



Combined Machine Learning and GRID-Independent Molecular Descriptor (GRIND) Models to Probe the Activity Profiles of 5-Lipoxygenase Activating Protein Inhibitors

Hafiza Aliza Khan and Ishrat Jabeen*

Research Centre for Modelling and Simulation (RCMS), NUST Interdisciplinary Cluster for Higher Education (NICHE), National University of Sciences and Technology (NUST), Islamabad, Pakistan

OPEN ACCESS

Edited by:

Adriano D. Andricopulo,
University of Sao Paulo, Brazil

Reviewed by:

Sajjad Gharaghani,
University of Tehran, Iran
Eslam Pourbasheer,
University of Mohaghegh Ardabili, Iran

*Correspondence:

Ishrat Jabeen
ishrat.jabeen@rcms.nust.edu.pk

Specialty section:

This article was submitted to
Experimental Pharmacology and Drug
Discovery,
a section of the journal
Frontiers in Pharmacology

Received: 30 November 2021

Accepted: 03 January 2022

Published: 01 March 2022

Citation:

Khan HA and Jabeen I (2022)
Combined Machine Learning and
GRID-Independent Molecular
Descriptor (GRIND) Models to Probe
the Activity Profiles of 5-Lipoxygenase
Activating Protein Inhibitors.
Front. Pharmacol. 13:825741.
doi: 10.3389/fphar.2022.825741

Leukotrienes (LTs) are pro-inflammatory lipid mediators derived from arachidonic acid (AA), and their high production has been reported in multiple allergic, autoimmune, and cardiovascular disorders. The biological synthesis of leukotrienes is instigated by transfer of AA to 5-lipoxygenase (5-LO) via the 5-lipoxygenase-activating protein (FLAP). Suppression of FLAP can inhibit LT production at the earliest level, providing relief to patients requiring anti-leukotriene therapy. Over the last 3 decades, several FLAP modulators have been synthesized and pharmacologically tested, but none of them could be able to reach the market. Therefore, it is highly desirable to unveil the structural requirement of FLAP modulators. Here, in this study, supervised machine learning techniques and molecular modeling strategies are adapted to vaticinate the important 2D and 3D anti-inflammatory properties of structurally diverse FLAP inhibitors, respectively. For this purpose, multiple machine learning classification models have been developed to reveal the most relevant 2D features. Furthermore, to probe the 3D molecular basis of interaction of diverse anti-inflammatory compounds with FLAP, molecular docking studies were executed. By using the most probable binding poses from docking studies, the GRIND model was developed, which indicated the positive contribution of four hydrophobic, two hydrogen bond acceptor, and two shape-based features at certain distances from each other towards the inhibitory potency of FLAP modulators. Collectively, this study sheds light on important two-dimensional and three-dimensional structural requirements of FLAP modulators that can potentially guide the development of more potent chemotypes for the treatment of inflammatory disorders.

Keywords: 5-lipoxygenase activating protein (FLAP) inhibitors, machine learning, molecular docking, grind, leukotrienes (LTs)

1 INTRODUCTION

The 5-LO pathway is responsible for the biological synthesis of leukotrienes (LTs) using arachidonic acid (AA) predominately by inflammatory cells like polymorphonuclear leukocytes, activated macrophages, and mast cells upon arrival of immunologic and non-immunologic stimuli (Hedi and Norbert 2004). Activation of leukocytes results in translocation of cytosolic protein

phospholipase A2 (PLA2) to membrane where it selectively hydrolyzes the sn-2 acyl bond of membrane phospholipids to release AA and lysophosphatidic acid. An integral membrane protein called FLAP (5-lipoxygenase activating protein) uptakes the AA and efficiently transfers it to the active site of 5-lipoxygenase (5-LO) enzyme, which catalyzes a series of reactions at a single active site (Peters-Golden 1998; Peters-Golden and Brock 2003). In the first step, in a calcium- and ATP-dependent reaction, AA is converted to a 5-lipoxygenase-specific hydroperoxide intermediate (5-HPETE), while in the second step, 5-LO performs synthase reaction for conversion of 5-HPETE to the epoxide intermediate, leukotriene A4 (LTA4) (Woods et al., 1995; Smyrniotis et al., 2014). LTA4 acts as a common precursor for biosynthesis of chemoattractant leukotriene B4 (LTB4) by a zinc-bound LTA4 hydrolase (LTA4H) and bronchoconstrictive cysteinyl leukotrienes (CysLTs or LTC4) with the help of LTC4 synthase (LTC4S) (Jakschik and Kuo 1983; Haeggström 2000). Both LTB4 and CysLTs are physiologically active final products of the 5-LO pathway and are exported out of the cell through specific transport proteins while extracellular peptidases metabolize LTC4 to LTD4, which is converted into LTE4 depending on type of inflammatory signal and cell demand (Jedlitschky and Keppler 2002). After export, LTs bind with respective G-protein-coupled receptors, e.g., LTB4 binds with BLT1 and BLT2, whereas CysLTs activate CysLT1 and CysLT2 receptors to incite further proinflammatory signaling cascades (Ghosh et al., 2016).

Since high levels of LTs have been reported in the pathophysiology of a wide range of inflammatory, cardiovascular, and autoimmune disorders, FLAP has become the focus of immense research because LT production can be stopped at the earliest level (Folco et al., 2000; Liu and Yokomizo 2015). Over the course of the last 3 decades, several FLAP modulators have been proposed including first generation of derivatives of indoles and quinolines for asthma treatment (Evans et al., 1991; Frenette et al., 1999). These inhibitors such as MK-886, MK-591, and BAY-X-1005 demonstrated efficiency in clinical trials in patients with inflammatory diseases in the mid-1990s but were not brought to market due to poor pharmacokinetics (Friedman et al., 1993; Diamant et al., 1995; Dahlén et al., 1997). Revelation of SAR data along with crystal structure expedites the drug discovery quest against FLAP, leading to the second generation consisting of derivatives of diarylalkanes, biaryl amino-heteroarenes, and benzimidazoles, proposed with renewed interest for treatment of cardiovascular diseases (Lemurell et al., 2015; Macdonald et al., 2008; N.; Stock et al., 2010). Moreover, several inhibitors proved to be promising readouts for preclinical and clinical studies such as AM103, AM803, BI665915, AZD5718, and AZD6642 and have been shown to ameliorate inflammation-related diseases (Bain et al., 2010; Lorrain et al., 2010; Antoniu, 2014; Ericsson et al., 2020). However, despite several practices, not a single inhibitor has won the race to the market as a drug to date. Therefore, development of more potent chemical entities against FLAP is highly desirable to provide relief to patients suffering from inflammatory disorders.

Mostly FLAP modulators were synthesized and pharmacologically tested and optimized through SAR (structure–activity relationship) studies. Some candidates were also identified by virtual screening from a ligand-based pharmacophore built upon smaller datasets (Banoglu et al., 2012; Temml et al., 2017; Olgac et al., 2020). Here, in this study, advanced machine learning (ML) techniques along with classical modeling strategies are adapted to shed light on important 2D and 3D anti-inflammatory properties of a diverse set of inhibitors targeting FLAP. For this reason, ML models based on most relevant 2D descriptors or features have been constructed. Further molecular docking was performed to establish a binding hypothesis of each class of inhibitors within the FLAP binding cavity followed by common scaffold clustering to obtain the most probable 3D binding solutions. The most probable 3D binding poses were utilized for GRID-independent molecular descriptor analysis (GRIND) to probe the important 3D binding features and associated mutual distances in active FLAP modulators.

2 MATERIALS AND METHODS

2.1 Machine Learning Modeling

2.1.1 Dataset Preparation

All compounds having activity values in IC_{50} against FLAP were retrieved from the ChEMBL database under target ID ChEMBL4550 followed by removal of compounds with similar canonical smiles resulting in a dataset of 658 compounds. The IC_{50} of the finalized 658 compounds ranged from 0.3 to 22,500 nM. Furthermore, the highly active and least active compounds were distinguished by the application of activity threshold, i.e., compounds having $IC_{50} < 10$ nM were categorized as highly active while compounds having $IC_{50} > 70$ nM were categorized as least active considering that FLAP inhibitors that have entered clinical trials usually possess values < 10 nM (Gür, Çalıřkan, and Banoglu 2018). Compounds with IC_{50} values in between >10 nM and <70 nM were labeled as intermediates and were removed. For ML classification model development, highly active compounds were labeled as one, while least active ones were labeled as 0. The final dataset was composed of 503 (253 highly actives and 250 least actives) compounds and was randomly divided into a training set (402 compounds: 201 highly actives and 201 least actives) and a test set (101 compounds: 52 highly actives and 49 least actives) by a ratio of 80% and 20% respectively using *train_test_split* function (*random_state* = 42) of model_selection library from the scikit-learn Python package (Pedregosa et al., 2012). Additionally, it was ensured that the ratio of the highly active to weakly active inhibitors remained equal in the training and test set.

2.1.2 Computation of 2D Chemical Descriptors

Initially, 4,179 2D descriptors were calculated using alvaDesc tool version 2.0.8 (Mauri 2020). The descriptors can be divided into 21 categories named constitutional indices, ring descriptors, topological indices, walk and path counts, connectivity indices,

information indices, 2D-matrix based descriptors, 2D autocorrelations, burden eigenvalues, P_VSA like descriptors, ETA indices, edge adjacency indices, fractional group counts, atom-centered fragments, atom-type-estate indices, pharmacophore descriptors, 2D atom pairs, charge descriptors, molecular properties, drug-like indices, MDE descriptors, and chirality descriptors. Descriptors with null values and variance near zero were removed. For the remaining 2,352 descriptors, Pearson autocorrelation coefficient was calculated and autocorrelated descriptors along with low dependency (correlation) on the target variable (inhibitory potency, IC₅₀) were discarded, resulting in a set of 442 descriptors. The final 442 features were further applied to train the ML models.

2.1.3 Machine Learning Modeling

For this study, six supervised ML classification models named support vector machine (SVM), random forest (RF), multilayer perceptron (MLP), decision tree (DT), logistic regression (LR), and gradient boost decision tree (GBDT) were developed. SVM, RF, MLP, DT, and LR were generated using the scikit-learn Python package whereas the GBDT was built by the XGBoost Python package (Cortes, Vapnik, and Saitta 1995; Liaw and Wiener 2002; Haykin, 2009; Quinlan 1986; T.; Chen and Guestrin 2016; McCullagh and Nelder 1989). To select the most relevant features from the set of 442 descriptors, the RFECV (Recursive Feature Elimination and Cross-Validation Selection) algorithm of scikit-learn was used (Guyon et al., 2002). Recursive feature elimination (RFE) is a wrapper-type feature selection that works by eliminating *n* features from a model by fitting the model multiple times and, at each step, removing the weakest features, determined by either the *coef_* (SVM and LR) or *feature_importances_* (RF, DT, and XGBoost) attribute of the fitted model (Guyon et al., 2002). Since there is no attribute available to estimate feature importance in MLP, XGBoost was used as the base estimator. The cross-validation (*cv*) parameter of RFECV was set at fivefold and was done by using the RepeatedStratifiedKFold method of the model_selection library from scikit-learn (Zeng and Martinez 2010). The GirdSearchCV library in scikit-learn was used to tune hyperparameters of the estimators based on a 10-fold cross-validation Matthews Correlation Coefficient (MCC). This process was repeated ten times. Moreover, to assemble data transformer (RFECV) and hyperparameter tuner (GirdSearchCV) with simultaneous cross-validation while setting different parameters, the pipeline module of scikit-learn was used.

An SVM constructs a maximum marginal hyperplane with the help of a kernel function to map the non-linear problem in multidimensional space for data separation. The performance of the SVM model is controlled by parameters such as kernel, capacity parameter (*C*), and gamma. Kernel represents sample distribution in the mapping space, *C* controls the trade-off between smooth decision boundary, and gamma controls the extent of curvature in decision boundary (Nekoei, Mohammadhosseini, and Pournasheer 2015; Pournasheer et al., 2017). For this project, linear kernel was utilized while all parameters were set at their default values except for tuning of

penalty parameter (*C*) (Chang and Lin, 2021). MLP is a feedforward artificial neural network and is trained using back propagation algorithm. It has an activation function that forms a linear combination according to weights of inputs to decide the output. The MLP model was controlled by tuning the following parameters: the number of neurons (*hidden_layer_sizes*) and activation function (*activation*), while the rest of the parameters were set at their default values (Glorot and Bengio, 2021). An LR model predicts a dependent data variable by analyzing the relationship through logic functions between one or more existing independent variables. It was controlled by tuning the following parameters: the way of regularization (*penalty*), strength of regularization (*C*), tolerance for stopping criteria (*tol*), and algorithm of optimization (*solver*), whereas other parameters not mentioned were set at their default values (Fan et al., 2008). A DT classifies data by splitting them into source nodes and then multiple child nodes using statistical probability. The DT model was optimized by tuning the following parameters: quality of split (*criterion*), split at each node (*splitter*), and number of features for the best split (*max_features*). The remaining parameters were set as their default values (Breiman and Olshen 2012). An RF builds multiple decision trees and merges them together to get an accurate and stable prediction. The RF model was controlled by tuning the following parameters: number of trees (*n_estimators*), quality of a split (*criterion*), features for the best split (*max_features*), and the minimum number of samples required for splitting (*min_samples_split*); the other parameters not mentioned were set at their default values (Breiman 2001). XGBoost is an ensemble tree method that applies the principle of boosting weak learners using the gradient descent architecture. For this project, gradient boost tree (GBDT) has been implemented, which uses decision trees as weak classifiers. The XGBoost model was controlled by tuning the following parameters: the maximum depth of a tree (*max_depth*), the number of the tree (*n_estimators*), minimum loss reduction required for partition on a node (*gamma*), minimum sum of instance weight needed to generate a child node (*min_child_weight*), strength of L1 regularization (*reg_alpha*), and learning rate (*learning_rate*). The other parameters not mentioned were set at their default values (T. Chen and Guestrin 2016).

The repeated stratified 5-fold cross-validation was used on the training set to select and evaluate the robustness of models, and the test set was used to evaluate the performance of models. Evaluation parameters include classification accuracy (ACC), true positive rate or sensitivity (SE), true negative rate or specificity (SP), and Matthews correlation coefficient (MCC) as mentioned in (Eqs 1–4) below:

$$\text{True Positive Rate (Sensitivity)} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{True Negative Rate (Specificity)} = \frac{TN}{TN + FP} \quad (2)$$

$$\text{Classification Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\begin{aligned} & \text{Matthews Correlation Coefficient (MCC)} \\ & = \frac{(TP \times TN - FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4) \end{aligned}$$

2.2 Molecular Modeling

2.2.1 Calculation of Lipophilic Efficiency (LipE) and cLogP

To estimate the druglikeness of the initially finalized 658 FLAP inhibitors (section 2.1.1), LogP was calculated by using Bio-Loom software (BioByte - Bio-Loom, 2021) followed by computation of LipE with the following equation:

$$\text{LipE} = \text{pIC}_{50} - \text{cLogP} \quad (5)$$

Briefly, lipophilicity or cLogP strongly impacts membrane passive permeability, which is required for oral absorption and access of the drug to intracellular compartments and tissue penetration (Arnott and Planey 2012). Lipophilic efficiency (LipE) is defined as normalization of the pIC₅₀ with respect to cLogP of the compound. Previously, Leeson et al. proposed that an ideal drug candidate should have a LipE value greater than five, which is obtained in case of high potency and low lipophilicity (Leeson and Springthorpe 2007). For the application of molecular modeling techniques (Docking and 3D QSAR GRIND), LipE and cLogP filter were used; i.e., compounds having LipE greater than one and cLogP greater than two were selected. The new dataset of compounds having LipE value greater than one and cLogP greater than two was divided into a training set (80%) and a test set (20%) by using the *train_test_split* function (*random_state* = 42) of the *model_selection* library from the *scikit-learn* Python package. Both training and test datasets were further employed in molecular modeling studies (docking studies and GRIND modeling).

2.2.1 Molecular Docking and Pose Analysis

To explore the binding interactions of structurally diverse FLAP inhibitors, and to obtain the most probable 3D binding conformations of ligands for GRIND analysis, inhibitors having LipE value greater than one and cLogP greater than two were docked into the binding pocket of the FLAP structure retrieved from the Protein Data Bank (PDB ID: 2Q7M) (Ferguson et al., 2007). Protein structure was prepared by energy minimization through the Amber99 force field of MOE (A. A. Chen and Pappu 2007). The energy-minimized structure was imported into GOLD software (version 5.6.1) (Jones et al., 1997) followed by determination of x, y, z coordinates around the single-solvent accessible point present in the center of the active site. The binding site area was kept at 12 Å radius, which included all important amino acid residues reported by previous studies. A total of 100 conformations for each ligand were generated, and GOLD fitness scoring function was used to rank each pose of ligands with subsequent energy minimization of ligand-protein docking complexes using LigX implemented in software MOE. Gold score fitness scoring function was calculated as:

$$\text{Fitness} = S_{(hb)_{ext}} + 1.3750 * S_{(vdw)_{ext}} + S_{(hb)_{int}} + 1.0000 * S_{(int)} \quad (6)$$

Based on structural similarity, common scaffold clustering (CSC) as proposed by Jabeen et al. (2012) was conducted to reduce the conformational space. For this purpose, RMSD matrix was generated through agglomerative hierarchical cluster analysis, and clusters with maximum docked ligands were selected for ligand-protein interaction profiling. Common interactions between each class were sorted out and binding hypothesis was generated for each class with respect to interaction pattern and position in binding pocket. Conformations from selected clusters were further utilized in GRIND analysis as training set.

2.2.3 Grid Independent Molecular Descriptors Analysis

Selected 3D molecular confirmations of ligands obtained from clusters containing maximum docked ligands along with their inhibitory potencies (pIC₅₀) were imported in Pentacle software version 1.06 to construct the GRIND model (Pastor et al., 2000). Calculation of molecular interaction fields (MIFs) was done by use of different probes, namely, N1, O, DRY, and TIP, where N1 (amide N) represents a hydrogen bond donor, O (sp² carbonyl O) denotes a hydrogen bond acceptor, DRY indicates a hydrophobic region, and TIP stands for steric hotspots within the virtual receptor site. A GRID was used to iteratively place these probes to calculate the total energy by addition of Lennard-Jones potential energy (Elj), hydrogen bond energies (Ehb), and the electrostatic energy (Eel), whereas with the help of the following equation, total interaction energy at each node was calculated:

$$E_{xyz} = \sum E_{hb} + \sum E_{lj} + \sum E_{el} \quad (7)$$

AMANDA algorithm was used to extract the most relevant and significant MIFs along with evaluation of structural characteristics of the dataset explained by GRIND descriptors (Durán, Martínez, and Pastor 2008). The default GRID space of 0.5 and the energy cutoff values, which are -4.2, -2.6, -0.5, and -0.74 for N1, O, DRY, and TIP, respectively, were used for discretization of MIFs, while nodes that did not meet the energy cutoff were discarded. The next encoding step involves CLACC algorithm that aided in selection of consistent nodes by adjustment of compounds according to their moment of inertia. The values obtained from encoding consist of a consistent set of variables whose values were directly represented in the form of correlogram plots. The final GRIND model with PLS (partial least square) analysis using LOO (leave one-out) method with statistically significant *R*², *q*², and standard error values (SDEP) was built on the training set followed by evaluation with the test set (section 2.2.1). Additionally, *r*²_m metrics (*r*²_m, Delta *r*²_m) was also generated for validation purposes according to the previously published studies (Roy et al., 2013; Gajo et al., 2016).

TABLE 1 | The layout of prediction performances of machine learning models assessed by stratified 5-fold cross-validation for the training set and test set.

Classifier	Training set (n = 402)				CV5 of training set (n = 402)				Test set (n = 101)			
	SE	SP	ACC	MCC	SE	SP	ACC	MCC	SE	SP	ACC	MCC
XGBoost (GBDT)	0.99	0.99	0.99	0.98	0.91	0.89	0.90	0.81	0.89	0.91	0.90	0.80
Random forest (RF)	0.99	1.00	1.00	0.99	0.85	0.90	0.87	0.75	0.94	0.88	0.91	0.82
Decision tree (DT)	0.88	0.94	0.91	0.82	0.83	0.83	0.83	0.66	0.83	0.84	0.84	0.68
Support vector machine (SVM)	0.96	0.98	0.97	0.93	0.84	0.77	0.80	0.61	0.75	0.80	0.78	0.56
Logistic regression (LR)	0.82	0.88	0.85	0.69	0.84	0.75	0.79	0.59	0.85	0.87	0.86	0.72
Multilayer perceptron (MLP)	0.79	0.81	0.80	0.60	0.72	0.78	0.75	0.50	0.70	0.71	0.70	0.40

3 RESULTS

3.1 Machine Learning Models

Six ML models were developed by different algorithms (SVM, LR, MLP, DT, RF, and GBDT) using two-dimensional structural features of FLAP inhibitors. The performance of these models on 5-fold repeated stratified cross-validation is explained in **Table 1**. The cross-validation accuracy of the training set ranged between 0.90 and 0.75, and the MCC ranged from 0.81 to 0.50. The prediction accuracy and MCC values of the test set ranged from 0.90 to 0.70 and 0.80 to 0.40, respectively. MCC is often used as a measure of quality of binary classification models. Two models (XGBoost and RF) exhibited an MCC value of >0.7 on training and test sets, which means these two algorithms have a relatively good ability to predict whether a compound was a highly active or a least active FLAP modulator. In terms of the best model, XGBoost outperformed all and the accuracy and MCC values were observed as 0.90 and 0.81, respectively. Additionally, a previous fingerprint-based ML study on FLAP modulators stated that the reliability of predicted results depends mainly on the compounds themselves rather than algorithms or fingerprints (Tu et al., 2020).

The lowest performance was shown by the MLP model with an accuracy value of 0.75 and an MCC value of 0.50. For the best model (XGBoost), RFECV curve jumps to a maximum accuracy when the 46 informative features are captured with feature importance values ranging from 0.01 to 0.4. These 46 features mainly belong to eight descriptor categories named topological indices, 2D matrix-based descriptors, 2D autocorrelations, P_VSA-like descriptors, edge adjacency indices, atom-type E-state indices, pharmacophore descriptors, and molecular properties descriptors. All 46 captured features of the best-performing model (XGBoost) along with description and feature importance values are given in **Supplementary Table S1**. Additionally, 84, 126, 89, and 90 features have been captured by RF, DT, SVM, and LR, respectively, and RFECV curves for all models with optimal number of selected features are illustrated in **Supplementary Figure S1**. We anticipate that these 46 2D descriptors have the largest impact to differentiate between highly active and least active FLAP inhibitors. Additionally, the tuned hyperparameters for each model can be found in **Supplementary Table S2**.

3.2 LipE and cLogP Calculation

LipE and cLogP demonstrate the druggability of a compound in lead optimization programs to evaluate the potential for better *in*

vivo efficacy and safety. A graph between pIC₅₀ and cLogP along with LipE values of the compounds in the training set is shown in **Supplementary Figure S2**. In the current dataset of 658 FLAP inhibitors, only 238 compounds out of 658 demonstrated LipE value greater than five, which is the optimal threshold with cLogP values in the range of -0.27 to 3.78. Moreover, only 136 compounds showed a cLogP value between optimal range of 2–3.5 as proposed by Leeson and Springthorpe (2007). Additionally, 349 compounds out of 658 exhibited values of LipE less than 1 (cLogP = 4.3–10.19) while the cLogP range for 309 compounds having a LipE value greater than one was observed as 0.27–7.88. Interestingly, several potent FLAP inhibitors such as MK-886 (pIC₅₀ = 8.65 cLogP = 8.58, LipE = 0.07), MK-591 (pIC₅₀ = 9.30 cLogP = 8.82, LipE = 0.48), AM-643 (pIC₅₀ = 8.69, cLogP = 7.72, LipE = 0.97), AM-679 (pIC₅₀ = 8.65, cLogP = 7.98, LipE = 0.67), AM-803 (pIC₅₀ = 8.53, cLogP = 8.97, LipE = 0.43), and BRP-7 (pIC₅₀ = 6.50, cLogP = 7.23, LipE = 0.72) displayed significantly low values of LipE. It seems that increase in potency of these compounds might be due to increase in lipophilicity. On the other hand, other FLAP modulators such as BI665915 (pIC₅₀ = 8.76, cLogP = 2.14, LipE = 6.62) and AZD6642 (pIC₅₀ = 8.31, cLogP = 1.72, LipE = 6.62) showed relatively high values of LipE.

Herein, a dataset of 187 compounds having LipE value greater than one and cLogP greater than two was selected for further application of molecular modeling studies as all FLAP inhibitors in clinical trials possess high values of lipophilicity (cLogP). The dataset of 187 compounds was subsequently divided into a training set (151 compounds, **Supplementary Table S3**) and a test set (36 compounds, **Supplementary Table S4**). Docking-guided GRIND analysis was performed on the training set followed by evaluation of the final GRIND model with the test set.

3.3 Molecular Docking and SAR-Guided Pose Analysis

The selected dataset of 187 compounds mainly consists of already published indoles, biaryl bicycloheptanes, oxadiazole, and benzimidazole-based compounds. The dataset was further divided into a training set (151 compounds) and a test set (36 compounds) and based on common scaffolds; the training set was classified into six distinct classes. Common scaffold along with activity, lipophilicity, and lipophilic efficiency ranges of the six classes is depicted in **Figure 1**. Furthermore, a binding hypothesis of each class within the FLAP binding cavity was established. The

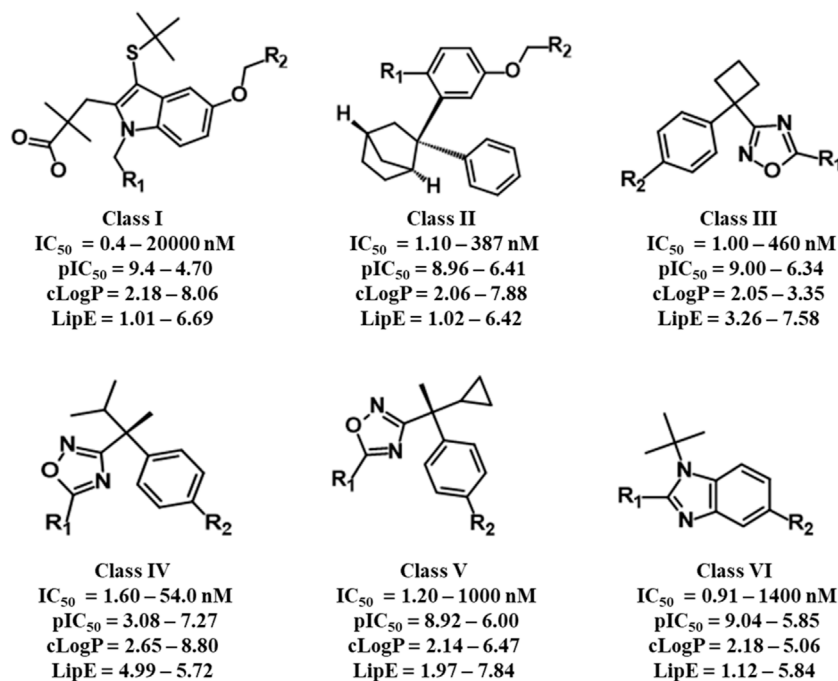


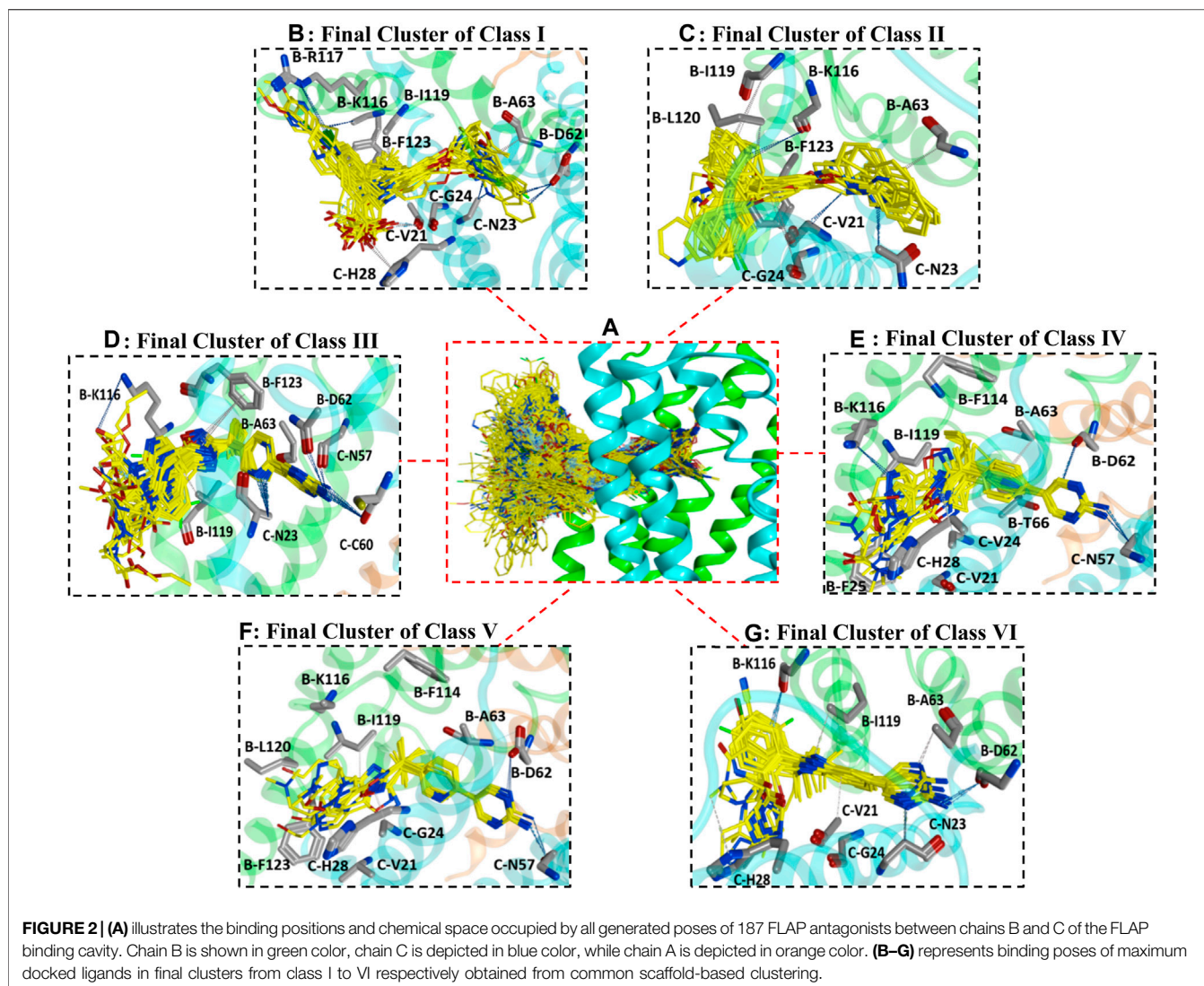
FIGURE 1 | Common scaffolds of six classes of FLAP inhibitors used for common scaffold clustering to obtain the most probable 3D binding poses for employment in GRIND studies.

distributions of compounds in each class along with common scaffolds are depicted in **Supplementary Table S5**.

All datasets of the selected 187 compounds were docked into the FLAP binding pocket, which included an area of 12 \AA selected by assigning x (65.7018), y (58.7512), and z (36.4565) coordinates between chains B and C near previously known interacting amino acid residues (B-F123, B-L120, B-I119, B-R117, B-K116, B-G115, B-F114, B-I113, B-Y112, B-T66, B-A63, B-D62, C-V61, C-C60, C-Q58, C-N57, C-H28, C-A27, C-F25, C-G24, C-N23, and C-V21) (Mancini et al., 1994; Ma et al., 2008). To remove any biases in the docking protocol, 100 poses per ligand were generated using the GOLD score fitness function. Further docking solutions were inspected by algoromatics hierarchical cluster analysis based on root mean square deviation (RMSD) at 3.5 \AA of the heavy atoms around a common scaffold. To follow the idea of similar binding mode for similar compounds, only those clusters that comprised the maximum number of docked ligands were selected. Overall, one cluster of binding conformations of compounds in all classes have been identified that contained the maximum number of docked ligands. The final selected cluster of each class and details of common scaffold clustering are depicted in **Supplementary Table S5**. Briefly, 26 out of 32 compounds for class I, 12 out of 20 for class II, 32 out of 35 for class III, 10 out of 10 for class IV, 13 out of 18 for class V, and 26 out of 36 for class VI were clustered out. Interestingly, the binding position of all final clusters was the same, and they bind between helix 4 ($\alpha 4$) and helix 2 ($\alpha 2$) of chain B and helix 1 ($\alpha 1$) and helix 2 ($\alpha 2$) of chain C, but a distinct binding pattern was observed for each class. The

binding region between chains B and C occupied by all generated poses of 187 ligands is shown in **Figure 2A**.

Briefly, class I compounds are derivatives of indole with dimethyl butanoic acid and S-tert-butyl substituents at positions two and three, respectively, as displayed in **Figure 1**, while R_1 at position one and R_2 at position five are generally occupied by heterogeneous 6-membered cyclic rings. The binding solutions for final cluster (cluster 1, **Supplementary Table S5**) of compounds in class I showed that dimethylbutanoic acid makes π -H-bond interactions with C-H28 and C-Val21, S-tert-butyl makes π -H-bond interactions with B-L120 and B-F123, while the indole scaffold is primarily involved in making π -H-bond interactions with C-G24 (**Figure 2B**). The R_1 substituents show hydrogen bonding with B-D62 and C-N23 and π -H interactions contact with B-A63 and C-N23, whereas N of the pyridine ring of R_2 shows a strong hydrogen bond with B-A63 while R_2 substituents show hydrogen bonding with B-R117 and B-K116 (**Figure 2B**). Overall, compounds of class I displayed a positive trend ($R^2 = 0.57$) between lipophilicity and inhibitory potency (**Supplementary Figure S3**) and exhibit a distinct SAR pattern. For instance, compound **1** ($IC_{50} = 0.4 \text{ nM}$, **Supplementary Table S3**) having the highest activity value ($cLogP = 8.06$, $LipE = 1.34$) among all the datasets contains 5-methylpyridine at R_1 and para-fluoro-2-phenylpyridine at R_2 as depicted in **Supplementary Table S6**. The final docking solution of compound **1** reveals that the pyridine ring present at R_1 shows a π -H-bond bonding interaction with the $-NH_2$ group of B-R117 (**Figure 4**). Compound **98** ($IC_{50} = 9.0 \text{ nM}$, **Supplementary Table S6**) has a similar structure to compound **1** except for the absence



of terminal 5-methyl on R_1 and the absence of fluorine on R_2 , rendering it low lipophilic ($c\text{LogP} = 3.35$, $\text{LipE} = 4.69$) and less active.

A study by Stock et al. also established that terminal 5-methyl on pyridine at R_1 significantly increases the inhibitory potency of compounds against FLAP (N. Stock et al., 2010; N. S. Stock, 2011). Interestingly, the hydrogen bonding between $-\text{NH}_2$ of B-R117 and nitrogen of the pyridine ring of R_2 of compound **98** has also been observed in the final docking solutions. However, the pyridine ring at R_1 did not seem to be involved in making any clear interactions. It was observed that the absence of terminal methyl on the pyridine ring of R_1 in compound **98** might reduce the exposure of pyridine ring to amino acids inside the FLAP binding cavity (Figure 3), leading to a substantial decrease in inhibitory potency of compound **98**. Moreover, the high LipE value of compound **98** as compared to compound **1** could be attributed only to its low logP (*o/w*) without an increase in biological activity.

Class II of FLAP antagonists contains 2,2-biaryl bicycloheptane as a common scaffold having diverse substituents at position 2 (R_1) and quinoline moiety at position 5 (R_2) of the exo aryl group (Figure 1). Ligand–protein interaction profiling of the final cluster (cluster 3, Supplementary Table S5) shows that the common scaffold orients itself towards the outer side-facing membrane and makes π -H interactions with B-L120, B-I119, and C-V21, whereas the quinoline moiety occupied the inside of the FLAP binding cavity and shows π -H interactions with B-A63, C-G24, and C-N23 (Figure 2D). Generally, R_1 is involved in making hydrogen bonds with B-F123 and B-K116 amino acid residues. Overall, a slight positive correlation ($R^2 = 0.27$) has been observed between inhibitory potency and lipophilicity for class III (Supplementary Figure S3). Moreover, a distinct SAR pattern was observed among compounds of class III. For instance, compound **10** ($\text{IC}_{50} = 1.1 \text{ nM}$, Supplementary Table S3), being the most potent and lipophilic ($c\text{LogP} = 7.88$, $\text{LipE} = 1.07$) member of this class, possesses oxadiazole-2-thione at R_1 ,

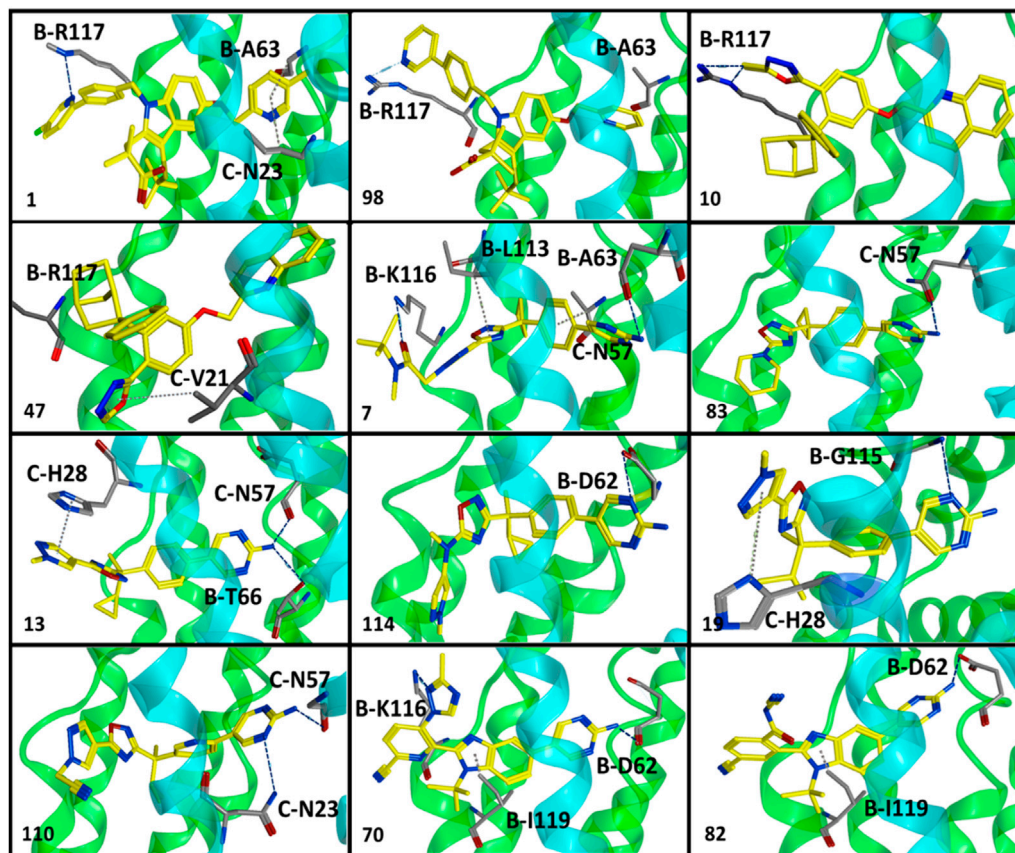


FIGURE 3 | Optimal binding poses of compounds displaying a distinct SAR pattern from all six classes of FLAP modulators. These poses were obtained from clusters with maximum docked ligands (common scaffold clustering) and were further employed for GRID-independent molecular descriptor (GRIND) analysis. Chain B is shown in green while chain C is depicted in blue color.

and absence of thiol substituent of oxadiazole at R_1 (**Supplementary Table S6**) resulted in approximately threefold decrease in inhibitory potency of compound **47** ($IC_{50} = 2.9$ nM, $cLogP = 7.37$, $LipE = 1.16$). The lipophilicities and $LipE$ values of compounds **10** and **47** are relatively the same and the difference in inhibitory potencies might be due to a distinct binding pattern. The final docking solution of compound **10** reveals the presence of two hydrogen bonds between the terminal sulfur of the oxadiazole-2-thione group at R_1 and $-NH_2$ of B-R117 (**Figure 3**). In compound **47**, only a π -H-bond interaction was found between the oxadiazole ring of R_1 and C-V21 that might be not very favorable and contribute to its low inhibitory. The positive contribution of negative ionizable moieties at the oxadiazole ring of R_1 towards inhibitory potency for class II FLAP antagonists is also evident from previous SAR studies (Chu et al., 2012).

Classes III, IV, and V are cyclobutylbenzene, cyclopropylethylbenzene, and dimethylpropylbenzene derivatives of oxadiazole, respectively (**Figure 1**). Unlike other three classes, no positive correlation between lipophilicity and inhibitory potency was observed for classes III, IV, and V (**Supplementary Figure S3**). It means that the difference in inhibitory potency might be due to the distinct interaction

pattern and $LipE$ values. All compounds of classes IV, V, and VI contain diverse substituents at R_1 and pyrimidinamine at R_2 (**Supplementary Table S6**), which occupies the inside of the FLAP binding cavity (**Figures 2E–G**). The final cluster of class III (cluster 3, **Supplementary Table S5**) reveals that the common scaffold shows π - π stacking with B-F114 and π -H-bond interactions with B-A63 (**Figure 2E**). R_2 forms hydrogen bonding with B-D62, C-C60, and π -H-bond interactions with C-N57 and C-N23, while R_1 seems to be involved in making hydrogen bonds with B-K116 and π -H interactions with B-I119. Compound **7** ($IC_{50} = 1.0$ nM, $cLogP = 3.13$, $LipE = 5.87$, **Supplementary Table S3**), being the most potent compound of class III, contains N-tert-butyl methylacetamide at the pyrazole ring of R_1 compared to compound **83** ($IC_{50} = 6.5$ nM, $cLogP = 2.68$, $LipE = 5.50$, **Supplementary Table S3**), which contains only piperidine ring at R_1 , resulting in a twofold decrease in its inhibitory potency (**Supplementary Table S6**). The final docking solution of compound **7** reveals that the carbonyl group of N-tert-butyl methylacetamide at R_1 forms a strong hydrogen bond with $-NH_2$ of B-K116 (**Figure 3**). However, for compound **83**, no interaction was observed between the piperidine ring of R_1 and amino acid residues of the FLAP binding cavity. The difference in the binding interaction

pattern of compounds **7** and **83** might be solely responsible for the difference in inhibitory potencies of both compounds as lipophilicities and LipE values are not significantly different. For class IV, ligand–protein interaction profiling of the final cluster (cluster 4, **Supplementary Table S2**) suggests that the common scaffold makes π -H interactions with B-T66, B-A63, C-G24, B-I119, and C-Val21, and π - π stacking with B-F114 (**Figure 2F**). Amino acid residues such as B-D62 and C-N57 present inside the FLAP binding cavity shows hydrogen bonding with R₂ whereas R₁ makes hydrogen bonds with B-K116 and π - π interactions with C-H28 and C-F25. Compounds **13** and **114** of class IV were selected to evaluate binding poses due to the distinct SAR pattern (**Supplementary Table S6**). Compound **13** (IC₅₀ = 1.3 nM, cLogP = 2.36, LipE = 6.52, **Supplementary Table S3**), being the most active from class IV, contains the terminal methyl at the pyrazole ring of R₁, while compound **114** (IC₅₀ = 29.0 nM, cLogP = 4.86, LipE = 2.67, **Supplementary Table S3**) contains N₁,1-dimethylpyrazol-4-amine at R₁, resulting in a twofold decrease in inhibitory potency. The final docking solution of compound **13** reveals that the pyrazole ring of R₁ is involved in making π - π stacks with C-H28, whereas no significant interaction was observed for terminal methyl (**Figure 3**). For compound **114**, the final binding pose suggests that the terminal pyrazole ring at R₁ is unable to show any interactions that might contribute to low inhibitory potency. The significantly low LipE value of compound **114** as compared to compound **13** suggests that gain in activity of compound **13** might be due to its distinct interaction pattern. Similarly, the ligand–protein interaction analysis of the final cluster (cluster 5, **Supplementary Table S5**) of class V points out that the scaffold makes π -H-bond interactions with B-L120, B-I119, B-A63, C-G24, and C-Val21, and π - π stacks with B-F123 and B-F114 (**Figure 2G**). R₂ is involved in making hydrogen bonds with B-D62 and C-N57 while R₁ makes hydrogen bonds with B-K116 and π - π contact with C-H28. Compound **19** (IC₅₀ = 1.6 nM, cLogP = 3.08, LipE = 5.71, **Supplementary Table S3**) contains terminal methyl at the pyrazole ring of R₁, whereas compound **110** (IC₅₀ = 23 nM, cLogP = 2.65, LipE = 4.98, **Supplementary Table S3**) possesses acetonitrile at the pyrazole ring of R₁ (**Supplementary Table S6**) and approximately two orders of magnitude decrease of inhibitory potency was observed for compound **110** as compared to compound **19**. The final binding pose of compound **19** reveals that the pyrazole ring of R₁ is involved in making π - π contact with C-H28 while terminal methyl could not make any interactions (**Figure 3**). The absence of interactions between acetonitrile at R₁ of compound **110** and amino acid residues of the FLAP binding pocket was likely the reason for the two orders of magnitude decrease in inhibitory potency of compound **110** as LipE and lipophilicity values of both compounds do not differ appreciably. Overall, compounds of classes III, IV, and V displayed better LipE values, but the high inhibitory potencies of highly active compounds are due to strong interactions among particular functional groups and amino acids of the FLAP binding cavity.

Class VI FLAP inhibitors are benzimidazole derivatives (**Figure 1**) having diverse substituents at R₁ and pyrimidinamine at R₂ around the benzimidazole scaffold. The compounds of class VI did not exhibit any correlation between

activity and lipophilicity (**Supplementary Figure S3**). The ligand–protein interaction profile of the final cluster indicated that the pyrimidinamine group orients itself towards the inner side of the FLAP binding cavity and is involved in making hydrogen bonds with B-D62 and C-N23, whereas the common scaffold occupies the between chains B and C and forms π -H bonding with B-I119, C-G24, and C-V21. The diverse R₁ is involved in making hydrogen bonds with B-K116 and π - π interactions with C-H28. Compounds of class VI showed a distinct SAR pattern; e.g., in compound **70** (IC₅₀ = 4.2 nM, cLogP = 2.54, LipE = 5.84), the pyridine moiety of R₁ contains methyl triazole at position three and its replacement with acetonitrile in compound **82** resulted in two orders of magnitude decrease in the inhibitory potency of compound **82** (IC₅₀ = 6.09 nM, cLogP = 2.18, LipE = 6.04). The selected binding pose of compound **70** indicated that the triazole ring is making hydrogen bonds with B-K116 whereas no interaction was observed between acetonitrile and amino acid residues of the FLAP binding cavity in the case of compound **82** (**Figure 3**). The compounds of class VI did not show any correlation with LipE, which means that the difference in binding interactions is the main driving factor behind the difference in activity.

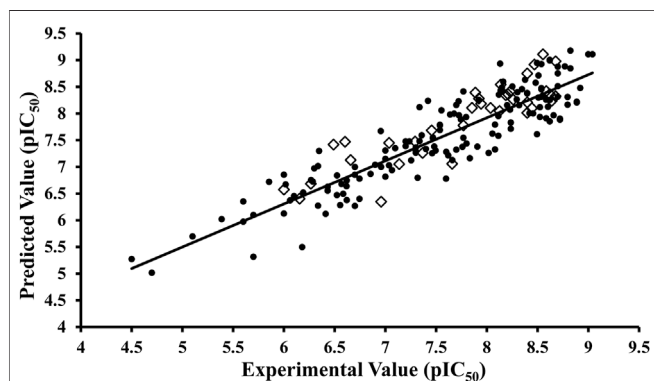
Overall, our criteria for the selection of compounds for molecular modeling studies were cLogP and LipE. However, our results indicate that only cLogP contributes slightly positively towards inhibitory potency for classes I and II, whereas for compounds of classes III, IV, V, and VI, the difference in interaction pattern might be exclusively responsible for the difference in inhibitory potency, as in these classes of FLAP inhibitors, the high LipE values were maintained due to loss in lipophilicity. In addition, our docking results suggest that heterocyclic moieties are involved in making π -H interactions with hydrophobic amino acid residues of the FLAP binding cavity. Therefore, the presence of pyridine, pyrimidine, pyridazine, pyrazole, triazole, and oxadiazole rings moderately increases not only the lipophilicity but also the inhibitory potency. Moreover, an increase or decrease in LipE values of FLAP inhibitors does not alter the inhibitory potencies in either way. Further docking poses obtained from multiple clusters with maximum docked ligands were employed to generate the vGRIND model.

3.4 GRID-Independent Molecular Descriptors Analysis

The selected binding poses of 151 (**Supplementary Table S3**) compounds of the training set obtained through common scaffold clustering of docking poses along with their inhibitory activity (pIC₅₀) values were implied in the pentacle v 1.07 software package that utilizes special alignment independent GRIND descriptors to develop a 3D-QSAR model. To correlate the inhibitory potencies with 3D structural features and to derive the most important pharmacophoric features of our training set, a partial least square model was developed on five principal components using the leave-one-out (LOO) cross-validation method, resulting in initial models with satisfactory values of variables. The inconsistent nodes were removed by one-

TABLE 2 | Statistical parameters obtained before and after application of fractional factorial design (FFD) on final GRIND model.

Fractional factorial design cycle (FFD)											
Datasets	Complete variable					Datasets	FFD1				
	R^2	q^2_{Loo}	SDEP	r^2_m	Delta r^2_m		R^2	q^2_{Loo}	SDEP	r^2_m	Delta r^2_m
Training set	0.71	0.60	0.49	0.703	0.004	Training set	0.82	0.66	0.47	0.775	0.001
Test set	0.63	0.58	0.49	0.517	0.028	Test set	0.77	0.64	0.47	0.686	0.012

**FIGURE 4** | Activity interactive graph plot between predicted and actual experimental activity values. The graph plot displays separate data series for training (filled circles) and test (rhombus) set. R^2 for training set was observed as 0.82 and 0.77 for test set.

time application of the fractional factorial design (FFD) variable selection algorithm. The final GRIND model was obtained with good values of performance measures, $q^2 = 0.66$ and $R^2 = 0.82$, while the standard error of prediction (SDEP) was 0.47. The before and after FFD application statistics along with the r^2_m metric is shown in **Table 2**. The difference between actual and predicted activity values was less than one log unit for all 151 inhibitors of the training set as shown in **Figure 4**. The test set (**Supplementary Table S4**) was used for the evaluation of the final GRIND model, which predicted inhibitory potencies of test set compounds with a difference of less than one log unit for all compounds between experimental and predicted pIC_{50} values with R^2 value observed as 0.77 (**Figure 4**).

A PLS coefficient correlogram of the GRIND variables is shown in **Figure 5A** and describes important 3D structural features that directly/inversely correlate with the inhibitory potencies of the training set compounds. The PLS coefficient correlogram depicts that DRY-DRY, DRY-N1, DRY-TIP, and N1-TIP pair of probes positively contribute towards the inhibitory potency of chemically diverse FLAP inhibitors whereas no inverse contribution was observed by any variable. These variables are located at a certain distance within active inhibitors between substitutions at R_1 , R_2 , and common scaffolds.

More explicitly, the DRY-DRY correlogram in **Figure 5A** shows the presence of two hydrophobic contours (HYD1 and HYD2) at a mutual distance of 16.00–16.40 Å in a virtual receptor

site of highly active FLAP inhibitors $pIC_{50} > 7.5$. For class I, the distance is present between the pyridine ring of R_1 and the phenyl ring of R_2 ; for class II, it is observed between the quinoline group and the endo aryl moiety of the common scaffold; for classes III, IV, and V, it is present between the pyrazole ring at R_1 position and the phenyl ring of the common scaffold; and for class VI, it was observed between phenyl of the common scaffold and pyrimidinamine of R_2 (**Figures 5B–E**). Furthermore, the backstage projection of the actual FLAP structure onto the identified hotspots revealed the presence of complementary hydrophobic amino acid residues such as B-A63, B-I119, and B-L120. This further strengthened our docking outcomes as all of these amino acid residues are involved in making extensive π -H interactions with dataset compounds. Additionally, a recent pharmacophore study of Olgac et al. revealed that four hydrophobic features are important in most potent indole- and oxadiazole-based FLAP inhibitors (Olgac et al., 2020).

Similarly, DRY-N1 (**Figure 5A**) explicates the positive contribution of one hydrophobic (HYD3) and one hydrogen bond acceptor (HBA1) at a mutual distance of 16.40–16.80 Å within active FLAP modulators. Interestingly, this distance was observed in all highly active FLAP modulators $pIC_{50} > 7.5$ and absent in all less-active compounds $pIC_{50} < 7.5$. Briefly, for class I, it is observed between the terminal negative ionizable moiety present at R_2 and the pyridine ring of R_1 ; in class II, it is observed between the quinoline group and pyrazole ring; for classes III, IV, and V, it is present between the pyrimidinamine group of R_2 and oxadiazole ring; and for class VI, it was observed between pyrimidinamine and pyridine of R_1 as displayed in **Figures 5B–E**. Projecting actual FLAP structure onto the identified virtual hotspots revealed the presence of hydrophobic amino acids B-F114, B-A63, and C-G24 and complementary amide groups in the B-K116 and B-R117 amino acid residues within the FLAP binding cavity that further complements the accuracy of our model. These results further reinforce our docking outcomes, which demonstrated the importance of B-A63 and B-K116 for the hydrophobic and hydrogen bonding interactions within the FLAP binding cavity. These outcomes are also in accord with another pharmacophore-based study that demonstrated the importance of hydrophobic and hydrogen bond acceptor features in the highly active indole- and biaryl bicycloheptane-based FLAP inhibitors (Temml et al., 2017).

Moreover, DRY-TIP correlogram (**Figure 5A**) portrays the presence of one hydrophobic (HYD4) and one shape-based feature (TIP1) that positively contribute towards the inhibitory potency of FLAP inhibitors. For the sharpest peak, the two

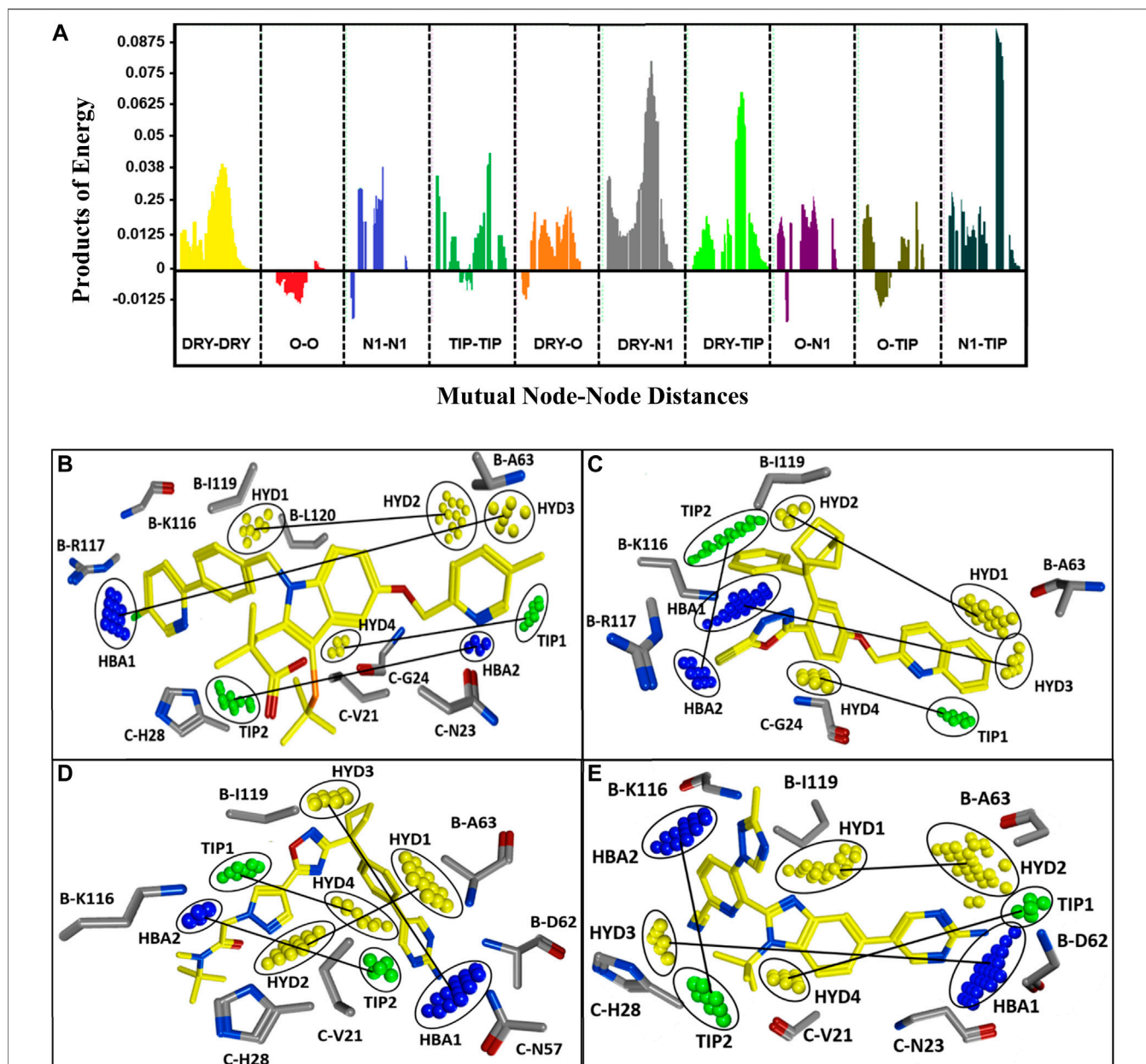


FIGURE 5 | (A) Correlogram of PLS coefficients representing the pair of probes contributing positively (peaks above 0) or negatively (peaks below 0) towards the inhibitory potencies of training set compounds. The positive contribution towards pIC_{50} of FLAP inhibitors has been depicted by DRY-DRY (two hydrophobic), DRY-N1 (one hydrophobic and one hydrogen bond acceptor), DRY-TIP (one hydrophobic and one steric), and N1-TIP (one hydrogen bond acceptor and one steric) variables. The variables are present in all highly active FLAP compounds and are located at mutual distances of 16.00–16.40 Å, 16.40–16.80 Å, 18.00–18.40 Å, and 17.20–17.60 Å, respectively. **(B)** The identified hotspots on most active indole-based FLAP inhibitor (compound 1) of training set with projection of actual FLAP structure. Hydrophobic features are depicted in yellow, hydrogen bond acceptors are in blue, while steric hotspots are depicted in green color. The two hydrophobic hotspots (HYD1 and HYD2) are located between two aromatic moieties, one hydrophobic (HYD3) and one hydrogen bond acceptor feature (HBA1) are present between aromatic rings and terminal negative ionizable substitution, one hydrophobic (HYD4) and steric feature (TIP1) can be spotted between aromatic ring and indole scaffold, while one hydrogen bond acceptor (HBA2) and one steric (TIP2) hotspot are present between dimethylbutanoic acid and pyridine ring. **(C)** The most active compound (compound 10) of class II with mapping of complemented amino acids on the recognized contours. **(D)** The most active of class III (compound 7), which is also the most active compound from oxadiazole-based FLAP antagonists (classes III, IV, and V) and mapped hotspots along with projection of complementary amino acids of FLAP binding cavity. Due to high structural similarity, the features were also observed at the same positions in all active compounds of classes IV and V. **(E)** The compound (70) from class VI with identified hotspots and corresponding amino acids.

contours are present at a mutual distance of 18.00–18.40 Å, indicating the distance between indole scaffold and R₁ for class I; exoaryl of common scaffold and quinoline for class II; phenyl ring of scaffold and pyrazole ring of R₁ for classes III, IV, and V; and pyrazole of common scaffold and pyrimidinamine for class VI (Figures 5B–E). Two identified contours (HYD4 and TIP) were mapped on the actual FLAP binding site, and interestingly, the hydrophobic region in all active compounds is complementary to hydrophobic amino acids C-V21, C-G24, and C-H28. It is also in accordance with our docking findings where many compounds of our dataset are involved in making π -H interactions and π - π stacking with these amino acids. The green contour elucidates a steric hotspot region, and it defines the 3D molecular shape of FLAP inhibitors.

The last selected peak N1-TIP (Figure 5A) represents the presence of one hydrogen bond acceptor (HBA2) and one shape-based feature (TIP2) at a mutual distance of 17.20–17.60 Å within highly active FLAP inhibitors. The two features at this distance contributes positively towards the inhibitory potency of compounds against FLAP. The hydrogen bond acceptor hotspot in the virtual receptor site of FLAP is complemented by the presence of B-K116, B-D62, C-N57, and C-N23 amino acids in the actual receptor site when we mapped the FLAP structure onto the identified N1 (HBA2) hotspot. The -NH₂ and carbonyl groups of these amino acids are involved in making hydrogen bonds with active FLAP modulators as evident from our docking studies and pose analysis. Moreover, these features have been observed in all active FLAP inhibitors pIC₅₀ > 7.5 while they are absent in less active compounds pIC₅₀ < 7.5. For class I, the distance is present between the pyrimidine ring at R₂ position and dimethylbutanoic acid; for class II it is observed between substituents at the pyrazole ring of R₁ and endo aryl of scaffold; for classes III, IV, and V, it is present between the pyrimidine amine of R₂ and pyrazole ring at R₁ position; and for class VI, it is present between the triazole of R₁ and tertbutyl of the common scaffold (Figures 5B–E). The TIP probe signifies the importance of a steric hotspot at a distance of 17.20–17.60 Å from the hydrogen bond acceptor feature.

Generally, our study provided a deeper understanding of three-dimensional requirements of diverse inhibitor binding within the FLAP binding cavity by mapping the mutual distances of important pharmacophoric features (four hydrophobic, two hydrogen bond acceptor, and two steric features) as well as the complementary distances of the important interacting amino acid residues (B-L120, B-I119, B-R117, B-K116, B-F114, B-A63, B-D62, C-H28, C-G24, C-N23, and C-V21). Previous docking studies also revealed that highly potent FLAP modulators result in π - π stacking with C-H28, hydrophobic interactions with B-L120, B-I119, and hydrogen bonding with B-R117, B-K116, and B-D62 (Ma et al., 2008). Overall, the binding hypothesis generated for each class within the FLAP binding cavity was complementary with our GRIND model, which predicted the inhibitory potencies of validation and test sets with reasonable accuracy, indicating the fitness of our model. Based on our current findings, we suggest that the high inhibitory potency of a compound against FLAP can be achieved by (1) increasing the hydrogen bond acceptor

strength on at least one substitution position (R₁ or R₂); (2) insertion of heterocyclic moieties such as pyridine, pyrimidine, pyridazine, pyrazole, and triazole at each side of the common scaffold to increase hydrophobic strength; and (3) maintaining a distance of 16.00–16.40 Å between two hydrophobic groups (aromatic rings) and 16.40–16.80 Å between hydrophobic and hydrogen bond acceptor groups.

4 DISCUSSION

Since high levels of leukotrienes have been reported in multiple pathophysiological conditions in the past 3 decades, leukotriene synthesis pathway has been targeted at many levels while FLAP has received the greatest focus because it initiates the biological synthesis of leukotrienes *via* leukotriene synthesis pathway (Massoumi and Sjölander 2007; Bryda and Wątroba 2018; Jo-Watanabe, Okuno, and Yokomizo 2019). Several practices have been made to propose potent FLAP modulators, and many of them have shown good clinical efficacy. However, not a single molecule could be able to change into the status of “drug”. The focus of the present study is to unveil the two- and three-dimensional structural requirements of FLAP modulators.

First to demonstrate the important two-dimensional structural features, supervised ML approach was adapted over classical 2D QSAR modeling. The preference was made for two reasons: (1) to escape the alignment step as FLAP modulators are highly diverse in nature, and (2) evidence from the past strengthens the adaptation of ML for quantitative structure–activity relationship studies (Tsou et al., 2020; Gupta et al., 2021). We developed multiple ML models including XGboost (GBDT), random forest (RF), decision tree (DT), support vector machine (SVM), logistic regression (LR), and multilayer perceptron (MLP), and in comparison, XGBoost and RF were able to classify our training set and predict the test set with significant classification and prediction accuracies. Moreover, recursive feature elimination with cross-validation (RFECV) captured relevant features or 2D descriptors, which are mainly participating in the classification of highly active and least active FLAP inhibitors.

Further molecular modeling studies were performed to vaticinate the important three-dimensional pharmacophoric features instead of ML. The preference was made because (1) three-dimensional structural properties are highly dependent on binding poses, and (2) the GRIND model not only explains the important molecular interaction fields but also distances between them along with important amino acid residues by creating a virtual receptor site (Shafi and Jabeen 2017). Before implication of molecular modeling strategies, the dataset was first subjected to calculation of LipE and cLogP. The purpose of LipE-based lead optimization is to improve LipE while maintaining an appropriate range of logP for the optimization of potency and ADME properties. The increased potency of a compound with eque-LipE to the reference ligand demonstrates that the increase in lipophilicity alone is responsible for the increased potency, although other factors associated with the specific structural change cannot be ruled out. Also, an increase in LipE of a

compound suggests that an increase in potency is beyond lipophilicity increases alone and other factors such as transport to the target and hydrogen bonding strength within the protein binding site could be associated with this response. In total, 238 out of 658 demonstrated the LipE value greater than five, which is the optimal threshold, and only 136 demonstrated cLogP between the optimal range of 2–3.5. Herein, we selected 187 compounds having LipE greater than one and cLogP greater than two because FLAP is an integral membrane protein, which means that compounds should possess a high lipophilicity value for efficient binding. Also, FLAP modulators in clinical trials usually possess lipophilicity >3. The dataset of 187 compounds was divided into six distinct classes based on the common scaffold with subsequent docking into the FLAP binding pocket. Our docking results indicated that the FLAP binding pocket can cater diverse anti-inflammatory compounds and they bind between chains B and C. The ligand–protein interaction profile of selected FLAP modulators revealed that mostly B-R117, BK-116, C-N57, C-N23, and B-D62 FLAP amino acid residues are involved in making hydrogen bonds; B-A63, B-L119, B-L120, B-V21, and C-G24 make π -H-bond interactions, whereas C-H28 is involved in forming π - π contact with FLAP modulators. Also, for classes I and II, a moderate correlation was observed between lipophilicity and inhibitory potency, whereas for compounds of classes III, IV, V, and VI, an increase or decrease in lipophilicity or LipE did not alter the inhibitory potency in either way or *vice versa*.

To select the most probable binding poses, common scaffold clustering was performed because using GRID-independent molecular descriptors, analysis of 3D structural features is highly dependent on 3D confirmations of the molecules (Pastor et al., 2000). Multiple clusters at 3.5 Å RMSD were generated and binding poses from clusters with maximum number of docked ligands were further used to build the GRIND model. The reliability of binding pose selection *via* common scaffold clustering for generation of the GRIND model can be explained by satisfactory statistical results obtained for the final GRIND model. Furthermore, the model signifies the positive contribution of four hydrophobic, two hydrogen bond acceptor, and two steric features towards the inhibitory potency of FLAP modulators. The identified hotspots or pharmacophoric features were successfully mapped onto the highly active FLAP modulators followed by projection of the actual receptor site, which revealed the presence of corresponding amino acid residues. Overall, our GRIND model suggested that (1) two hydrophobic features should be present at a mutual distance of 16.00–16.40 Å, (2) one hydrophobic and hydrogen bond acceptor feature should be present at a distance of 16.40–16.80 Å, (3) the distance between hydrophobic and steric feature should be 18.00–18.40 Å, and (4) it should be 17.20–17.60 Å between hydrogen bond acceptor and steric features. The importance of hydrophobic and hydrogen bond acceptor features has also been demonstrated by previous studies (Temml et al., 2017; Olgac et al., 2020).

Based on these findings, further analyses will focus on virtual screening from both ML and GRIND models followed by selection of common compounds. The common hits can further be structurally tuned and optimized. The ML and GRIND model will allow internal inspection of FLAP modulators, before validating them using predictions on vendor libraries, purchase, and testing.

5 CONCLUSION

The current study deals with the development of ML models and a GRIND model on a diverse series of FLAP inhibitors. First of all, our ML models signify some important 2D descriptors, and the best-performing model (XG-Boost) has successfully classified the active and inactive compounds present in our training set exhibiting 91% overall classification accuracy. The subsequent screening of test set from the model resulted in 90% prediction accuracy, which further accentuates the efficiency of the model. Secondly, docking studies reveal that hydrogen bonding and hydrophobic interactions are critical for binding of FLAP inhibitors. Further common scaffold-based clustering revealed the optimal binding mode of structurally diverse inhibitors and aided in determination of their molecular basis of interaction within the FLAP binding cavity. Thirdly, the most probable binding poses were utilized for GRIND model development, which showed valid statistical results having an R^2 of 0.82 and a q^2 of 0.66. Additionally, the GRIND model predicted all compounds of training and test set with an activity difference of less than one log unit. Overall, our GRIND model illustrated that four hydrophobic, two hydrogen bond acceptor, and two steric features are critical for achieving high inhibitory potency against FLAP. All the features were successfully complemented by the docking studies highlighting the significance of respective amino acid residues such as B-L120, B119, B-A63, C-H28, C-G24, and C-V21 for hydrophobic interactions and B-R117, B-K116, D-62, and C-N57 for hydrogen bonding. In general, application of ML, docking analysis, common scaffold clustering, and GRIND modeling to predict the 2D structural requirements as well as the 3D molecular basis of interaction of diverse FLAP inhibitors could potentially guide the development of more potent chemotypes for the treatment of inflammatory disorders requiring anti-leukotriene therapy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization, HK and IJ; methodology, HK and IJ; software, IJ; validation, HK and IJ; formal analysis, HK; investigation, HK; resources, IJ; data curation, HK; writing—original draft preparation, HK; writing—review and editing, IJ; visualization, HK; supervision, IJ; project administration, IJ.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2022.825741/full#supplementary-material>

REFERENCES

- Antoniou, S. A. (2014). Targeting 5-Lipoxygenase-Activating Protein in Asthma and Chronic Obstructive Pulmonary Disease. *Expert Opin. Ther. Targets* 18 (11), 1285–1292. doi:10.1517/14728222.2014.945425
- Arnott, J. A., and Planey, S. L. (2012). The Influence of Lipophilicity in Drug Discovery and Design. *Expert Opin. Drug Discov.* 7, 863–875. doi:10.1517/17460441.2012.714363
- Bain, G., King, C. D., Rewolinski, M., Schaab, K., Santini, A. M., Shapiro, D., et al. (2010). Pharmacodynamics and Pharmacokinetics of AM103, a Novel Inhibitor of 5-Lipoxygenase-Activating Protein (FLAP). *Clin. Pharmacol. Ther.* 87 (4), 437–444. doi:10.1038/clpt.2009.301
- Banoglu, E., Çalişkan, B., Luderer, S., Eren, G., Özkan, Y., Altenhofen, W., et al. (2012). Identification of Novel Benzimidazole Derivatives as Inhibitors of Leukotriene Biosynthesis by Virtual Screening Targeting 5-Lipoxygenase-Activating Protein (FLAP). *Bioorg. Med. Chem.* 20 (12), 3728–3741. doi:10.1016/j.bmc.2012.04.048
- BioByte - Bio-Loom (2021). BioByte - Bio-Loom. Available at: <http://biobyte.com/bb/prod/bioloom.html> (Accessed November 10, 2020).
- Breiman, L. (2001). Random Forests. *Machine Learn.* 45 (1), 5–32. doi:10.1023/A:1010933404324
- Brieman, F., and Olshen, S. (2012). Classification A Nd R Egression Trees. *Chapman & Hall/Crc* 66, 37–39.
- Bryda, J., and Wątroba, S. (2018). The Proinflammatory Role of Lipoxygenases in Rheumatoid Arthritis. *J. Pre Clin. Clin. Res.* 12 (4), 129–134. doi:10.26444/JPCRR/99597
- Chang, C.-C., and Lin, C.-J. 2021 “LIBSVM: A Library for Support Vector Machines.”
- Chen, A. A., and Pappu, R. V. (2007). Parameters of Monovalent Ions in the AMBER-99 Forcefield: Assessment of Inaccuracies and Proposed Improvements. *J. Phys. Chem. B* 111 (41), 11884–11887. doi:10.1021/jp0765392
- Chen, T., and Guestrin, C. (2016). “XGBoost.” Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 13-17: 785–794. doi:10.1145/2939672.2939785
- Chu, L., Armstrong, H. M., Chang, L. L., Cheng, A. F., Colwell, L., Cui, J., et al. (2012). Evaluation of Endo- and Exo-Aryl-Substitutions and Central Scaffold Modifications on Diphenyl Substituted Alkanes as 5-Lipoxygenase Activating Protein Inhibitors. *Bioorg. Med. Chem. Lett.* 22 (12), 4133–4138. doi:10.1016/j.bmcl.2012.04.064
- Cortes, C., Vapnik, V., and Saitta, L. (1995). Support-Vector Networks. *Mach Learn.* 20 (3), 273–297. doi:10.1007/BF00994018
- Dahlén, B., Kumlin, M., Ihre, E., Zetterström, O., and Dahlen, S. E. (1997). Inhibition of Allergen-Induced Airway Obstruction and Leukotriene Generation in Atopic Asthmatic Subjects by the Leukotriene Biosynthesis Inhibitor BAYx 1005. *Thorax* 52 (4), 342–347. doi:10.1136/thx.52.4.342
- Diamant, Z., Timmers, M. C., van der Veen, H., Friedman, B. S., De Smet, M., Depré, M., et al. (1995). The Effect of MK-0591, a Novel 5-Lipoxygenase Activating Protein Inhibitor, on Leukotriene Biosynthesis and Allergen-Induced Airway Responses in Asthmatic Subjects *In Vivo*. *J. Allergy Clin. Immunol.* 95 (1–2), 42–51. doi:10.1016/S0091-6749(95)70151-6
- Durán, A., Martínez, G. C., and Pastor, M. (2008). Development and Validation of AMANDA, a New Algorithm for Selecting Highly Relevant Regions in Molecular Interaction Fields. *J. Chem. Inf. Model.* 48 (9), 1813–1823. doi:10.1021/ci800037t
- Ericsson, H., Nelander, K., Heijer, M., Kjaer, M., Lindstedt, E. L., Albayaty, M., et al. (2020). Phase 1 Pharmacokinetic Study of AZD5718 in Healthy Volunteers: Effects of Coadministration with Rosuvastatin, Formulation and Food on Oral Bioavailability. *Clin. Pharmacol. Drug Dev.* 9 (3), 411–421. doi:10.1002/cpdd.756
- Evans, J. F., Lévellé, C., Mancini, J. A., Prasit, P., Thérien, M., Zamboni, R., et al. (1991). 5-Lipoxygenase-Activating Protein Is the Target of a Quinoline Class of Leukotriene Synthesis Inhibitors. *Mol. Pharmacol.* 40 (1), 22–27.
- Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., and Lin, C.-J. (2008). LIBLINEAR: A Library for Large Linear Classification. *J. Machine Learn. Res.* 9, 1871–1874.
- Ferguson, A. D., McKeever, B. M., Xu, S., Wisniewski, D., Miller, D. K. Ting. Ting. Yamin., Yamin, T. T., et al. (2007). Crystal Structure of Inhibitor-Bound Human 5-Lipoxygenase-Activating Protein. *Science* 317 (5837), 510–512. doi:10.1126/science.1144346
- Folco, G., Rossoni, G., Buccellati, C., Berti, F., Maclouf, J., and Sala, A. (2000). Leukotrienes in Cardiovascular Diseases. *Am. J. Respir. Crit. Care Med.* 161 (2 II), S112–S116. doi:10.1164/ajrccm.161.supplement_1.lta-22
- Frenette, R., Hutchinson, J. H., Léger, S., Thérien, M., Brideau, C., Chan, C. C., et al. (1999). Substituted Indoles as Potent and Orally Active 5-Lipoxygenase Activating Protein (FLAP) Inhibitors. *Bioorg. Med. Chem. Lett.* 9 (16), 2391–2396. doi:10.1016/S0960-894X(99)00399-6
- Friedman, B. S., Bel, E. H., Buntinx, A., Tanaka, W., Han, Y. H., Shingo, S., et al. (1993). Oral Leukotriene Inhibitor (MK-886) Blocks Allergen-Induced Airway Responses. *Am. Rev. Respir. Dis.* 147 (4), 839–844. doi:10.1164/ajrccm/147.4.839
- Gajo, G. C., de Assis, T. M., Assis, L. C., Leticia, C. A., Ramalho, T. C., da Cunha, E. F. F., et al. (2016). Quantitative Structure-Activity Relationship Studies for Potential Rho-Associated Protein Kinase Inhibitors. *J. Chem.* 2016, 1–12. doi:10.1155/2016/9198582
- Ghosh, A., Chen, F., Thakur, A., and Hong, H. (2016). Cysteinyl Leukotrienes and Their Receptors: Emerging Therapeutic Targets in Central Nervous System Disorders. *CNS Neurosci. Ther.* 22 (12), 943–951. doi:10.1111/cns.12596
- Glorot, X., and Yoshua, B. (2021) Understanding the Difficulty of Training Deep Feedforward Neural Networks.
- Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., and Kumar, P. (2021). Artificial Intelligence to Deep Learning: Machine Intelligence Approach for Drug Discovery. *Mol. Divers.* 25 (April), 1315–1360. doi:10.1007/S11030-021-10217-3
- Gür, Z. T., Çalişkan, B., Banoglu, E., and Banoglu, Erden. (2018). Drug Discovery Approaches Targeting 5-Lipoxygenase-Activating Protein (FLAP) for Inhibition of Cellular Leukotriene Biosynthesis. *Eur. J. Med. Chem.* 153, 34–48. doi:10.1016/j.ejmech.2017.07.019
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene Selection for Cancer Classification Using Support Vector Machines. *Machine Learn.* 46 (1 461), 389–422. doi:10.1023/A:1012487302797
- Haeggström, J. Z. (2000). “Structure, Function, and Regulation of Leukotriene A4 Hydroxylase” *American Journal of Respiratory and Critical Care Medicine*. American Lung Association. doi:10.1164/ajrccm.161.supplement_1.lta-6
- Haykin, S. (2009). *Neural Networks and Learning Machines*. Third Edition. New York, Boston San, Francisco London, Toronto Sydney, Tokyo Singapore, Madrid Mexico.
- Hedi, H., and Norbert, G. (2004). 5-Lipoxygenase Pathway, Dendritic Cells, and Adaptive Immunity. *J. Biomed. Biotechnol.* 2004, 99–105. doi:10.1155/S1110724304310041
- Jabeen, I., Pleban, K., Rinner, U., Chiba, P., and Ecker, G. F. (2012). Structure-Activity Relationships, Ligand Efficiency, and Lipophilic Efficiency Profiles of Benzophenone-type Inhibitors of the Multidrug Transporter P-Glycoprotein. *J. Med. Chem.* 55 (7), 3261–3273. doi:10.1021/jm201705f
- Jakschik, B. A., and Kuo, C. G. (1983). Characterization of Leukotriene A4 and B4 Biosynthesis. *Prostaglandins* 25 (6), 767–782. doi:10.1016/0090-6980(83)90002-3
- Jedlitschky, G., and Keppler, D. (2002). “Transport of Leukotriene C4 and Structurally Related Conjugates” *Vitamins and Hormones*. Academic Press. doi:10.1016/s0083-6729(02)64005-1
- Jo-Watanabe, A., Okuno, T., and Yokomizo, T. (2019). The Role of Leukotrienes as Potential Therapeutic Targets in Allergic Disorders. *Int. J. Mol. Sci.* 20 (14). doi:10.3390/IJMS20143580
- Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* 267 (3), 727–748. doi:10.1006/jmbi.1996.0897
- Leeson, P. D., and Springthorpe, B. (2007). The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discov.* 6 (11), 881–890. doi:10.1038/nrd2445
- Lemurell, M., Ulander, J., Winiwarter, S., Dahlén, A., Davidsson, Ö., Emtenäs, H., et al. (2015). Discovery of AZD6642, an Inhibitor of 5-Lipoxygenase Activating Protein (FLAP) for the Treatment of Inflammatory Diseases. *J. Med. Chem.* 58 (2), 897–911. doi:10.1021/jm501531v
- Liaw, A., and Wiener, M. (2002). Classification and Regression by. *RandomForest* 2 (3).

- Liu, M., and Yokomizo, T. (2015). The Role of Leukotrienes in Allergic Diseases. *Allergol. Int.* 64, 17–26. doi:10.1016/j.alit.2014.09.001
- Lorrain, D. S., Bain, G., Charles, C., Correa, L. D., Santini, A. M., Chapman, C., et al. (2010). Pharmacology of AM803, a Novel Selective Five-Lipoxygenase-Activating Protein (FLAP) Inhibitor in Rodent Models of Acute Inflammation. *Eur. J. Pharmacol.* 640 (1–3), 211–218. doi:10.1016/j.ejphar.2010.05.003
- Ma, X., Zhou, L., Zuo, Z.-L., Liu, J., Yang, M., and Wang, R.-W. (2008). Molecular Docking and 3-D QSAR Studies of Substituted 2,2-Bisaryl-Bicycloheptanes as Human 5-Lipoxygenase-Activating Protein (FLAP) Inhibitors. *QSAR Comb. Sci.* 27 (9), 1083–1091. doi:10.1002/qsar.200810053
- Macdonald, D., Brideau, C., Chan, C. C., Falguyret, J. P., Falguyret, J. P., Frenette, R., et al. (2008). Substituted 2,2-Bisaryl-Bicycloheptanes as Novel and Potent Inhibitors of 5-Lipoxygenase Activating Protein. *Bioorg. Med. Chem. Lett.* 18 (6), 2023–2027. doi:10.1016/j.bmcl.2008.01.105
- Mancini, J. A., Coppolino, M. G., Klassen, J. H., Klassen, J. H., Charleson, S., and Vickers, P. J. (1994). The Binding of Leukotriene Biosynthesis Inhibitors to Site-Directed Mutants of Human 5-Lipoxygenase-Activating Protein. *Life Sci.* 54 (9), PL137–42. doi:10.1016/0024-3205(94)00872-8
- Massoumi, R., and Sjölander, A. (2007). The Role of Leukotriene Receptor Signaling in Inflammation and Cancer. *ScientificWorldJournal* 7 (September), 1413–1421. doi:10.1100/TSW.2007.200
- Mauri, A. (2020). AlvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints. *Methods Pharmacol. Toxicol.*, 801–820. doi:10.1007/978-1-0716-0150-1_32
- McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models*, 511.
- Nekoei, M., Mohammadhosseini, M., and Pourbasheer, E. (20152015). QSAR Study of VEGFR-2 Inhibitors by Using Genetic Algorithm-Multiple Linear Regressions (GA-MLR) and Genetic Algorithm-Support Vector Machine (GA-SVM): A Comparative Approach. *Med. Chem. Res.* 24 (7), 3037–3046. doi:10.1007/S00044-015-1354-4
- Olgac, A., Carotti, A., Kretzer, C., Zergiebel, S., Seeling, A., Garscha, U., et al. (2020). Discovery of Novel 5-Lipoxygenase-Activating Protein (FLAP) Inhibitors by Exploiting a Multistep Virtual Screening Protocol. *J. Chem. Inf. Model.* 60 (3), 1737–1748. doi:10.1021/acs.jcim.9b00941
- Pastor, M., Cruciani, G., McLay, I., Pickett, S., and Clementi, S. (2000). GRIND-Independent Descriptors (GRIND): A Novel Class of Alignment-independent Three-Dimensional Molecular Descriptors. *J. Med. Chem.* 43 (17), 3233–3243. doi:10.1021/jm000941m
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Bertrand, T., Grisel, O., et al. (2012). Scikit-Learn: Machine Learning in Python. *J. Machine Learn. Res.* 12 (January), 2825–2830.
- Peters-Golden, M. (1998). Cell Biology of the 5-Lipoxygenase Pathway. *Am. J. Respir. Crit. Care Med.* 157 (6 Pt 1), S227–S228. doi:10.1164/ajrccm.157.6.mar4
- Peters-Golden, M., and Brock, T. G. (2003). 5-Lipoxygenase and FLAP. *Prostaglandins, Leukot. Essent. Fatty Acids* 69 (2–3), 99–109. doi:10.1016/S0952-3278(03)00070-X
- Pourbasheer, E., Vahdani, S., Malekzadeh, D., Aalizadeh, R., and Ebadi, A. (2017). QSAR Study of 17 β -HSD3 Inhibitors by Genetic Algorithm-Support Vector Machine as a Target Receptor for the Treatment of Prostate Cancer. *Iran J. Pharm. Res.* 16 (3), 966–980.
- Quinlan, J. R. (1986). Induction of Decision Trees. *Mach Learn.* 1 (1), 81–106. doi:10.1007/BF00116251
- Roy, K., Chakraborty, P., Mitra, I., Ojha, P. K., Kar, S., and Das, R. N. (2013). Some Case Studies on Application of "r(m)²" Metrics for Judging Quality of Quantitative Structure-Activity Relationship Predictions: Emphasis on Scaling of Response Data. *J. Comput. Chem.* 34 (12), 1071–1082. doi:10.1002/JCC.23231
- Shafi, T., and Jabeen, I. (2017). Grid-Independent Descriptors (GRIND) Analysis and SAR Guided Molecular Docking Studies to Probe Selectivity Profiles of Inhibitors of Multidrug Resistance Transporters ABCB1 and ABCG2. *Curr. Cancer Drug Targets* 17 (2), 177–190. doi:10.2174/1568009616666160901094140
- Smyrniotis, C. J., Barbour, S. R., Xia, Z., Hixon, M. S., and Holman, T. R. (2014). ATP Allosterically Activates the Human 5-Lipoxygenase Molecular Mechanism of Arachidonic Acid and 5(S)-Hydroperoxy-6(E),8(Z),11(Z),14(Z)-Eicosatetraenoic Acid. *Biochemistry* 53 (27), 4407–4419. doi:10.1021/bi401621d
- Stock, N., Baccei, C., Bain, G., Chapman, C., Correa, L., Darlington, J., et al. (2010). 5-Lipoxygenase-activating Protein Inhibitors. Part 3: 3-[3-Tert-Butylsulfanyl-1-[4-(5-Methoxy-Pyrimidin-2-Yl)-Benzyl]-5-(5-Methylpyridin-2-Ylmethoxy)-1h-Indol-2-Yl]-2,2-Dimethyl-Propionic Acid (AM643)-A Potent FLAP Inhibitor Suitable for Topical Administration. *Bioorg. Med. Chem. Lett.* 20 (15), 4598–4601. doi:10.1016/j.bmcl.2010.06.011
- Stock, N. S., Bain, G., Zunic, J., Li, Y., Ziff, J., Roppe, J., et al. (2011). 5-Lipoxygenase-Activating Protein (FLAP) Inhibitors. Part 4: Development of 3-[3-Tert-Butylsulfanyl-1-[4-(6-Ethoxy-pyridin-3-Yl)Benzyl]-5-(5-Methylpyridin-2-Ylmethoxy)-1 H -Indol-2-Yl]-2,2-Dimethylpropionic Acid (AM803), a Potent, Oral, once Daily FLAP in. *J. Med. Chem.* 54 (23), 8013–8029. doi:10.1021/jm2008369
- Temml, V., Garscha, U., Romp, E., Schubert, G., Gerstmeier, J., Kutil, Z., et al. (2017). Discovery of the First Dual Inhibitor of the 5-Lipoxygenase-Activating Protein and Soluble Epoxide Hydrolase Using Pharmacophore-Based Virtual Screening. *Sci. Rep.* 7 (January), 42751–42813. doi:10.1038/srep42751
- Tsou, L. K., Yeh, S.-H., Ueng, S.-H., Chang, C.-P., Song, J.-S., Wu, M.-H., et al. (2020). Comparative Study between Deep Learning and QSAR Classifications for TNBC Inhibitors and Novel GPCR Agonist Discovery. *Sci. Rep.* 10, 1–11. doi:10.1038/s41598-020-73681-1
- Tu, G., Qin, Z., Huo, D., Zhang, S., and Yan, A. (2020). Fingerprint-Based Computational Models of 5-Lipoxygenase Activating Protein Inhibitors: Activity Prediction and Structure Clustering. *Chem. Biol. Drug Des.* 96 (3), 931–947. doi:10.1111/cbdd.13657
- Woods, J. W., Coffey, M. J., Brock, T. G., Singer, I. I., Singer, I. I., and Peters-Golden, M. (1995). 5-Lipoxygenase Is Located in the Euchromatin of the Nucleus in Resting Human Alveolar Macrophages and Translocates to the Nuclear Envelope upon Cell Activation. *J. Clin. Invest.* 95 (5), 2035–2046. doi:10.1172/JCI117889
- Zeng, X., and Martinez, T. R. (2010). Distribution-Balanced Stratified Cross-Validation for Accuracy Estimation 12, 1–12. doi:10.1080/095281300146272

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Khan and Jabeen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.