivatives which are then split into a training set of 30 compounds for generating QSAR models and a test set of 11 compounds for validating the quality of the models. Remaining 7 compounds having IC_{50} value greater than 10 μM were removed. The compounds in the test set were manually selected from the original pool of structures based on Y-response (dependent variable). This approach is based on the activity (Y-response) sampling. For maintaining uniform distribution, molecules with low, moderate and high activity were placed in both sets. Most active and least active molecules were retained in training set for better performance. All the structures and associated inhibitory activities are listed in Table 1.

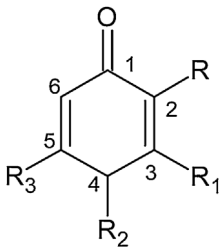

Fig. 1. 2D chemical structure of benzoquinone core.

2.2. Molecular modeling

All the 3D structures were drawn and built by Gauss View 05. Gas phase geometries were optimized using Density Functional Theory (DFT) [20] of three-parameter compound of Becke (B3LYP) [21,22] employing 6–31 G(d, p) basis set using Gaussian 09 program package [23]. Harmonic energy calculations are carried out for confirming frequencies are all real. The lowest energy conformer of each compound was used for further analysis.

2.3. 2D-QSAR methodology

2.3.1. Calculation of 2D molecular descriptors

Descriptors are the mathematical representation of a molecule which contains different sources of chemical information transformed and coded to deal with chemical, biological and pharmacological problems. For the development of 2D QSAR models, various physicochemical descriptors are calculated for each of the compounds in the dataset by means of three different software such as e-DRAGON [24], PowerMV [25] and Gaussian 09. Different sets of 0D, 1D, 2D and 3D molecular descriptors are calculated with the help e- DRAGON software. Pharmacophore Fingerprint descriptors and Weighted Burden Number descriptors were computed by PowerMV software. Pharmacophore Fingerprint descriptors were built based on bioisosteric principles (Two atoms or groups that are expected to have roughly the same biological effect are called bioisosteres). Electronic and quantum chemical descriptors such as highest occupied molecular orbital (HOMO) energies, lowest unoccupied molecular orbital (LUMO) energies and molecular dipole moment were calculated by Gaussian 09. DFT based global reactivity descriptors provide the reactivity of chemical species regarding electronic features. Ionization potential (IP), electron affinity (EA), electronegativity (χ), electrophilicity (ω), softness (S), hardness (η) and chemical potential (μ) are the commonly used global reactivity descriptors. Koopman's theorem approximates the negative value of HOMO and LUMO as ionization potential and electron affinity respectively. For the development of 3D-QSAR model steric and electrostatic descriptors were considered. The 2D-Descriptors used in the study given in Table 2. 2D-QSAR models were generated using WEKA software [26].

Table 1
Structural formulae of compounds and their IC_{50} values.

Compound	R	R_1	R_2	R_3	IC_{50} (μM)
17a	—OCH_3	n-butyl	=O	—OCH_3	3.3 ± 0.9
18a	—OCH_3	n-hexyl	=O	—OCH_3	2.6 ± 1.2
19a	—OCH_3	n-octyl	=O	—OCH_3	1.2 ± 0.3
20a	—OCH_3	n-decyl	=O	—OCH_3	1.8 ± 0.7
21a	—OCH_3	n-undecyl	=O	—OCH_3	0.93 ± 0.13
22a	—OCH_3	n-dodecyl	=O	—OCH_3	0.61 ± 0.08
23a	—OCH_3	n-tridecyl	=O	—OCH_3	0.26 ± 0.01
24a	—OCH_3	n-tetradecyl	=O	—OCH_3	0.21 ± 0.03
25a	—OCH_3	n-pentadecyl	=O	—OCH_3	0.27 ± 0.08
26a	—OCH_3	n-hexadecyl	=O	—OCH_3	1.6 ± 0.2
27a	—OCH_3	geranyl	=O	—OCH_3	3.3 ± 0.7
28a	—OCH_3	farnesyl	=O	—OCH_3	1.7 ± 0.7
17b	—OH	n-butyl	=O	—OCH_3	>10
18b	—OH	n-hexyl	=O	—OCH_3	>10
19b	—OH	n-octyl	=O	—OCH_3	3.0 ± 0.4
20b	—OH	n-decyl	=O	—OCH_3	4.3 ± 0.3
21b	—OH	n-undecyl	=O	—OCH_3	3.8 ± 0.5
22b	—OH	n-dodecyl	=O	—OCH_3	0.74 ± 0.08
23b	—OH	n-tridecyl	=O	—OCH_3	0.92 ± 0.47
24b	—OH	n—tetradecyl	=O	—OCH_3	0.42 ± 0.01
25b	—OH	n-pentadecyl	=O	—OCH_3	0.27 ± 0.10
26b	—OH	n-hexadecyl	=O	—OCH_3	>10
27b	—OH	geranyl	=O	—OCH_3	>10
28b	—OH	farnesyl	=O	—OCH_3	5.6 ± 0.1
17c	=O	n-butyl	—OCH_3	—OCH_3	>10
18c	=O	n-hexyl	—OCH_3	—OCH_3	2.6 ± 0.3
19c	=O	n-octyl	—OCH_3	—OCH_3	0.33 ± 0.05
20c	=O	n-decyl	—OCH_3	—OCH_3	0.13 ± 0.01
21c	=O	n-undecyl	—OCH_3	—OCH_3	0.09 ± 0.04
22c	=O	n-dodecyl	—OCH_3	—OCH_3	0.13 ± 0.12
23c	=O	n-tridecyl	—OCH_3	—OCH_3	0.08 ± 0.01
24c	=O	n-tetradecyl	—OCH_3	—OCH_3	0.04 ± 0.02
25c	=O	n-pentadecyl	—OCH_3	—OCH_3	0.62 ± 0.14
26c	=O	n-hexadecyl	—OCH_3	—OCH_3	n.d.
27c	=O	geranyl	—OCH_3	—OCH_3	0.6 ± 0.08
28c	=O	farnesyl	—OCH_3	—OCH_3	0.3 ± 0.07
17d	—OH	n-butyl	=O	—OH	>10
18d	—OH	n-hexyl	=O	—OH	4.0 ± 1.1
19d	—OH	n-octyl	=O	—OH	0.38 ± 0.04
20d	—OH	n-decyl	=O	—OH	0.18 ± 0.01
21d	—OH	n-undecyl	=O	—OH	0.06 ± 0.001
22d	—OH	n-dodecyl	=O	—OH	0.17 ± 0.03
23d	—OH	n-tridecyl	=O	—OH	0.22 ± 0.09
24d	—OH	n-tetradecyl	=O	—OH	0.17 ± 0.08
25d	—OH	n-pentadecyl	=O	—OH	0.23 ± 0.08
26d	—OH	n-hexadecyl	=O	—OH	0.19 ± 0.04
27d	—OH	geranyl	=O	—OH	1.8 ± 0.2
28d	—OH	farnesyl	=O	—OH	2.5 ± 1.4

Table 2
Descriptors used in the 2D-QSAR study.

Descriptor Types	Descriptor description	Total No. of Descriptor
e-dragon descriptors	Constitutional descriptors, Walk and path counts, Connectivity indices, Information indices, Topological indices, Charge descriptors, Molecular properties	584
PowerMV descriptors	Pharmacophore fingerprints, Weighted burden number, Property descriptors	179
Quantum chemical descriptors	HOMO, LUMO, HOMO-LUMO gap, dipole moment, IP, EA, χ, ω, S, η and μ	11

2.3.2. Feature selection

Feature Selection methods have been employed for the selection of suitable descriptors from a vast number of raw descriptors contain little information or are correlated with other descriptors without incurring much loss of information. CFS (Correlation-based Feature Selection) [27] is a highly effective feature selection method in reducing dimensionality, removing redundant descriptors, increasing learning accuracy, and improving the performance of machine learning algorithms. The heart of CFS is a heuristic for evaluating the worth or merit of a subset of features by the following hypothesis: G̈ood feature subsets contain features highly correlated with the classification, yet uncorrelated to each other.T̈he following equation (Eq. (1)) gives the merit of a feature subset S consisting of k features:

$$Merit_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \quad (1)$$

Here, r_{cf} is the average value of all feature-classification correlations, and r_{ff} is the average value of all feature-feature correlations. The Eq. (1) is, in fact, Pearson's correlation where all variables have been standardized.

2.3.3. Multiple linear regression model

Multiple linear regression or MLR is a conventional and commonly used method in QSAR due to its simplicity, flexibility, reproducibility, and easy interpretability. The MLR attempts to model the relationship connecting a group of explanatory variables X and a response variable Y fitting a linear equation to observed data. The MLR model has a following mathematical form, given n observations:

$$Y = C_0 + \sum C_n X_N \quad (2)$$

Where Y is the pIC_{50} of the benzoquinone derivatives, C_0 is the intercept and C_n are the regression coefficients of the descriptors X_n. Although MLR is computationally simple and the prediction models give a strong mechanistic interpretation, it is criticized for its lack of robustness in handling the non-linear data even though MLR models serve as the basis for some multivariate methods [28].

2.3.4. Support vector machines

Support vector machines (SVM) [29] are a bunch of supervised learning methods mostly used for classification and regression challenges. Classification of data using SVM involves by looking for a hyperplane in high dimensional space of independent variables that separate positive and negative data at an optimal distance using a non-linear kernel function. SVM methodology purely laid on maximizing the margin between a small subset of training instances (the support vectors) and the hyperplane. SVM methods are one of the most popular machine-learning methods in chemoinformatics. Several good reviews highlighting the application SVM in QSPR/QSAR studies particularly in drug design have been published [30]. The main advantages of SVM are: results are stable, reproducible, and largely independent of the optimization algorithm. The choice of a correctly configured kernel function is an important parameter to a successful SVM model. Polynomial kernel and Radical basis function (RBF) kernel are the two widely used kernels for solving classification problems. In this study, we employed a Polynomial kernel rather than RBF. WEKA Implements John Platt's Sequential Minimal Optimisation (SMO) [31], an algorithm for training a support vector classifier. SMO normalizes all attributes; replace all missing values and transforms nominal attributes into binary ones.

2.3.5. Random forest

Random forest (RF) [32] (forest of decision trees) is an ensemble learning method of unpruned classification and regression trees such that each tree depends on the values of a random vector sampled independently and the same distribution of all trees in the forest. After a large number of trees is generated, each tree casts a unit vote for the most popular class. RF drawn sub-samples from original data with a replacement called as bootstrap sampling and

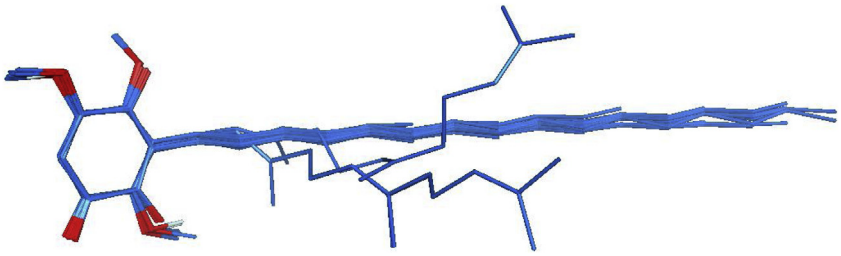

Fig. 2. Alignment of 48 benzoquinone derivatives.

fits trees to these samples. Here one-third of the data is left out of the bootstrap sample and used to testing while the rest form a training set. The prediction error internally estimates the performance of the developed model for the objects left out in the bootstrap procedure (out-of-bag estimation, OOB). Unlike Decision Tree which has relatively low prediction accuracy, RF displays some unique features that make it highly suitable for QSAR tasks. These include built-in estimation of prediction accuracy, the power of handling large dataset with higher dimensionality, measures of importance for each descriptor in the model, and a measure of similarity between molecules. This study explored the influence of the number of trees for the model's predictive ability. Results showed that no significant change in the model's predictive ability by increasing the number of trees beyond 1000. For all the descriptor set, the optimal tree values in the forest were set to 100.

2.4. Molecular modeling for 3D-QSAR

2.4.1. Molecular alignment and conformational analysis

Proper alignment of the compounds relative to one another is one of the most important steps in the 3D-QSAR analysis for obtaining valid molecular interaction field model. Energy minimized structures of molecules were aligned by the template-based method. The mixed alignment procedure in the combination of the Atom-based fashion and pharmacophore-based fashion was performed using Open3DALIGN software (version 2.27), an open-source tool capable of carrying out the multi-conformational, unsupervised rigid-body alignment of 3D molecular structures [33].The alignment procedure was executed by using all available molecules as possible templates. Hence, 48 alignments were produced, each obtained by superimposition on the corresponding template molecule. For each alignment, an O3 A score derived from the source code of the Open3DALIGN program is computed which indicates the quality of the superimposition. O3 A score for each alignment is given in supporting information S1. The alignment corresponding to the highest cumulative O3 A score was selected for further analysis. Fig. 2 shows the best alignment in which compound 26a was selected as the template.

2.4.2. CoMFA analysis

CoMFA is a versatile method to describe 3DQSAR quantitatively. Open3DQSAR software (version 2.282) is open-source software available for high-throughput chemometric analysis of molecular interaction fields (MIFs) [34]. This study used Open3DQSAR to perform CoMFA analysis. The best alignment with compound '26a' as a template is placed in a 3D cubic lattice with 2 Å grid size and a 5.0 Å outgap. The steric fields were computed using sp^3 hybridized carbon atom probe with +1 charge. Similarly, electrostatic fields were computed using a volume-less probe. These steric and electrostatic interaction energies were considered as independent variables (CoMFA descriptors). Before creating of CoMFA model following pre-treatment operations were carried out to reduce the noise hidden in PLS matrix and hence reduced the computational time: 1) the minimum and maximum energy values of steric and electrostatic were set to a cutoffs value −30.0 and + 30.0 kcal/mol, respectively. This pretreatment avoids infinity of energy values inside the molecule. 2) Low energy values (<0.05 kcal/mol) were set to zero in both fields. 3) Standard deviation set to <0.1 in order to improve the signal-to-noise ratio. 4) N-level variables that are variables which assume only N values across the training set were removed, most of which distributed on a small number of objects. This process avoids overweighting the importance of particular substituents present in a single molecule. Otherwise, it might negatively affect the whole model. 5) The whole block of *X* or *Y* variables scaled by block unscaled weighting (BUW) technique.

Predictivity of CoMFA model can be significantly improved by appropriate variable clustering and selection procedures such as smart region definition (SRD) and fractional factorial design (FFD). These variable selection techniques selectively remove noisy variables with no predictability. The SRD procedure carries out variable grouping based on their closeness in 3D space in order to reduce the redundancy arising from the existence of multiple nearby descriptors which mainly encode the same kind of information [35]. FFD aims at selecting the variables which significantly increase the predictive ability (using the LOO, LTO or LMO paradigms), and can operate on both single variables or groups identified by a previous SRD run, thereby removing uninformative variables groups as performed in GOLPE [36].

PLS analysis implemented in Open3DQSAR was employed to obtain a correlation between the descriptors derived by CoMFA (independent variables) and pIC_{50} values (dependent variable). Open3DQSAR generate a PLS model through the Nonlinear iterative partial least squares (NIPALS) algorithm [37].The statistical parameters like coefficient of determination (R^2), Standard Deviation Error in Calculation (SDEC), Standard Deviation Error in Predictivity (SEDP) and F-ratio test were computed the overall significance of model (Eq. (3)–(5)). Moreover, the CoMFA color contour maps are derived for the steric and electrostatic fields.

$$SDEC = \left[\frac{\sum_{i=1}^{n}(y_{obs,i} - y_{calc,i})}{n}\right]^{\frac{1}{2}} \tag{3}$$

$$SDEP = \left[\frac{\sum_{i=1}^{n}(y_{obs,i} - y_{pred,i})}{n}\right]^{\frac{1}{2}} \tag{4}$$

$$F(n, n-p-1) = \frac{(n-p-1)R^2}{p(1-R^2)} \tag{5}$$

Where, $y_{obs,i}$ is the experimental activity, $y_{calc,i}$ is the estimated y in the calibration step and $y_{pred,i}$ predicted the activity of the test set. The value corresponding to 'n' and 'p' is the number of samples in the training set and the number of components in the PLSR model respectively.

2.5. Statistical analysis and model validation

2.5.1. Internal validation

Three main cross-validations (CV) techniques such as leave-one-out (LOO), leave-two-out (LTO) and leave-many-out (LMO) were used to explore the reliability of statistical models. In LOO-CV, each time one compound is removed from the original training set, and a new model is built based on the rest of the set and this model is used to predict the activity of the omitted one. This procedure is repeated for whole compounds of data set. In LTO CV, two compounds are removed instead of one and the remaining procedure repeated as same as that of LOO. In the LMO method, each time 20% of compounds were removed randomly and the procedure was repeated 20 times and predicted their activities *via* the reduced model. Golbraikh and Tropsha reported that the LMO CV is much more robust than LOO-CV and also a high value of Q^2 is essential and important but not adequate for a predictive model [38]. The cross-validated R^2, *i.e.*, Q^2 is given in Eq. (6).

$$R^2_{cv} = Q^2 = 1 - \left(\frac{RESS}{SSY}\right) \tag{6}$$

The term PRESS is the sum of squared difference between experimentally observed activity and the activity predicted by a regression model estimated when the i^{th} sample was left out from the training set and the SSY is the sum of squared differences between the experimental activity and the average experimental activity. According to Hawkins et al., a valid statistical model should have high Q^2 value ($Q^2 > 0.5$) and is evidence of the high predictive ability of the model [39].

2.5.2. External validation

The predictive power of the generated model was evaluated using an external test set of 13 molecules. The predictive correlation coefficient (R^2_{pred}) was determined according to the equation shown below, *i.e.*, Eq. (7).

$$R^2_{pred} = \left(\frac{SD - PRESS}{SD}\right) \tag{7}$$

SD is defined as the sum of the squares of the deviation between the experimentally observed activity of the test set compounds and the mean activity of the training set molecules.

2.6. Docking studies

Molecular docking methodology has been used to predict the best binding orientation of ligand molecule in the active site of receptor targets and recently it is the most used computational tool for drug designing and virtual screening. In this study, the crystal structure of human 5-LOX obtained from PDB database having PDB ID 3o8y was used for the docking study. AutoDock Vina [40] software was used to carry out molecular docking analysis. Optimized the dimension of the grid box to 20 × 20 × 25 Å with centre at −8.374, 66.379, −1.009 for x, y and z respectively. AutoDock Vina uses a sophisticated gradient optimization in its local optimization procedure for rigid-flexible molecular docking. The AutoDock Vina represents the output of the docking result as Gibbs free energy of binding (ΔG). The top nine docking poses of each compound were visually analyzed and emphasized by ΔG values to rank the different conformations of the receptor-ligand complex. The ligand confirmation with lowest ΔG values was taken for further study. Protein-ligand complexes were visualized and analyzed using three different molecular modeling software LigPlus [41], Autodock tool 1.5.6 [42], Chimera [43] and PyMol [44].

Table 3
Statistical data of optimal 2D-QSAR models.

Statistical quality parameter	MLR	RF	SVM
R	0.86	0.98	0.86
MAE	0.23	0.11	0.21
RMSE	0.29	0.15	0.31
R^2	0.75	0.93	0.70
Q^2(LOO)	0.66	0.52	0.61
Q^2(LMO)	0.61	0.51	0.62
R^2_{pred}	0.54	0.71	0.56
n	30.0	30.0	30.0

3. Results and discussion

3.1. Feature selection and the 2D-QSAR prediction model

In order to develop accurate, robust and efficient 2D-QSAR models to predict 5-LOX inhibition activity of benzoquinone derivatives, a small subset of descriptors which represent the total set of descriptors has been identified through this work. We have employed the CFS optimization technique with the help CfsSubsetEval attribute evaluator of WEKA. Five descriptors show maximum correlation such as LUMO (lowest unoccupied molecular orbital- a quantum chemical descriptor), HBD_06_HBA (Pharmacophore fingerprints), WBN_EN_H_0.75 (Weighted Burden Number), X5A (average connectivity index of order 5) and JGI3 (mean topological charge index of order 3-2D autocorrelation). Selected descriptors used in 2D-QSAR model with values were provided in supporting information S2. Based on these optimal features subset, the possibility of predicting 5-LOX inhibition activity was investigated with the help of MLR, SVM, and RF regression methods. The statistical performance of the optimum MLR, SVM and RF models using default parameters, is summarized in Table 3.

3.2. Statistical analysis of 2D-QSAR models

MLR Regression is the easiest and simplest technique to construct QSAR models but is also probably the least powerful. It can be used as a preliminary step before moving onto more complex algorithms. Linear regression works by calculating the coefficients for a line or hyperplane that best fits the training data. The best MLR equation obtained for the pIC_{50} of the benzoquinone derivatives is based on the LUMO, HBD_06_HBA and JGI3 descriptors is given in Eq. (8).

$$pIC_{50} = -29.1948 \times LUMO + -0.5634 \times HBD\_06\_HBA + -62.3572 \times JGI3 + 6.5459 \tag{8}$$

Constructed MLR equations indicate the negative contribution of the above descriptors towards the prediction of 5-LOX activity of benzoquinone derivatives, *i.e.*, 5-LOX receptor binding activity of these inhibitors might be decreased by increasing the LUMO energy value and HBD_06_HBA value as well as JGI3 value. The LUMO is the lowest energy level with an empty electron in the molecule. The energy of the LUMO is directly related to the electron affinity, *i.e.*, when a molecule undergoes a bonding interaction with protein; incoming electron pairs are received in its LUMO. Thus, the LUMO descriptor measures the electrophilicity of a molecule. Molecules with low LUMO energy are more able to accept electrons than those with high LUMO energy. The regression equations show that the low LUMO energy values positively influence the 5-LOX inhibition activity of benzoquinone derivatives. HBD_06_HBA is pharmacophoric descriptors account the hydrogen bond donor-acceptor features of benzoquinone that are thought to be responsible for pharmacological action. JGI3 is third order mean

Galvez topological charge descriptor, the negative influence of this descriptor from this class to the 'activity' suggested that a higher order charge index would not be beneficiary to the activity.

By analyzing the statistical parameter given in Table 3, we can say that the MLR model explains about 75% of the 5-LOX inhibitory activity variance of the training set. The predictive power of the MLR model was validated internally through cross-validation correlation coefficient Q^2 having 0.66 for LOO and 0.61 for LMO and externally through test set prediction using R^2_{pred} value, 0.54. The R^2_{pred} value of the test set is less than 0.6; this indicates the MLR model is not satisfactory for prediction of the external test set. However, Stability of the QSAR models was analyzed by progressive Y-scrambling by taking the MLR model as a representative system. Y-scrambling (Y-randomisation) was applied to exclude the probability that our QSAR model performance could have occurred by chance. The Y-vector (pIC_{50}) of the 30 compounds in the training set are sorted according to decreasing pIC_{50} value, then these values were shuffled randomly and a new model was constructed. The shuffling within blocks is repeated 50 times. For each scrambling, an MLR and a CV model were computed. The obtained R^2 and Q^2 (LOO) values for random models should be less than that of obtained for the main initial model, then only we can accept the reliability of the main model. The results of the Y-scrambling test are given in supporting information S3. In all cases, the obtained random models have much lower prediction accuracies than the model based on the real data, indicating no apparent chance correlation in the QSAR model. The Average R, R^2 and Q^2 (LOO) for the 5-LOX inhibitors random model are around 0.41, 0.18 and -0.32 respectively.

SMO regression results were obtained using a popular algorithm-Sequential Minimization Optimization for regression (SMOreg). Statistical quality parameters were analyzed and it is found that Root means squared error (RMSE) generated by SMOreg were slightly higher than MLR model but this small difference may not be relevant and all other quality parameter values were slightly better than MLR model. R^2_{pred} values of the test set are still less than 0.6; this indicates the SVM model is also not entirely satisfactory for prediction of the external test set. Hence it can be concluded that both MLR and SVM algorithms can be equally useful for internal prediction but not so good for external prediction. The experimental and predicted activities of the training set and the independent test set of MLR and SVM model are listed in Tables 5 and 6 respectively and corresponding scatter plot of observed *vs.* predicted values of pIC_{50} of both the training and test set are shown in Fig. 3A and B respectively. This data shows that the experimental and the predicted activities of inhibitors are not very close to each other. Most of the molecules show residual values greater than 0.4.

Then we made an RF regression model which works by constructing an ensemble of decision trees using training set and outputting mean prediction of the individual trees. One hundred trees and one seed were used for building the RF model for the study. Mean absolute error (MAE) measurement of the obtained model gives the magnitude of the error in prediction. For the training set, an MAE of 0.1154, RMSE of 0.1545 and R^2, of 0.93 is obtained indicating that the better predictions. The cross-validated R^2 (Q^2) is 0.52, and the R^2_{pred} is 0.71, indicating good internal and external predictions of the developed RF model. The predicted pIC_{50} values by RF model for training and test set are listed in Tables 5 and 6 respectively. Fig. 3C demonstrates the correlation between experimental and predicted pIC_{50} values by the RF model. These plots further reveal that the 2D-QSAR model based on the RF method is much better than that based on MLR and SVM methods. From all these results it can be concluded that the present RF model exhibits excellent predictive power from both the internal and external points of view concerning the prediction of the test sets.

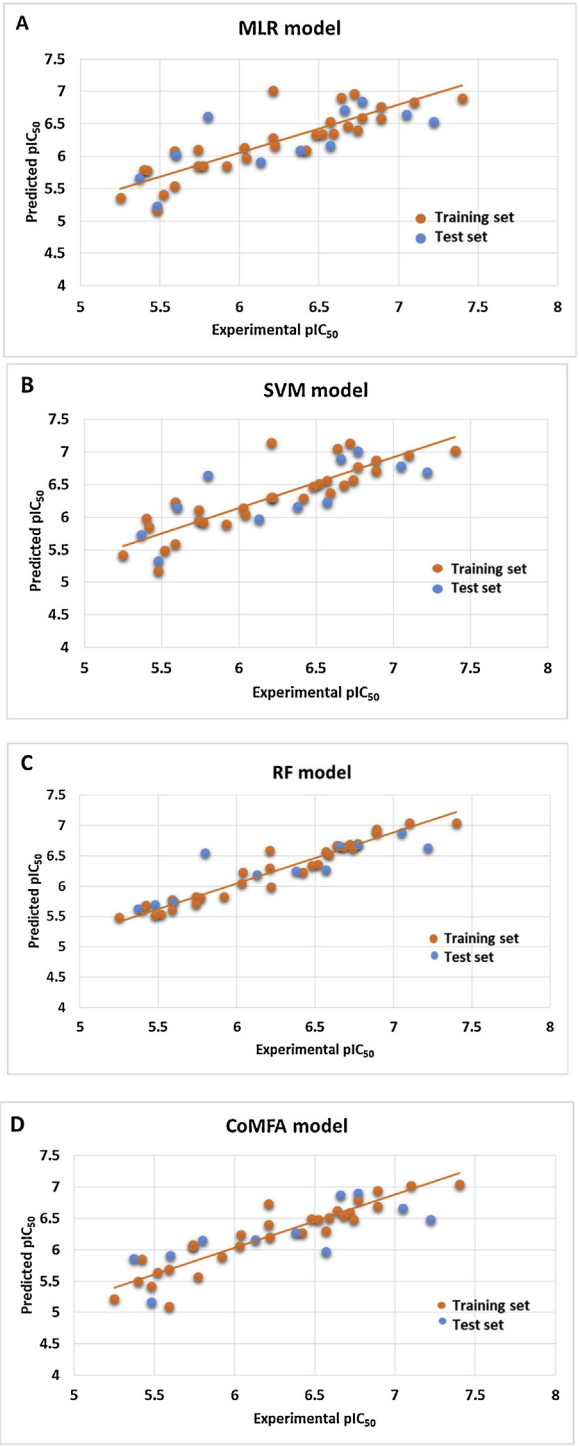

Fig. 3. Activity plots of observed *vs.* predicted pIC_{50} of training and test set of 5-LOX inhibitors by the A) MLR, B)SVR, C)RF and D) CoMFA models.

3.3. Statistical analysis of CoMFA models

Using the training set of 30 benzoquinone derivatives, CoMFA model with five PLS components was built and then, the external test set including 11 compounds was used to evaluate the reliability and applicability of the built model. Statistical quality parameters associated with CoMFA models based on FFD procedures for noise reduction in the input data is listed in Table 4. The analysis of these parameters revealed that the best CoMFA model was obtained with a combination of steric and electrostatic fields. The

Table 4
Statistical data of optimal CoMFA model.

Statistical parameter	CoMFA
R^2	0.8489
Q^2(LOO)	0.5976
Q^2(LTO)	0.5851
Q^2(LMO)	0.5361
R^2_{pred}	0.6000
F	26.9743
SDEC	0.2203
SDEP	0.3651
Steric contribution	0.7715
Electrostatic contribution	0.2285

percentage contribution of the steric field and electrostatic field to the PLS model is 77 and 23% respectively. *i.e.*, more than 75% contribution was observed from the steric field, indicating that steric interaction is essential to the binding of benzoquinone analogs with 5-LOX. CoMFA model gave good cross-validated correlation coefficient (Q^2) for LOO, LTO, and LMO as 0.5976, 0.5851 and 0.5361 respectively, indicating an excellent internal predictive power of the established model. The non-cross-validated PLS analysis with the five components resulted in a conventional R^2 of 0.8489, F of 26.97 and SDEP of 0.2203 for CoMFA model was found to reasonable. The values of experimental and predicted activities along with the residual values of the training set and test set molecules are summarized in Tables 5 and 6 respectively. The scatter plot of observed *vs.* predicted values of pIC_{50} of both the training and test set of CoMFA models is shown in Fig. 3D. This data shows that the experimental and the predicted activities of inhibitors are very close to each other. Most of the molecules show residual values less than 0.4. This graphical representation again conforms the good predictive power of the established model and also indicated that the developed CoMFA model is reliable, and could be used in designing new inhibitors.

3.4. Graphical interpretation of the CoMFA contour maps

The most significant advantage of CoMFA is that it generates 3D contour plots around the molecules. These contour maps are used to identify regions in MIFs of the molecules included in the training

Table 5
The experimental and predicted pIC_{50} values of the training set.

compound id	actual	MLR		SVM		RF		CoMFA	
		predicted	error	predicted	error	predicted	error	predicted	error
28c	6.52	6.35	−0.17	6.52	0.00	6.37	−0.15	6.48	0.04
28b	5.25	5.36	0.11	5.42	0.17	5.50	0.25	5.22	0.03
28a	5.77	5.85	0.08	5.93	0.16	5.82	0.05	5.57	0.20
27d	5.74	5.85	0.11	5.96	0.22	5.71	−0.03	6.05	−0.31
27c	6.22	6.16	−0.06	6.31	0.09	5.99	−0.23	6.21	0.01
27a	5.48	5.17	−0.31	5.19	−0.30	5.52	0.04	5.42	0.06
26d	6.72	6.97	0.25	7.14	0.42	6.70	−0.02	6.60	0.12
25d	6.64	6.91	0.27	7.06	0.42	6.67	0.03	6.63	0.01
25c	6.21	7.02	0.81	7.15	0.94	6.59	0.38	6.73	−0.52
25a	6.57	6.54	−0.04	6.57	0.00	6.58	0.01	6.30	0.27
24c	7.40	6.90	−0.50	7.03	−0.37	7.05	−0.35	7.05	0.35
24a	6.68	6.47	−0.21	6.49	−0.19	6.64	−0.04	6.55	0.13
23c	7.10	6.84	−0.27	6.96	−0.15	7.05	−0.05	7.03	0.07
23b	6.04	5.98	−0.07	6.04	0.00	6.23	0.19	6.25	−0.21
23a	6.59	6.35	−0.24	6.38	−0.21	6.53	−0.07	6.51	0.08
22d	6.77	6.60	−0.17	6.77	0.00	6.70	−0.07	6.79	−0.02
22c	6.89	6.77	−0.12	6.88	−0.01	6.95	0.06	6.95	−0.06
22a	6.21	6.29	0.08	6.30	0.09	6.31	0.10	6.41	−0.20
21b	5.42	5.79	0.37	5.85	0.43	5.70	0.28	5.86	−0.44
21a	6.03	6.14	0.11	6.14	0.11	6.05	0.02	6.07	−0.04
20d	6.74	6.41	−0.33	6.58	−0.16	6.63	−0.11	6.49	0.25
20c	6.89	6.59	−0.30	6.71	−0.18	6.89	0.00	6.69	0.20
20a	5.74	6.11	0.37	6.12	0.38	5.83	0.09	6.08	−0.34
19d	6.42	6.10	−0.32	6.30	−0.12	6.24	−0.18	6.27	0.15
19c	6.48	6.34	−0.14	6.48	0.00	6.35	−0.13	6.50	−0.02
19b	5.52	5.42	−0.11	5.49	−0.03	5.54	0.02	5.64	−0.12
19a	5.92	5.85	−0.07	5.90	−0.02	5.83	−0.09	5.89	0.03
18d	5.40	5.79	0.39	5.99	0.59	5.62	0.22	5.51	−0.11
18c	5.59	6.09	0.50	6.24	0.65	5.79	0.20	5.70	−0.11
18a	5.59	5.55	−0.05	5.59	0.00	5.61	0.02	5.10	0.49

Table 6
The experimental and predicted pIC_{50} values of the test set.

Compound	Actual	MLR		SVM		RF		CoMFA	
		Predicted	error	Predicted	error	Predicted	error	Predicted	error
28d	5.60	6.03	0.43	6.15	0.55	5.75	0.15	5.92	−0.32
26a	5.80	6.62	0.82	6.65	0.85	6.56	0.76	6.15	−0.35
25b	6.57	6.16	−0.41	6.24	−0.33	6.28	−0.29	5.98	0.59
24d	6.77	6.85	0.08	7.01	0.24	6.68	−0.09	6.91	−0.14
24b	6.38	6.10	−0.28	6.16	−0.22	6.25	−0.13	6.27	0.11
23d	6.66	6.72	0.06	6.89	0.23	6.67	0.01	6.88	−0.22
22b	6.13	5.91	−0.22	5.97	−0.16	6.19	0.06	6.16	−0.03
21c	7.05	6.65	−0.40	6.79	−0.26	6.88	−0.17	6.67	0.38
20b	5.37	5.66	0.29	5.73	0.36	5.63	0.26	5.86	−0.49
17a	5.48	5.23	−0.25	5.33	−0.15	5.71	0.23	5.17	0.31
2	7.22	6.54	−0.69	6.70	−0.52	6.64	−0.58	6.49	0.73

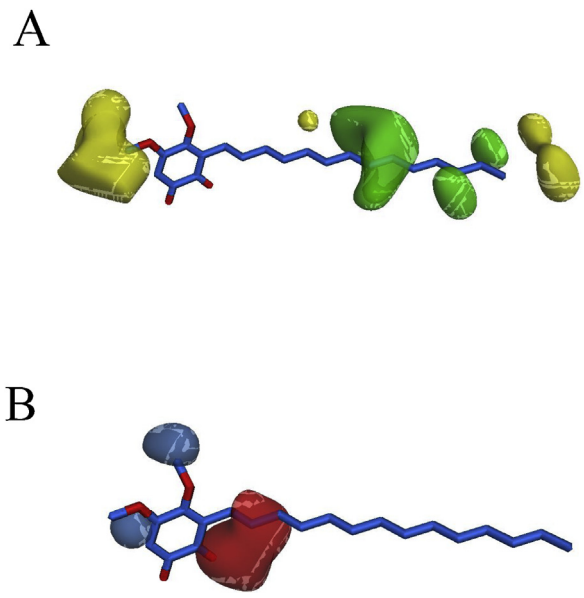

Fig. 4. PLS contours obtained from 3D-QSAR model of 5-LOX inhibitors. (A) CoMFA steric contour maps (B) CoMFA electrostatic contour maps.

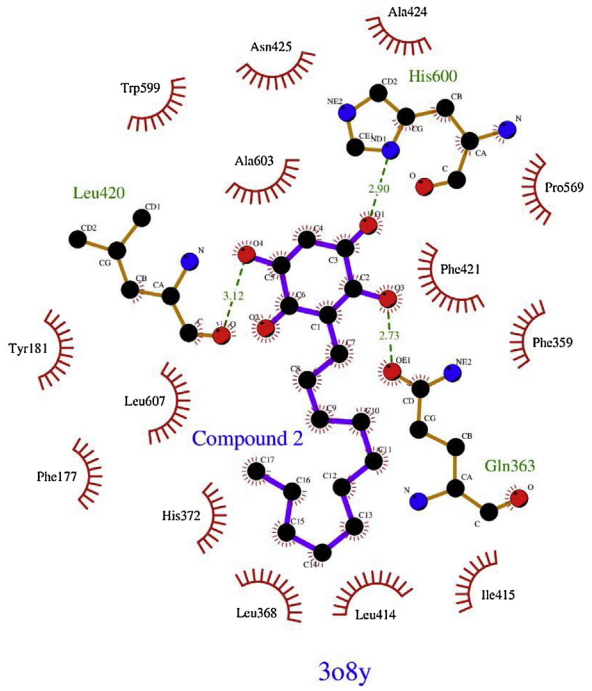

Fig. 5. 2D view of the binding interaction of Compound 2 (Embelin) with 5-LOX.

set where any change in the steric and electrostatic field might affect the biological activity and they also provide hints for the modification required to design new molecules with better activity. The CoMFA steric and electrostatic contour maps are shown in Fig. 4. The Green and yellow contours represent the steric fields. In detail, the green region in the steric contour maps indicates an area where the bulky groups are favored for activity while the yellow contours represent regions where the bulky groups are not favored for the activity. The red and blue contour represent electrostatic contour maps. The blue contour defines a region of space where positively charged substituent increases activity, whereas the red contour defines a region of space where negatively charged substituent increases activity.

These contour maps give us some general insight into the nature of the receptor-ligand binding region. Three green plots were found around the middle of the 'n-alkyl' residue in position 3 indicate large groups in this region (C10-, C11-, C12-, C13-, C14-, C15- and C16-) is favorable to increase the activity of the ligand. To justify this, we could say that the compounds 20d, 22d, 24d, and 26d with C10-, C12-, C14-, and C16-*n*-alkyl chains, respectively inhibited 5-LOX with low IC_{50} values between 0.17 and 0.19 μM than those of the compounds 17d, 18d and 19d with C4-, C6- and C8-alkyl substituent respectively. The long, bulky alkyl group in these regions are significant for a potentially active ligand. This observation is in agreement with general findings of parent literature showing that the potency of (poly) phenol-based 5-LOX inhibitors is often enhanced due to increasing the lipophilicity. This sterically crowded alkyl group may bring a hydrophobic nature to the parent benzoquinone thereby enhances the activity. The small yellow contours at the tail portions of the alkyl residue in position 3 show that too long alkyl chains (greater than C16-) might have a negative influence on its activity. A large yellow contour at the 5th position indicates that the —OH group is preferable at this position as compared to —OCH_3. This result is confirmed by the lower activity of compounds in 'b series' like 18b, 26b, and 27b (Table 1) which are the derivative of benzoquinone methylated the hydroxyl in 5th position.

In an electrostatic field, red and blue contours are mostly distributed near to the core of the benzoquinone. So the electronegative and electropositive substituents in these regions are likely to boost biological activity. The red contours surrounded the 2nd and 3rd position suggests that the electron-rich substituents in this region are likely to enhance biological activity. It indicates alkoxy (—OCH_3) groups are preferably in the 2nd position than —OH group because —OCH_3 group is more electron rich and it allows the electrons to be donated easily but not capable of ionization. This observation has been further confirmed by the higher activity of "a-series" with —OCH_3 at the said position than "b-series" and comparable to the activity of the unmethylated "d-series" with —OH at the same position. Even n-butyl-derivative 17a was active (IC_{50} = 3.3 mM), in contrast to the unmethylated analogs 17b and 17d. The blue electrostatic contour near the 4th and 5th position indicates the presence of an electron deficient group is favorable at this position. Higher activities of "d-series" and "c-series" with —OH group at the 5th position and —OCH_3 at 4th position respectively support this result.

3.5. Molecular docking analysis

Interaction of benzoquinone derivative with 5-LOX was observed to get the view of ligand conformational change when undergoes docking. Since a co-crystal ligand was absent for 5-LOX, Prediction of the size and spatial orientation of the ligand binding sites of proteins was a major challenge. The active site of 5-LOX crystal structure was reported in literature characterized by an elongated cavity which is surrounded by 5-LOX specific amino acids Tyr 181, Ala 603, Ala 606, His 600 and Try 364 and all LOX conserved residues like Leu- 368, 373, 414, 607 and Ile-406 [45]. We also have confirmed the active site by identifying the possible cavity of 5-LOX with the help of Pocket-Cavity Search Application POCASA and optimized the dimension of the grid box [46].

Using the optimized grid box and through molecular docking process, the interaction between protein 5-LOX and benzoquinone derivatives were deduced in the form of binding affinity value given in Supporting information S4. The binding mode between the natural derivative of benzoquinone 'Embelin' (compound 2) and 5-LOX (Fig. 5) reveals that inhibitory mechanism of compounds is almost similar to the typical inhibitory mechanism of a good inhibitor for 5-LOX, which should have a polar head and a hydrophobic body.

The polar OH- and C=O groups at the benzoquinone head portion interact with the polar amino acids of 5-LOX by forming hydrogen bonds with His 600, Gln 363 and Leu 420 residues and can be seen in green dotted line. This observation is compatible with CoMFA electrostatic contours found around benzoquinone head portion indicating these regions are favorable may be due to its interaction with polar amino acids of the target protein. This compound also forms hydrophobic interactions with the protein through its non-polar long alkyl part with residues like Leu 368, Ile 415, Phe 421, Phe 359, Leu 414, Leu 607, Phe 177. These findings again support the CoMFA result which has shown the importance of long, bulky alkyl group at the position 3 for a potentially active ligand.

4. Conclusions

Here, we report the systematic investigation of 2D and 3D-QSAR on benzoquinone derivatives as 5-LOX inhibitors. The 2D descriptors were obtained from e-Dragon, PowerMV and Gaussian 09 software packages. The most relevant descriptors were filtered with the help Correlation feature selection (CFS) technique. Different 2D-QSAR methods such as RF, SVM and MLR were employed to formulate a statistically reliable relationship between 5-LOX activity and the descriptors. The consistency of the developed 2D-QSAR models was verified using the Q^2 and R^2_{pred} values and the result shows that among the three models constructed, RF model gave the better predictive power with an R^2 value of 0.93, Q^2 (LMO) value of 0.51 and R^2_{pred} value of 0.71. The 3D-QSAR based on CoMFA methodology was also applied to the same set of benzoquinone derivatives. Reasonable Q^2, R^2 and R^2_{pred} values were obtained for the CoMFA indicating that the model constructed has an excellent internal predictive power and good external predictive ability. The contour maps extracted gives an idea of the effects of steric and electrostatic fields around the aligned molecules on their 5-LOX inhibitory activity. The lowest energy binding pose obtained from docking analysis provides a qualitative representation of ligand and and protein interactions, which are Compactable with CoMFA maps and the 2D-QSAR model. Thus, both CoMFA and docking studies indicate the long, bulky alkyl residue at third position is very important for a potentially active ligand and the polar or hydrogen bonded interactions of benzoquinone's —OH group with polar amino acids of the target protein were also favorable for 5-LOX inhibition. 2D-QSAR also supports the results that the activity of the ligands was mainly depended on hydrogen bond donor/acceptor features and electron affinity values that might be responsible for its pharmacological action. Our findings would seem to show that RF and CoMFA models together could very profitable to uncover more 5-LOX active benzoquinone derivatives.

Conflict of interest

No conflict of interest to declare

Acknowledgments

The authors T.K. Shameera Ahamed and Vijisha K. Rajan express their sincere gratitude to Human Resource Development Group Council of Scientific & Industrial Research (CSIR), India and University Grants Commission (UGC), India, respectively for the financial support. The authors are thankful to the Central Sophisticated Instrumentation Facility (CSIF) of the University of Calicut for the Gaussian 09 software support.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fshw.2019.02.001.

References

[1] B.A. Jakschik, L.H. Lee, Enzymatic assembly of slow reacting substance, Nature 287 (1980) 51–52.
[2] C.D. Funk, Prostaglandins and leukotrienes: advances in eicosanoid biology, Science 294 (2001) 1871–1875.
[3] J.Z. Haeggström, C.D. Funk, Lipoxygenase and leukotriene pathways: biochemistry, biology, and roles in disease, Chem. Rev. 111 (2011) 5866–5896, http://dx.doi.org/10.1021/cr200246d.
[4] L.G. Melstrom, D.J. Bentrem, M.R. Salabat, T.J. Kennedy, X.Z. Ding, M. Strouch, S.M. Rao, R.C. Witt, C.A. Ternent, M.S. Talamonti, R.H. Bell, T.A. Adrian, Overexpression of 5-lipoxygenase in colon polyps and cancer and the effect of 5-LOX Inhibitors in vitro and in a murine model, Clin. Cancer Res. 14 (2008) 6525–6530, http://dx.doi.org/10.1158/1078-0432.CCR-07-4631.
[5] R. Hennig, P. Grippo, X.Z. Ding, S.M. Rao, M.W. Buchler, H. Friess, M.S. Talamonti, R.H. Bell, T.E. Adrian, 5-Lipoxygenase, a marker for early pancreatic intraepithelial neoplastic lesions, Cancer Res. 65 (2005) 6011–6016, http://dx.doi.org/10.1158/0008-5472.CAN-04-4090.
[6] J. Ghosh, C.E. Myers, Arachidonic acid stimulates prostate cancer cell growth: critical role of 5-lipoxygenase, Biochem. Biophys. Res. Commun. 235 (1997) 418–423, http://dx.doi.org/10.1006/bbrc.1997.6799.
[7] J. Ghosh, Inhibition of arachidonate 5-lipoxygenase triggers prostate cancer cell death through rapid activation of c-Jun N-terminal kinase, Biochem. Biophys. Res. Commun. 307 (2003) 342–349, http://dx.doi.org/10.1016/S0006-291X(03)01201-4.
[8] Y. Chen, Y. Hu, H. Zhang, C. Peng, S. Li, Loss of the Alox5 gene impairs leukemia stem cells and prevents chronic myeloid leukemia, Nat. Genet. 41 (2009) 783–792, http://dx.doi.org/10.1038/ng.389.
[9] R.N. Young, Inhibitors of 5-lipoxygenase: a therapeutic potential yet to be fully realized? Eur. J. Med. Chem. 34 (1999) 671–685, http://dx.doi.org/10.1016/S0223-5234(99)00225-1.
[10] O. Werz, D. Steinhilber, Development of 5-lipoxygenase inhibitors - Lessons from cellular enzyme regulation, Biochem. Pharmacol. 70 (2005) 327–333, http://dx.doi.org/10.1016/j.bcp.2005.04.018.
[11] S.E. Wenzel, A.K. Kamada, Zileuton: the first 5-lipoxygenase inhibitor for the treatment of asthma, Ann. Pharmacother. 30 (1996) 858–864.
[12] M.C. Liu, L.M. Dube, J. Lancaster, M. Blumenthal, P.B. Boggs, N. K, Clinical aspects of allergic disease Acute and chronic effects of a 5-1ipoxygenase inhibitor in asthma: a 6-month randomized multicenter trial, J. Allergy Clin. Immunol. 98 (1996) 859–871.
[13] D. Steinhilber, B. Hofmann, Recent advances in the search for novel 5-lipoxygenase inhibitors, Basic Clin. Pharmacol. Toxicol. 114 (2014) 70–77, http://dx.doi.org/10.1111/bcpt.12114.
[14] G. Eren, A. MacChiarulo, E. Banoglu, From molecular docking to 3D-quantitative structure-activity relationships (3D-QSAR): insights into the binding mode of 5-Lipoxygenase inhibitors, Mol. Inform. 31 (2012) 123–134, http://dx.doi.org/10.1002/minf.201100101.
[15] M.C. Sharma, Molecular modeling studies of substituted 3,4-dihydroxychalcone derivatives as 5-lipoxygenase and cyclooxygenase inhibitors, Med. Chem. Res. 23 (2014) 1797–1818, http://dx.doi.org/10.1007/s00044-013-0745-7.
[16] P. Aparoy, G.K. Suresh, K.K. Reddy, P. Reddanna, CoMFA and CoMSIA studies on 5-hydroxyindole-3-carboxylate derivatives as 5-lipoxygenase inhibitors : Generation of homology model and docking studies, Bioorg. Med. Chem. Lett. 21 (2011) 456–462, http://dx.doi.org/10.1016/j.bmcl.2010.10.119.
[17] Z. Ul-Haq, N. Khan, S.K. Zafar, S.T. Moin, Active site characterization and structure based 3D-QSAR studies on non-redox type 5-lipoxygenase inhibitors, Eur. J. Pharm. Sci. 88 (2016) 26–36, http://dx.doi.org/10.1016/j.ejps.2016.03.014.
[18] T.K. Shameera Ahamed, K. Muraleedharan, A ligand-based comparative molecular field analysis (CoMFA) and homology model based molecular docking studies on 3′, 4′-dihydroxyflavones as rat 5-lipoxygenase inhibitors: design of new inhibitors, Comput. Biol. Chem. 71 (2017) 188–200, http://dx.doi.org/10.1016/j.compbiolchem.2017.08.010.
[19] R. Filosa, A. Peduto, A.M. Schaible, V. Krauth, C. Weinigel, D. Barz, C. Petronzi, F. Bruno, F. Roviezzo, G. Spaziano, B.D. Agostino, M. De Rosa, O. Werz, Novel series of benzoquinones with high potency against 5- lipoxygenase in human polymorphonuclear leukocytes, Eur. J. Med. Chem. 94 (2015) 132–139, http://dx.doi.org/10.1016/j.ejmech.2015.02.042.
[20] C.J. Cramer, Essentials of Computational Chemistry Theories and Models, 2nd ed., John Wiley and Sons Ltd, 2004.
[21] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
[22] C. Lee, W. Yang, R.G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B 37 (1998) 785–789.
[23] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K.

Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09 Revision A.02, Gaussian, Inc., Wallingford CT, 2009.
[24] I.V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V.A. Palyulin, E.V. Radchenko, N.S. Zefirov, A.S. Makarenko, V.Y. Tanchuk, V.V. Prokopenko, Virtual computational chemistry laboratory - design and description, J. Comput. Aided Mol. Des. 19 (2005) 453–463, http://dx.doi.org/10.1007/s10822-005-8694-y.
[25] K. Liu, J. Feng, S.S. Young, PowerMV: a software environment for molecular viewing, descriptor generation, data analysis and hit evaluation, J. Chem. Inf. Model. 45 (2005) 515–522, http://dx.doi.org/10.1021/ci049847v.
[26] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software, ACM SIGKDD Explor. Newsl. 11 (2009) 10–18, http://dx.doi.org/10.1145/1656274.1656278.
[27] M.A. Hall, Correlation-based Feature Selection for Machine Learning, 1999.
[28] A.K. Saxena, P. Prathipati, Comparison of MLR, PLS and GA-MLR in QSAR analysis*, SAR QSAR Environ. Res. 14 (2003) 433–445, http://dx.doi.org/10.1080/10629360310001624015.
[29] C. Cortes, V. Vladmir, Support vector networks, Mach. Learn. 20 (1995) 273–297, http://dx.doi.org/10.1007/BF00994018.
[30] J.P. Doucet, F. Barbault, H. Xia, A.P, B. Fan, Nonlinear SVM approaches to QSPR/QSAR studies and drug design, Curr. Comput. Drug Des. 3 (2007) 263–289, http://dx.doi.org/10.2174/157340907782799372.
[31] J. Platt, Fast training of support vector machines using sequential minimal optimization, Adv. Kernel Methods (1998) 41–64.
[32] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, http://dx.doi.org/10.1023/A:1010933404324.
[33] P. Tosco, T. Balle, F. Shiri, Open3DALIGN: an open-source software aimed at unsupervised ligand alignment, J. Comput. Aided Mol. Des. 25 (2011) 777–783, http://dx.doi.org/10.1007/s10822-011-9462-9.
[34] P. Tosco, T. Balle, Open3DQSAR: a new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields, J. Mol. Model. 17 (2011) 201–208, http://dx.doi.org/10.1007/s00894-010-0684-x.
[35] M. Pastor, G. Cruciani, S. Clementi, Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships, J. Med. Chem. 40 (1997) 1455–1464, http://dx.doi.org/10.1021/jm9608016.
[36] S.C. Massimo Baroni, Gabriele Costantino, Gabriele Cruciani, Daniela Riganelli, Roberta Valigi, Generating optimal linear PLS estimations (GOLPE): an advanced chemometric tool for handling 3D-QSAR problems, Quant. Struct. Relat. 12 (1993) 9–20, http://dx.doi.org/10.1002/qsar.19930120103.
[37] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. Syst. 58 (2001) 109–130, http://dx.doi.org/10.1016/S0169-7439(01)00155-1.
[38] A. Golbraikh, A. Tropsha, Beware of q2!, J. Mol. Graph. Model. 20 (2002) 269–276, http://dx.doi.org/10.1016/S1093-3263(01)00123-1.
[39] D.M. Hawkins, S.C. Basak, D. Mills, Assessing model fit by cross-validation, J. Chem. Inf. Comput. Sci. 43 (2003) 579–586.
[40] O. Trott, A. Olson, AutoDock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461, http://dx.doi.org/10.1002/jcc.21334.
[41] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786, http://dx.doi.org/10.1021/ci200227u.
[42] G. Morris, R. Huey, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791, http://dx.doi.org/10.1002/jcc.21256.AutoDock4.
[43] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF chimera - a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612, http://dx.doi.org/10.1002/jcc.20084.
[44] W.L. Delano, The PyMOL Molecular Graphics System, 2002.
[45] A.R. Brash, M.E. Newcomer, The structure of human 5-lipoxygenase, Science 331 (2011) 217–219, http://dx.doi.org/10.1126/science.1197203.
[46] T.K. Shameera Ahamed, V.K. Rajan, K. Sabira, K. Muraleedharan, QSAR classification-based virtual screening followed by molecular docking studies for identification of potential inhibitors of 5-lipoxygenase, Comput. Biol. Chem. 77 (2018) 154–166, http://dx.doi.org/10.1016/j.compbiolchem.2018.10.002.