Regular Article

PHYSICAL CHEMISTRY RESEARCH
Published by the
*Iranian Chemical Society*
www.physchemres.org
info@physchemres.org

*Phys. Chem. Res., Vol. 11, No. 4, 783-800, December 2023*
**DOI: 10.22036/pcr.2022.365408.2218**

# Structural Insight into 2-Aryl-4-Quinoline Carboxylic Acid-Based Dihydroorotate Dehydrogenase (DHODH) and its Potential Anti-SARS-CoV-2 Activity Through Pharmacophore Modeling, Multidimensional QSAR, ADME, and Docking Studies

B.R. Thorat[a], V.P. Purohit[b], R.S. Yamgar[c], D. Bhagat[d], S.D. Wavhal[b] and S. Mali[e,*]

[a]*Department of Chemistry, Government College of Arts and Science, Aurangabad, (M.S.)-431001, India*
[b]*Ismail Yusuf College of Arts, Science, and Commerce College, Mumbai-400060, India*
[c]*Department of Chemistry, Patkar Varde College, Goregaon, Mumbai, India*
[d]*Institute of Forensic Science, Aurangabad, India*
[e]*Government College of Pharmacy, Karad, Maharashtra, India*
*(Received 12 October 2022, Accepted 12 November 2022)*

Dihydroorotate dehydrogenase (DHODH) is a rate-limiting enzyme in the biosynthesis of pyrimidine, which catalyzes the oxidation of dihydroorotate to orotate. Uridine monophosphate is biosynthesized by orotate. DHODH inhibitors have been shown to have antiviral activity against cytomegalovirus, Ebola, influenza, Epstein-Barr virus, and picornavirus. The anti-SARS-CoV-2 activity of DHODH inhibitors has also been investigated. DHODH inhibitors, including leflunomide and its metabolite teriflunomide, have been found to have anti-SARS-CoV-2 activity. In relation to the importance of this enzyme (*i.e.*, DHODH) in drug design, the present study aimed to develop statistically robust and interpretable 2D- and 3D-quantitative structure-activity relationship (QSAR) models based on a dataset of 92 molecules of biologically active 2-aryl-4-quinoline carboxylic acid analogs, reported as DHODH inhibitors. The correlation coefficient ($R^2$) values of the training set of the partial least squares (PLS) and all five Kernel-based PLS models for the respective fingerprints were found to be 0.7091, 0.8336 (linear), 0.7586 (radial), 0.8606 (dendritic), 0.6832 (desc), and 0.7670 (Molprint 2D), respectively ($R^2 \approx 0.9$). However, the external validation coefficient ($Q^2$) values of the test set were found to be 0.7009, 0.7503 (linear), 0.7737 (radial), 0.8250 (dendritic), 0.6756 (desc), and 0.7533 (Molprint 2D), respectively ($Q^2 > 0.6$). The developed 4-point pharmacophore model (ARRR_1), with one hydrogen bond acceptor and three aromatic rings, was found to be crucial in preserving the activity of 2-aryl-4-quinoline carboxylic acid analogs as DHODH inhibitors. Furthermore, the molecular docking of DHODH inhibitors against SARS-CoV-2 target proteins revealed the significant role of DHODH inhibitors.

**Keywords:** Dihydroorotate dehydrogenase, Molecular modeling, QSAR, CADD, Structural features, 2-Aryl-4-quinoline carboxylic acid analogs

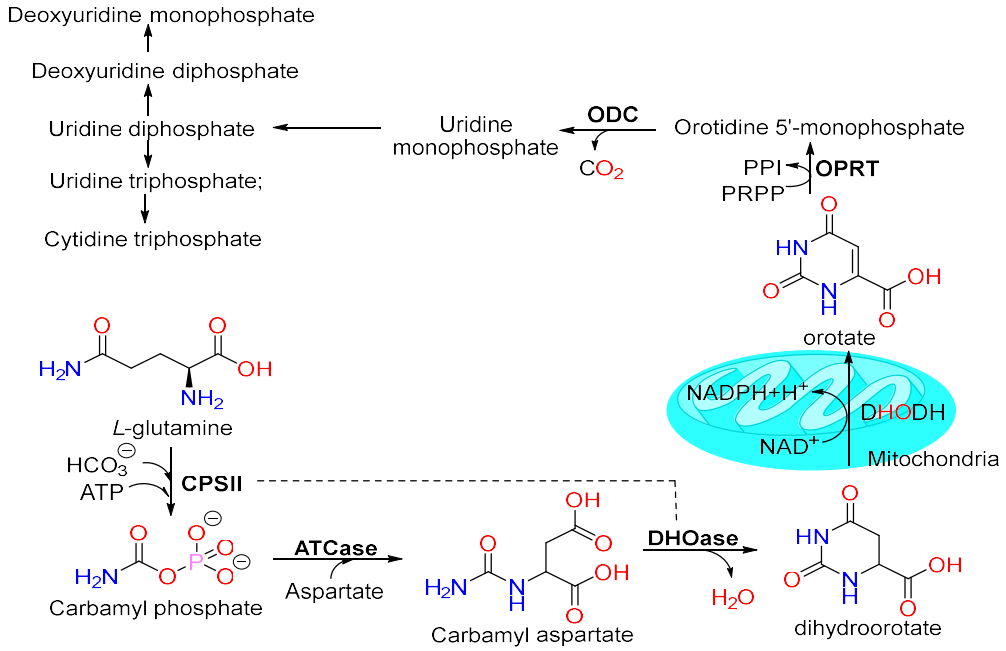
## INTRODUCTION

Dihydroorotate dehydrogenase (DHODH), as an iron-containing flavin-dependent enzyme found in the inner membrane of mitochondria, is the most researched and distinct therapeutic target among several viral agents [1]. Dihydroorotate was converted to orotate by the fourth enzyme for the de novo pyrimidine biosynthetic process [2]. These pyrimidines are necessary for the production of certain phospholipids and nucleic acids, such as RNA and DNA [2]. De novo synthesis and salvage are the two main mechanisms for the biosynthesis of pyrimidines.

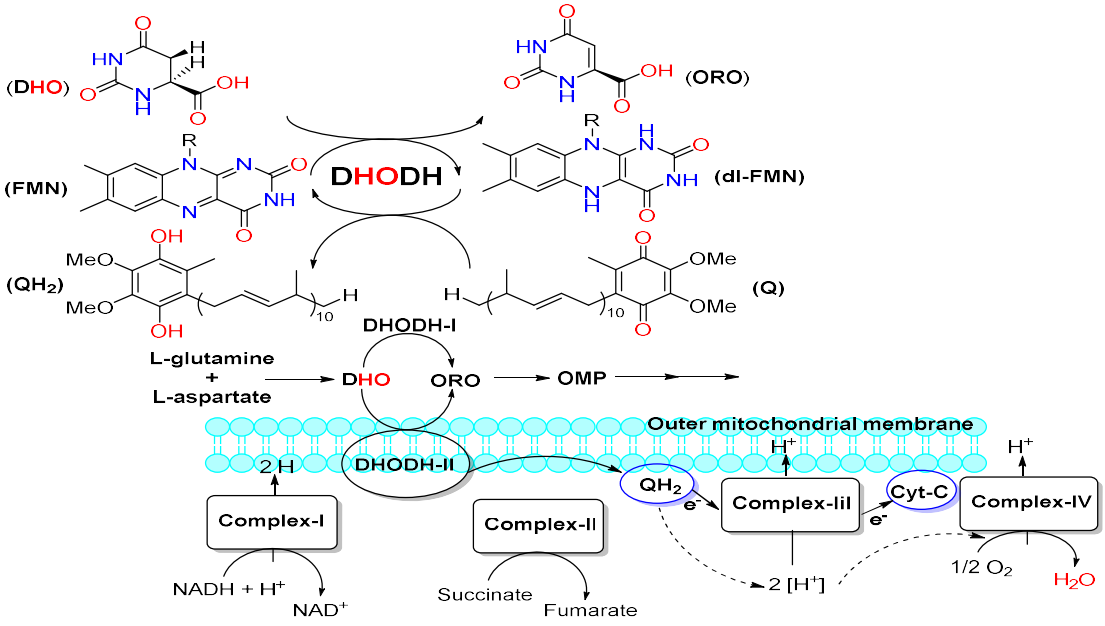
Pyrimidines are biosynthesized *via* the salvage synthesis pathway in resting or fully differentiated cells whereas highly proliferative cells, such as tumor cells, where the demand for

*Corresponding author. E-mail: mali.suraj1695@gmail.com

the nucleic acid precursors and other cellular components is higher, have been reported to be more active in the de novo biosynthesis [3]. DHODH inhibitors impact pyrimidine biosynthesis, lower pyrimidine levels, and delay or stop cell cycle progression at the S-phase, which calls for an adequate concentration of pyrimidines [4]. Numerous treatments for many diseases, including parasite disorders, have been triggered by pyrimidine depletion *via* DHODH inhibition in host cells. The detailed pathway involved in DHODH biosynthesis is illustrated in Schemes 1a and 1b [4].

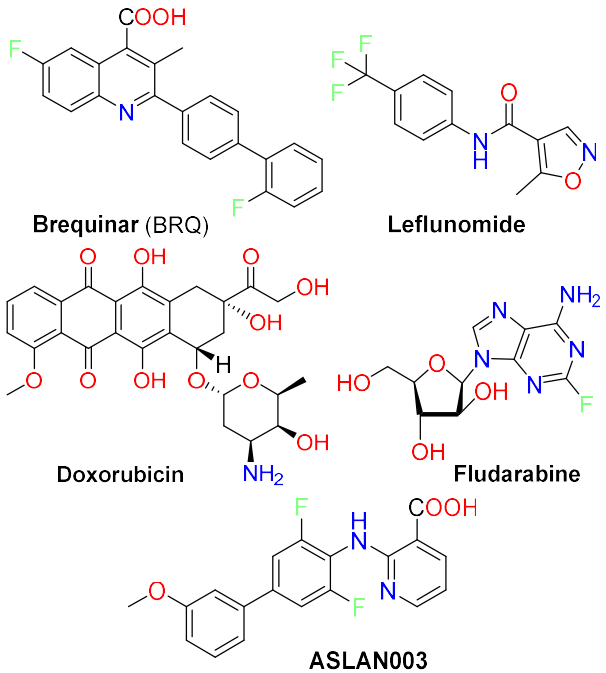



*Scheme 1.* (a) Pyrimidine de novo biosynthesis pathway [4]

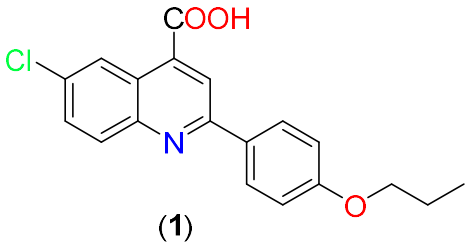


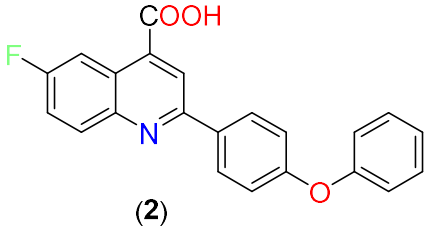

*Scheme 1.* (b) The role of dihydroorotate dehydrogenase (DHODH) in the de-novo pyrimidine biosynthesis and the mitochondrial respiratory chain [4]

Due to the fast DNA replication and nucleic acid production, the need for nucleic acid precursors increases rapidly in proliferative cells such as activated T cells. This process triggers the synthesis of pyrimidines activating the de novo pathway [4]. As a result, the role of DHODH in rapidly proliferating cells, such as cancer cells that are especially vulnerable to nucleotide biosynthesis inhibition, has gained increasing prominence. Cancer cells are far more susceptible to DHODH inhibitors than normal cells in a nutrient-depleted microenvironment and tumor hypoxia. Brequinar, leflunomide (HR486), which lowers the proliferation of CRC cells, doxorubicin, fludarabine, ASLAN003 (Phase II clinical trial), and other RNA interference or DHODH-blocking inhibitors have been shown to either stop cell growth or sensitize cells, also known as the anti-proliferation effect [5]. A wide range of quinoline-based DHODH inhibitors have been previously investigated [6]. Brequinar was found to inhibit human (h) DHODH in A549 cells infected with dengue virus and other flaviviruses, such as West Nile virus, yellow fever virus, vesicular stomatitis virus, and Western equine encephalitis virus. Brequinar is one of the broad-spectrum anti-viral drugs that inhibit the replication of evolutionarily diverse viruses by inhibiting a cellular factor or host pathways [7].




In addition, brequinar inhibits dengue virus RNA synthesis by disrupting intracellular pyrimidine pools. Another 4-quinoline carboxylic acid analog (1) was shown to have antiviral activity ($EC_{50}$ = 4650 nM against vesicular stomatitis virus (VSV)) and inhibit hDHODH ($IC_{50}$ = 260 nM), which is linked to antiviral activity. Based on the structure-activity relationship (SAR) study, analog (2) demonstrated a strong inhibitory effect against viruses ($EC_{50}$ = 2 nM for VSV and 41 nM for viral influenza (WSN)) and hDHODH ($IC_{50}$ = 1.0 nM), and its effect was more potent than that of brequinar.




In the present study, robust and statistically validated quantitative structure-activity relationship (QSAR) models, along with a pharmacophore model, were developed for the reported series of DHODH-blocking inhibitors using the Schrodinger suite (2020) for drug discovery [8-17]. To identify the structural requirements for the current scaffold, the following models were applied: 1D/2D and 3D-QSAR models, such as multi-linear regression, partial least squares (PLS), and Kernel-based PLS (KPLS) models as 2D QSAR models, and atom- and field-based 3D-QSAR models, as well as pharmacophore alignment-based 3D-QSAR models [8-10].

Furthermore, we used molecular docking analysis to examine whether these DHODH inhibitors had plausible correlations with different varieties of SARS-CoV-2 targets.

## MATERIALS AND METHODS

Dataset A series of biologically active quinoline carboxylic acid analogs that were previously reported as inhibitors of human DHODH were selected for the QSAR study [6-7]. All 92 molecules (*e.g.,* $R_1$ = COOH, COOR, CHROH, $CHCF_2$, COR, aryl, CON; $R_2$ = H, Me; $R_3$ = H, F, Cl, $CF_3$; $R_4$ = alkyl, alkoxy, aryloxy, halo, amino; $R_5$ & $R_6$ = H, alkyl, alkoxy; and X = CH/C/N) had diverse substitution patterns around the quinoline ring (Fig. 1). All these compounds were reported as DHODH inhibitors against the host cell targeted by different viruses or infected with different types of cancer. In this study, we focused on DHODH inhibition, which is reported in SI. Table S1 as half maximal inhibition concentration ($IC_{50}$ reported in nM) values. All 92 selected molecules were divided into two sets of test and training sets, and their biological activities were converted to $pIC_{50}$.

### Ligand Preparation and Docking Methodology

Structures of all the molecules were drawn in ChemBioDraw Ultra 12.1, and their 'SMILES' were copied and then utilized for the structure generation in the MASTERO interface (version 12.6) [16-17]. Geometry optimization, hydrogen addition, energy minimization, and ligand preparation were done by the same tool (Maestro 12.6). The lowest energy ionization of conformers was also considered. This led to the development of lowest-energy conformers structures for a given set of ligands at a defined range of pH (7.0 ± 2.0), and the developed structures were optimized for the optimized potentials for liquid simulations (OPLS-2005) force field for the simulation of ligands. Molecular docking was then carried out for the selected molecules using Glide (v. 8.9). Overall, molecular docking comprises the following five main steps: 1) protein preparation, 2) ligand preparation, 3) receptor grid generation, 4) ligand docking procedure, and 5) examination of the docking results. Properties generated by Prime molecular mechanics/generalized Born surface area (MM/GBSA) are helpful in determining the energies of ligand-receptor complexes. The detailed protocol with coordinates is provided in the Supplementary materials.

### QSAR Modeling, Validation, and Common Pharmacophore Hypothesis (CPH) Generations

After the preparation of ligands, as described above, different types of regression, atom- and field-based QSAR models were developed using the PHASE module (2021). All the models were developed through the random/seed selection of the test set and training set into 30%:70% by software, as described in the literature [16-17]. In all models, activity ($pIC_{50}$ value) was considered as the dependent variable, and other descriptors were considered as independent variables. Details of dataset splitting and QSAR calculations are discussed in the 'Results and Discussion Section'. The validation of QSAR models was carried out using the statistical validation criteria reported in the literature [16-17].

### Adsorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) Predictions by QikProp

QikProp module (QikProp, v. 6.6, Schrödinger, LLC)

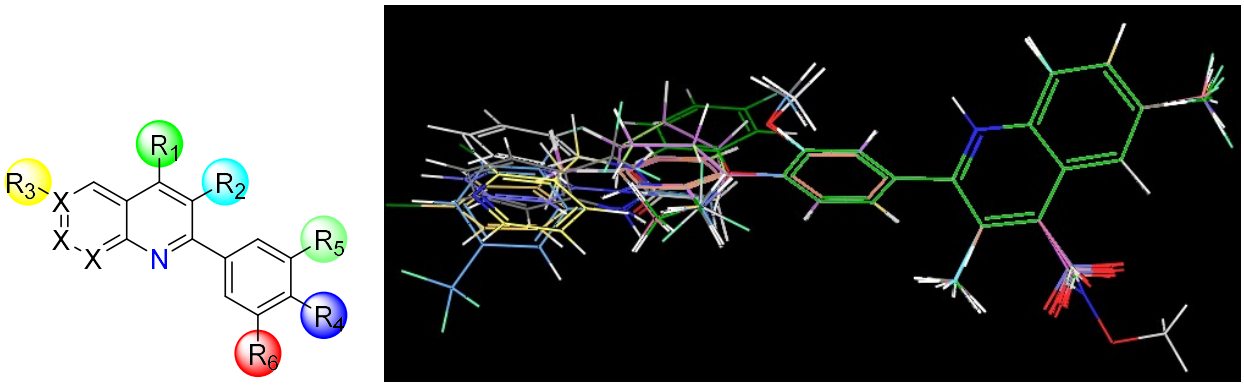



**Fig. 1.** Basic and aligned structures of 2-aryl-4-quinoline carboxylic acid analogs used in the current study.

was used to predict ADMET properties [16-17]. Today, QikProp is used to make predictions about numerous physically and pharmaceutically significant descriptors of organic compounds. In addition to making predictions about the chemical properties of compounds, QikProp provides bounds for contrasting the properties of a certain molecule with those of 95% of known medications. Additionally, QikProp informs users of 30 different reactive functional group types that might lead to erroneous results in high-throughput screening tests [16-17].

## RESULTS AND DISCUSSION

### 1D/2D-Based QSAR Model Development and Validation

QSAR models, with 12 descriptors, were developed and initially considered in the test set for 27 molecules and later in the training set for 65 molecules. The values of experimental activities, predicted activities, and errors of these models are reported in Table S2. All prepared ligands (output files of LigPrep module) were opened in the CANVAS module of Schrodinger and their 819 descriptors were obtained from the molecular descriptor tool. Then, the $pIC_{50}$ values were added and descriptors were further allowed for the feature selection utility to remove the constant, highly correlated, and informative descriptors ($R^2 > 0.8$). A total of 12 descriptors were found to be useful and other descriptors were removed from the data automatically. The 12 descriptors included aOm_Cnt, aasC_Cnt, ssCH2_Cnt, ALOGP6, ALOGP9, PEOE13, PEOE14, PEOE7, Total_structure_connectivity, average_connectivity_index_chi-2, mean_topological_charge_index_of_order_4, and mean_topological_charge_index_of_order_8. These descriptors did not have any direct correlation with their inhibitory effect and subsumed physicochemical, quantum chemical, topological, and constitutional parameters. The values of these descriptors are reported in Table 1.

The QSAR model was developed by performing stepwise regression analyses and using the above 12 descriptors, as independent variables, and activity ($pIC_{50}$), as the dependent variable. All the QSAR data were analyzed using the multiple linear regression (MLR) method.

The quality and utility of regression models depend on different parameters such as $R^2$, $Q^2$, SD, RMSE, r-Pearson, etc. The quality of the developed model was evaluated using the squared correlation coefficient ($R^2 > 0.67$), $Q^2 > 0.5$, and root mean square error (RMSE). In the developed model, $R^2$ indicated how the equation fitted the data while $Q^2$ reflected the predictive performance and stability of the predicted QSAR model. Also, the cross-validation of the model parameters (*i.e.*, $R^2_{cv}$ or $Q^2$) was considered. The average

**Table 1.** The Values of Descriptors for Models 1 and 2 Developed by the MLR Method

| Variable | Coefficient | Std. Err. | T |
|---|---|---|---|
| Intercept | 5.4280e+000 | 2.1994e+000 | 2.4679 |
| aOm_Cnt | 8.7843e-001 | 1.8449e-001 | 4.7614 |
| aasC_Cnt | 5.8101e-001 | 1.1426e-001 | 5.0850 |
| ssCH2_Cnt | 3.6916e-002 | 1.8444e-002 | 2.0015 |
| ALOGP6 | -4.5091e-003 | 2.4123e-003 | 1.8692 |
| ALOGP9 | 6.8270e-003 | 2.5830e-003 | 2.6431 |
| Average_connectivity_index_chi-2 | -1.1122e+001 | 4.5563e+000 | 2.4410 |
| Mean_topological_charge_index_of_order_4 | -2.5891e+001 | 6.1408e+000 | 4.2162 |
| Mean_topological_charge_index_of_order_8 | -8.1773e+001 | 2.2845e+001 | 3.5795 |
| PEOE13 | 2.2274e-001 | 6.0819e-002 | 3.6623 |
| PEOE14 | -6.1389e-002 | 2.6792e-002 | 2.2913 |
| PEOE7 | -2.6604e-002 | 6.3641e-003 | 4.1803 |
| Total_structure_connectivity | -1.4776e+003 | 4.0945e+002 | 3.6087 |

ranking score of model 1 was found to be 0.687155, with the number of independent variables being 06. Other parameters are presented in Table 2. To validate the data set selected for the model development, the molecules were divided into two sets of the training set and the test set for external predictions. The best regression model was predicted by observing the standard deviation (s) and the multiple correlation coefficients ($R^2$).

The applicability domain of the developed model was visualized by the Williams plot of the training and test sets of molecules for the MLR model (Fig. 2). Figure 2 shows the goodness of fit between the actual (observed) and calculated (predicted) activities of both test and training sets of molecules.

QSAR equations were developed by considering different combinations of the available descriptors. A total of 27 2-aryl-4-quinoline carboxylic acid analogs were selected and placed in the test set. As discussed earlier, 65 analogs were put in the training set. Also, as mentioned previously, 12 independent variables (descriptors) were used to develop the MLR model and establish the correlation between the $pIC_{50}$ and the selected descriptors. The correlation coefficient (standard deviations) value for each descriptor was examined, and it was found that the highest correlation coefficient belonged to the aOm_Cnt. Model 1 was developed to obtain the inhibitory concentrations of 2-aryl-4-quinoline carboxylic acid analogs using 92 molecules. Model 1, along with its six descriptors, is presented below.

$pIC_{50}$ (pred.) = 5.4280e + 000(2.1994e+000) + 8.7843e-001(1.8449e-001) aOm_Cnt + 5.8101e-001(1.1426e-001) aasC_Cnt + 3.6916e-002(1.8444e-002) ssCH2_Cnt - 4.5091e-003(2.4123e-003) ALOGP6 + 6.8270e-003(2.5830e-003) ALOGP9 - 2.6604e-002(6.3641e-003) PEOE7 Model 1

**Table 2.** PLS Parameters of the MLR Models of 2-Aryl-4-quinoline Carboxylic Acid Analogs

| MLR | Training set | | Test set | |
|---|---|---|---|---|
| Model | S.D. | $R^2$ | RMSE | $Q^2$ |
| | 0.7871 | 0.6743 | 0.6870 | 0.7161 |

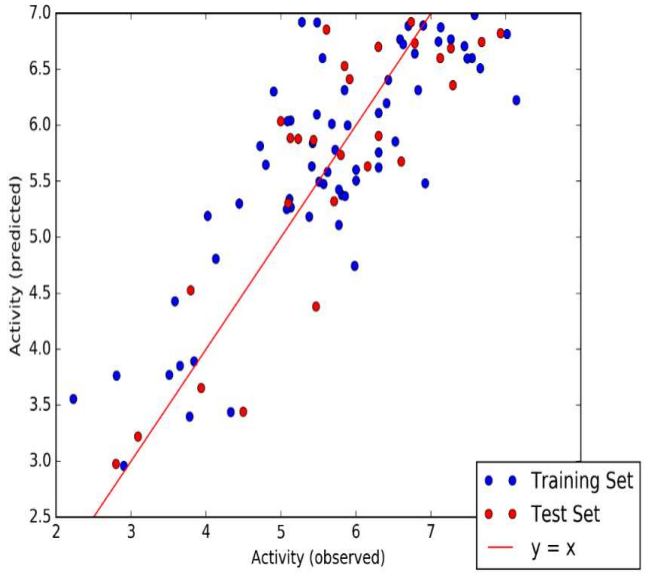



**Fig. 2.** A 2D-MLR model plot of observed activity *vs.* predicted activity ($pIC_{50}$) of A, P01-P91 of 2-aryl-4-quinoline carboxylic acid analogs.

All statistical parameters for the model were obtained within the acceptable limits explaining the external and internal predictive potential of the developed QSAR model. The lower RMSE value predicted that the developed model was good enough for further consideration and use. Fingerprints used in PLS and KPLS models were a set of descriptors (binary, numeric, categorical, *etc*.) that could be easily assembled into a string (characteristics of the molecules). A fingerprint-based method was used to generate a wide variety of possibilities including fragment-based similarity searching, bit scaling rules, atom/bond typing schemes, and similarity indices. PLS and KPLS-QSAR models were generated using the above-mentioned fingerprints as molecular descriptors. Both PLS and KPLS predictive 2D-models were generated in the CANVAS module of the Schrodinger suite.

In its original domain of chemometrics, the PLSR method has been referred to as one of the most popular regression, classification, and discrimination techniques. It is a popular statistical tool specifically designed for dealing with multiple regression problems where the number of observations is limited, missing data are nonignorable, and the correlation between the predicted variables is high. It is basically more

applicable when the number of observed variables is much higher than the number of observations and a high multicollinearity exists among the variables. Similarly, in the Kernel-based model and regression development, the original data are mapped to a high-dimensional feature space. In a high-dimensional predictive model, overfitting may lead to a decrease in the prediction accuracy of the external data. Both PLS and nonlinear PLS algorithms have been utilized for model development.

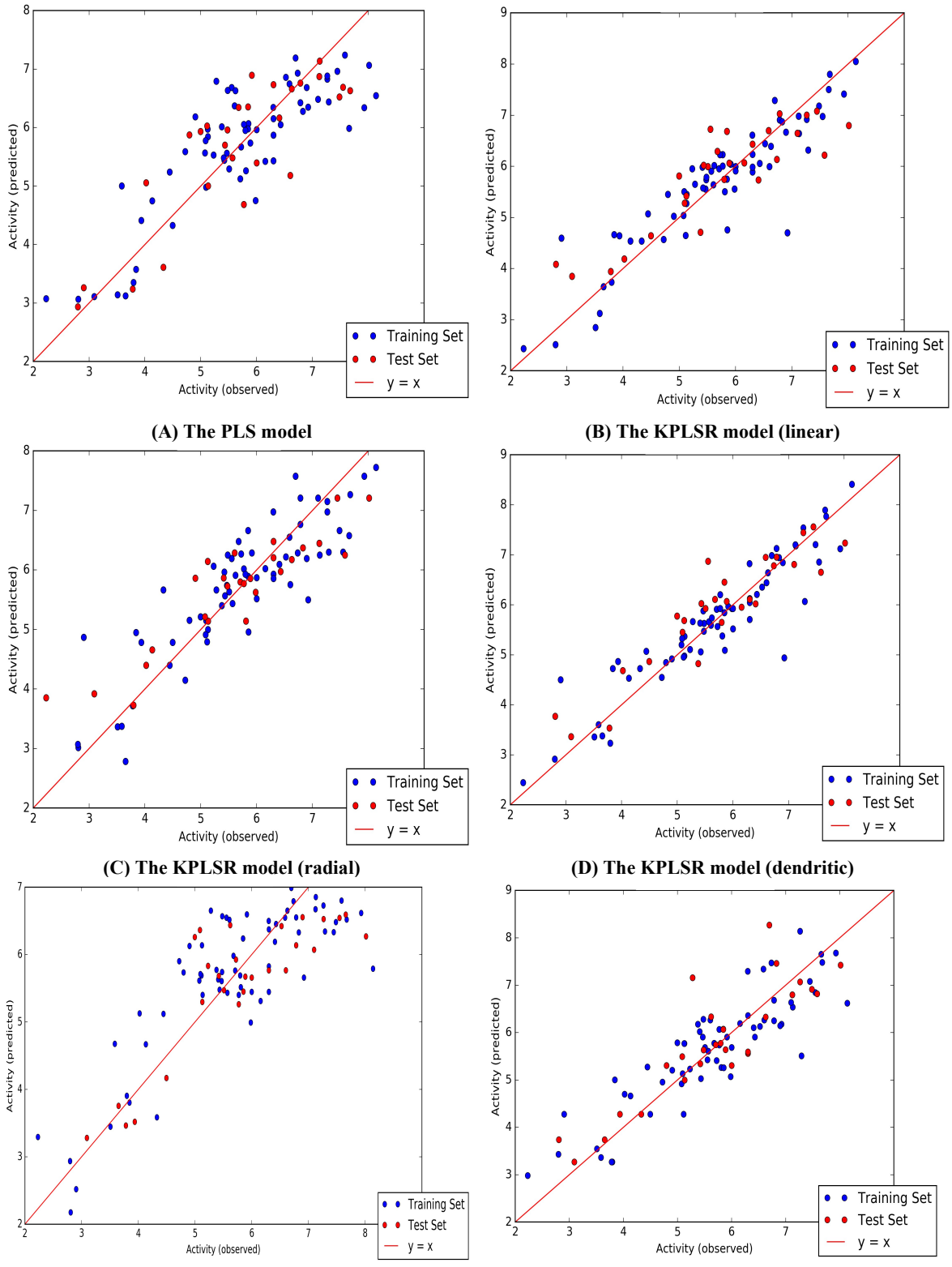
A KPLS model was found to be more competitive and even more stable than other kernel algorithms, such as kernel ridge regression and support vector machines. In the present study, 92 (2-aryl-4-quinoline carboxylic acid) analogs, which their minimum inhibition concentration for DHODH inhibition was reported previously in the literature, were used. The structures of all these compounds are available in *the supplementary materials*. In total, 12 descriptors were considered and the model was trained by the training set (65 molecules) and validated independently by the test set (27 molecules) (Table S3). All the molecules were energy-minimized and then used to define the 2D fingerprint, which was used as an X variable combined with the $pIC_{50,}$ for the KPLS QSAR model. A KPLS model was developed by keeping the Y-variable as the experimental activity and X-variable as the fingerprint with selected independent variables (descriptors) depending on the fingerprint. Five (hashed) types of fingerprints, including linear/radial/dendritic/desc/molprint, of the seven available fingerprints, were selected and considered as the structural MACCS keys of the study. The number of descriptors selected was large enough to develop the intended models. The maximum number of KPLS factors was set to 1 and a nonlinear kernel was defined at 0.05. The selected molecules were divided randomly into training and test sets with a ratio of 70%:30%. Bootstrapping was conducted for the random sampling of the training set with replacement to create a new set with the same size. This was repeated until the expected results for the QSAR descriptors were obtained (Table S4).

KPLS is a nonlinear Gaussian model with a higher predictive ability than conventional property-based models. In other words, it is an extension of the PLS regression that introduces some non-linearity to the scalar products of X-variables. The best KPLS regression model for the fingerprints was selected. This was shown by the good values of the parameters of the final model, the high predictive ability of the model for the external test set, and low uncertainties. The models predicted the DHODH inhibitors with the highest accuracy. The ADME and other properties of the analogs are reported in supplementary files. All five fingerprint descriptors resulted in the best models of the dendritic fingerprint with a better ability for the test set molecules and lower uncertainty. The correlation coefficient ($R^2$) values of the training set of the PLS and all five KPLS models for the respective fingerprints were found to be 0.7091, 0.8336 (linear), 0.7586 (radial), 0.8606 (dendritic), 0.6832 (desc), and 0.7670 (Molprint 2D), respectively, while the external validation coefficient ($Q^2$) values of the test set were calculated as 0.7009, 0.7503 (linear), 0.7737 (radial), 0.8250 (dendritic), 0.6756 (desc), and 0.7533 (Molprint 2D) with lower values of uncertainties, respectively. The residual errors, SD, and RMSE, presented in Table S3, were overall acceptable for all the above models (Fig. 3). Among the five fingerprints, the dendritic fingerprint gave the best performance. Therefore, it was assumed that fingerprint descriptors (with the priority given to the linear and dendritic fingerprints) could be used for modeling and that they could lead to optimization with the selected set of molecules as DHODH inhibitors.

Among the five models, the KPLS model with the dendritic fingerprint was found to be the best model with good coefficient and KPLS regression curves with a high predictive ability for an external test set of data. Figure 3 shows model correlation plots for 2-aryl-4-quinoline carboxylic acid analogs for PLS and KPLS for five fingerprints. The contributing effects of each atom and fragment are also shown in 2D plots. The analysis of the plot confirmed that the linear and dendritic fingerprint model gave the best agreement between the experimental and predicted DHODH inhibition of 2-aryl-4-quinoline carboxylic acid analogs.

## 3D QSAR Modeling (Atom- and Field-Based QSAR)

After the preparation of ligands, all ligands were superimposed on each other using 'flexible ligand alignment' from MAESTRO. All 92 analogs of 2-aryl-4-quinoline carboxylic acid were used for the generation of atom-based and field-based QSAR models using Phase. All the molecules were divided into the test and training sets based on a 30:70




**(A) The PLS model**

**(B) The KPLSR model (linear)**

**(C) The KPLSR model (radial)**

**(D) The KPLSR model (dendritic)**

**(E) The KPLSR model (desc)**

**(F) The KPLSR model (molprint2D)**

**Fig. 3.** A 2D-MLR model plot of observed activity *vs*. predicted activity ($pIC_{50}$) of A, P01-P91 of 2-aryl-4-quinoline carboxylic acid analogs.

ratio (*i.e.*, 27 test molecules and 65 training molecules). All training and test sets of molecules with a PLS factor of 3 were considered for the development of atom-based QSAR models. In addition, three field-based QSAR models were developed (Phase, Schrödinger, 2018). To develop the free energy force field (FEFF) 3D-QSAR model, 27 molecules in the test set and 65 molecules in the training set were used. The observed and predicted activities of all three models are shown in Tables S5 and S6. Grid spacing was kept as 1 A extended by 3 Å beyond the training set atoms. The steric force field within 2 Å of any training set atom was ignored while the PLS factor for the atom-based QSAR and FEFF 3D-QSAR was kept at 3.

Different molecular features, such as H-bond donor, hydrophobic/non-polar, and electron-withdrawing, were used for the generation of an atom-based QSAR model. A Gaussian-based QASR model was used for the generation of field-based QASR models, which included five different features such as Gaussian hydrophobic, Gaussian electrostatic, Gaussian steric, Gaussian H-bond acceptor, and Gaussian H-bond donor groups. In developing field-based models, a truncate steric force field of 30.0 kcal $mol^{-1}$ was used, and variables with StdDev < 0.01 and with | t value | < 2.0 were eliminated. Gaussian intensities (descriptors) were utilized as independent variables in field-based models. The best developed QSAR models were further selected on the basis of the robustness of statistical parameters (Tables S7 and S8).

## Validation of 3D-QSAR Models

The stability, robustness, and predictive power of the models were examined by leave-one-out (LOO) cross-validation. The best QSAR model was selected for QSAR visualizations. The predictive power of the selected model was determined from the RMSE, $Q^2$, and Pearson's r of the test set molecules as well as the $R^2$ and $R^2_{CV}$/scramble of the training set (Table S9) [16-17].

## Mechanistic Interpretations of the Generated QSAR Models

**Field-based 3D-QSAR models.** Scatter plots for the predicted vs. actual activities of the training, test, and all sets of molecules for the best atom- and field-based models are visualized in Fig. 4. Contour map analysis was used for the

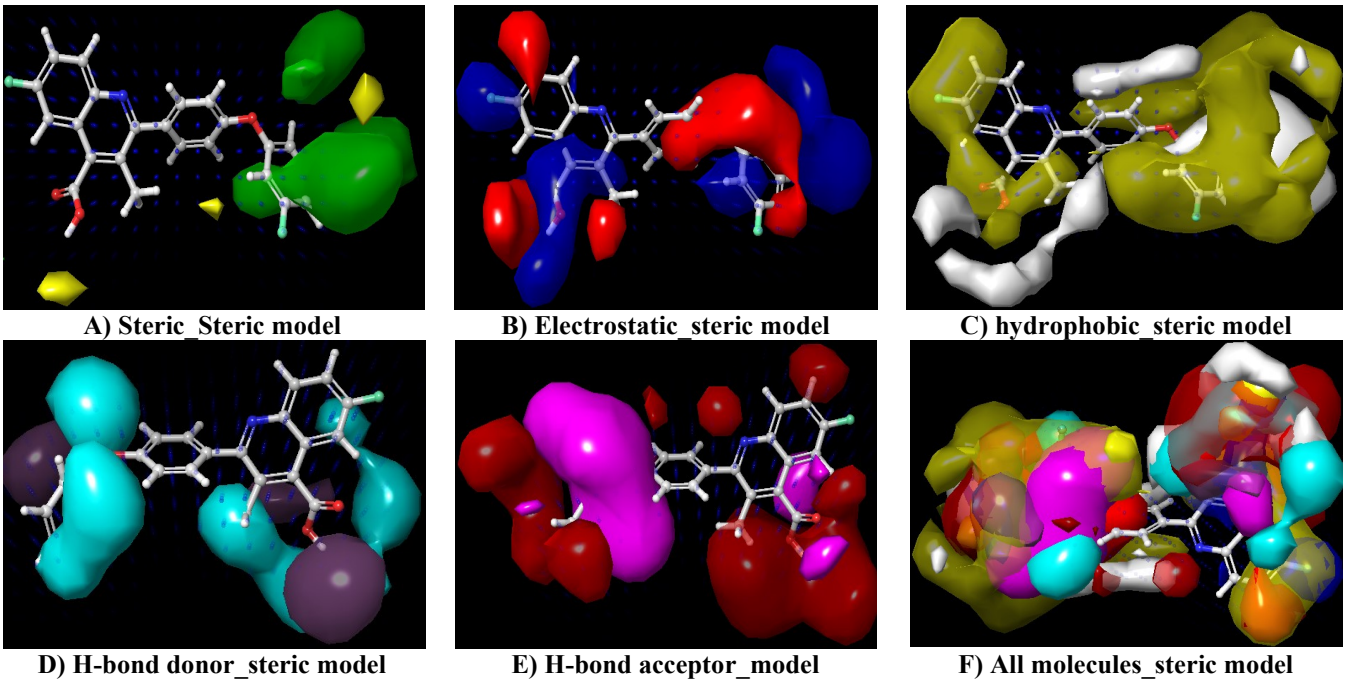



**Fig. 4.** Field-based QSAR models A) Gaussian steric field: favored (green), unfavorable (yellow); B) Gaussian electrostatic fields: favored electropositive (blue), disfavored electronegative (red); C) Gaussian hydrophobic field: favored (yellow), disfavored (white); D) Gaussian Hydrogen-bond acceptor field: favored (blue), disfavored (magenta); E) Gaussian Hydrogen-bond acceptor field: favored (red), disfavored (magenta); and F) the presence of all types of force fields.

development of the best model (Fig. 4). Regions with green color in the counter map are favored and show a positive contribution to the activity while yellow regions are disfavored and show a negative contribution to the activity. Sterically hindered groups on the $R_4$ position that are attached to aryl rings enhanced the activity while other sterically hindered groups did not affect the effect of DHODH inhibition. H-bond donor and acceptor groups attached to the side chain at $R_2$-groups ($C_3$-position), $R_1$-groups (CO-position), and $R_4$-groups (aryl substitutions) altered the value of activity; that is, change in activity was dependent on the nature of H-donor and acceptor groups, as shown in Figure 4. The groups or atoms with hydrophobic effects marginally affected activity at a maximum position, which is thought to be controlled or balanced by other electronic substituent properties. The electrostatic effect at the $R_2$, $R_1$, $R_4$, and $R_5$ positions of the aryl ring denoted a negative contribution while the C3 and C5 groups of the aryl ring at $R_4$, $R_1$, and $R_3$ positions displayed a positive contribution in the activity. The statistical parameters and field fractions of the field-based QSAR model showed good results regarding $Q^2$ and $R^2$ [16-17].

The scatter plots of predicted activity vs. actual activity of all three models and the training and test set molecules of the model for the field-based QSAR models are shown in Fig. 5.

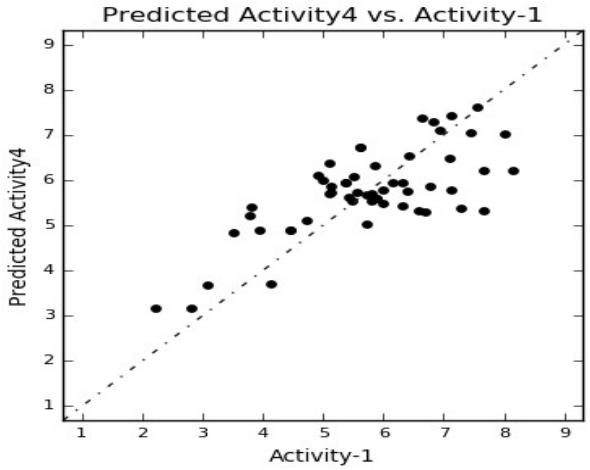
**Atom-based 3D-QSAR models.** In the atom-based QSAR models, an increase in activity was correlated with blue-colored cubes while red-colored cubes indicated a decrease in activity. A region with blue color in the counter map is highly important for the activity index. In Schrodinger software, output results or occlusion maps of the atom-based QSAR model are in the form of blue, red, or both color cubes. Herein, we considered molecule A (standard) as a reference for the prediction of models. In model 3 of molecule A, it was clearly observed that the quinoline ring, with substituents at $C_4$ ($R_1$-group), $C_7$, and $C_8$ positions and the aryl group (all positions except $C_2$), attached to the $R_4$-position represented blue color cubes, indicating that electron-withdrawing functionalities may favor the activities. However, electron-withdrawing groups at the hydrogen of the carboxylic group, $C_5$, and $C_6$-positions of quinoline rings and the $C_2$-position of the aryl ring attached to the $R_4$-position represented the red color maps, indicating that further electron-withdrawing substituents may disfavor the activity. Also, red cubes around the carboxylic group and $C_7$ position of the quinoline ring and the $C_3$-position of the aryl ring attached to the $R_4$-position indicated that the H-bond donor atom at this position disfavored or denoted negative contribution to the activity while the H-bond donor atom at the –OH group of carboxylic acid group, $C_8$-position of the quinoline ring, and $C_2$-position of the aryl ring attached to the $R_4$-position enhanced the activity. Hydrophobic groups attached to the quinoline ring at N, $C_5$, and $C_6$-positions and the alkyl/aryl group attached to the $R_4$-position favored the activity while hydrophobic substitutions at other positions of the quinoline ring disfavored the activity. A visual representation of the atom-based QSAR model for reference molecule A is shown in Fig. 6.

The scatter plots of predicted activity *vs*. actual activity of all three models and the training and test set molecules of model 3 for the 3D atom-based QSAR models are shown in Fig. 7.

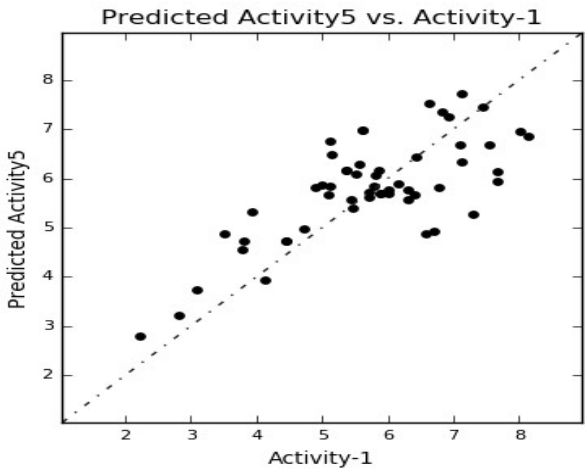
Both 3D atom-based and field-based QSAR models demonstrated good predictive abilities with good robustness. These models may be useful in designing quinoline-based DHODH inhibitors for the treatment of various infections.

## Pharmacophore Modeling and Analysis

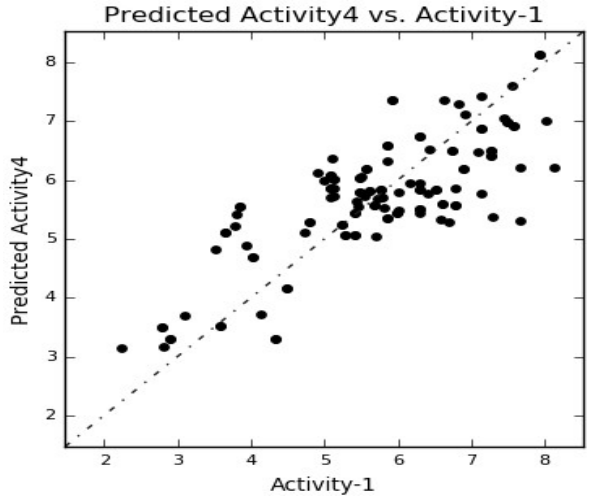
A different molecular environment, such as hydrogen, atoms, and skeletons, usually becomes an acceptor, donor, ionizable, hydrophobic, or aromatic ring and is important for designing a pharmacophore model [18-21]. A PHASE tool can be used for the generation of H-bond donor (D), H-bond acceptor (A), hydrophobic group (H), negatively ionizable (N), positively ionizable (P), and aromatic ring (R) features. In this study, a five-point pharmacophore model (ARRR_1), with one H-bond acceptor and three aromatic rings, was generated. All molecules were merged using flexible ligand arrangement and classified into two sets, such as active and inactive sets, with those molecules having $pIC_{50} > 2.20$ being considered as active 2-aryl-4-quinoline carboxylic acid analogs (92 molecules) and those having $pIC_{50} < 2.20$ being considered as inactive molecules. A minimum of four and a maximum of six pharmacophore features were selected for the development of the study hypothesis. The hypothesized difference was kept at 0.50 by default.



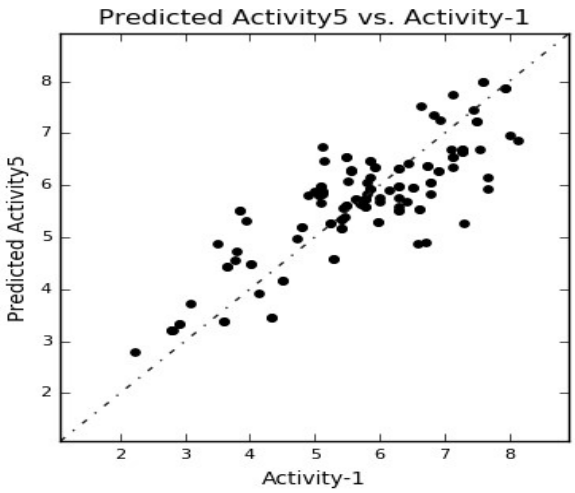

**A) QSAR plot of the test set of model 4**



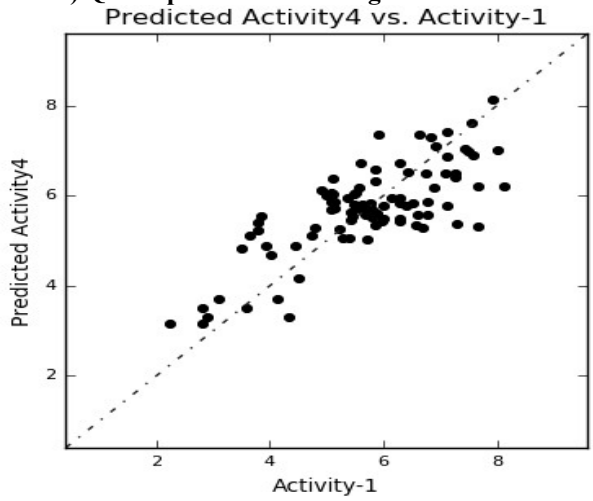

**B) QSAR plot of the test set of model 5**



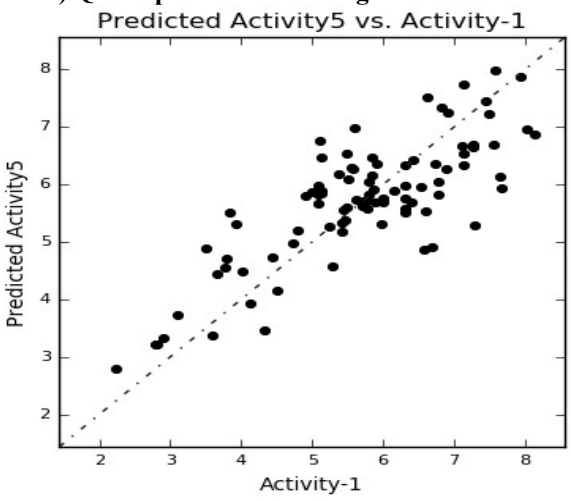

**C) QSAR plot of the training set of model 4**




**D) QSAR plot of the training set of model 5**




**E) QSAR plot of model 4**




**F) QSAR plot of model 5**

**Fig. 5.** Field-based QSAR model plots of actual vs. predicted $pEC_{50}$ values of A) QSAR plot of the test set of model 4, B) QSAR plot of the test set of model 5, C) QSAR plot of the training set of model 4, D) QSAR plot of the training set of model 5, E) QSAR plot of model 4, and F) QSAR plot of model 5.

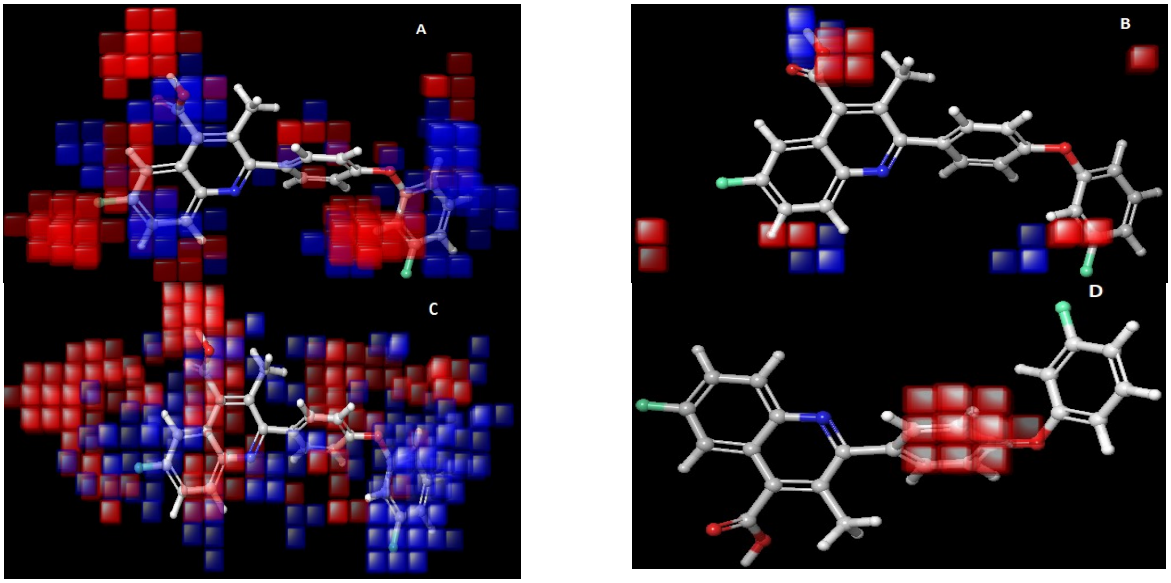
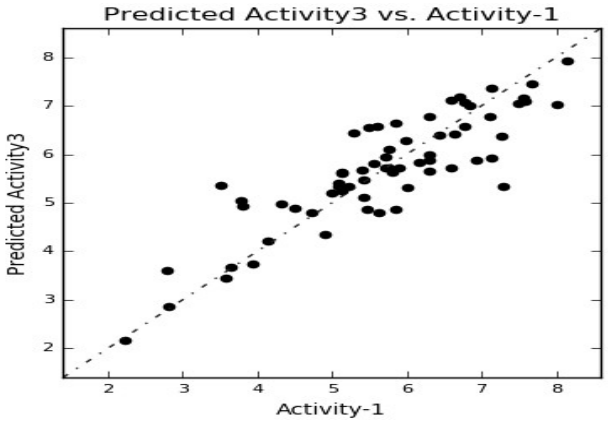
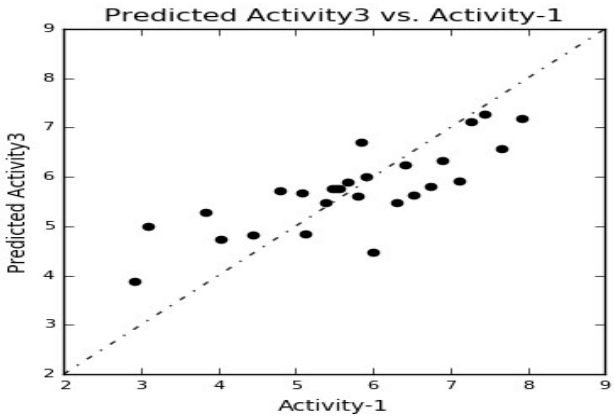


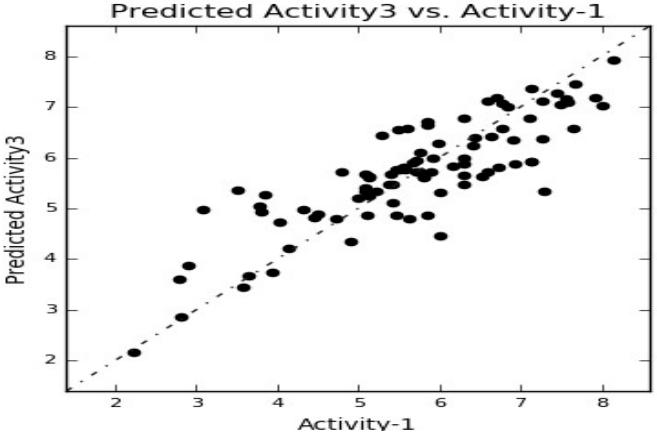

**Fig. 6.** Visual representation of atom-based PHASE 3D-QSAR model of the reference molecule A: A) electron withdrawing, B) H-bond donor, C) hydrophobic effect, and D) other effects. Blue cubes indicate a positive coefficient or an increase in activity while red cubes indicate a negative coefficient or a decrease in activity.




**A) QSAR plot of the training set of model 3**




**B) QSAR plot of the test set of model 3**




**C) QSAR plot of model 3**

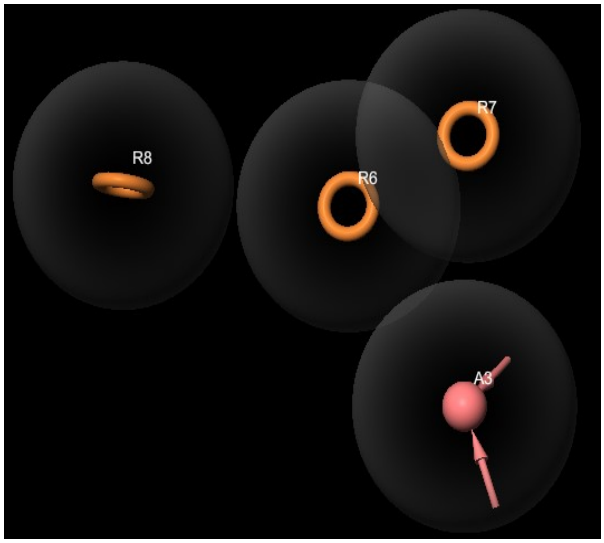
**Fig. 7.** 3D atom-based QSAR model plot of actual *vs*. predicted $pEC_{50}$ values of A) training set of model-3, B) test set of model-3, and C) model-3.

A total of 20 hypotheses were developed, and a top hypothesis was selected on the basis of the BEDROC score. All merged images of all active molecules were aligned on the common pharmacophore hypothesis ARRR_1. All 20 hypotheses were arranged based on their BEDROC scores, as presented in Table 3.

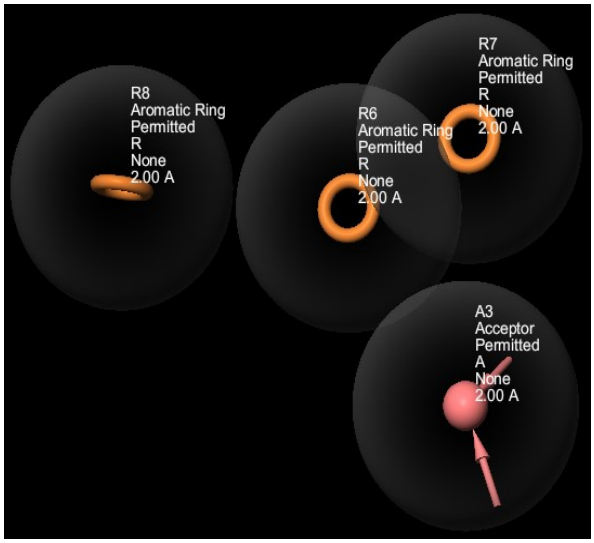
ARRR_1 pharmacophore model was selected based on its BEDROC score = 0.996 (99.6% accuracy), survival score = 5.726, and site score = 0.905 for developing 3D-contours with all the aligned active molecules. All 2-aryl-4-quinoline carboxylic acid analogs were aligned on the selected ARRR_1 pharmacophore model to examine whether the structural features correlated with the developed ranked hypothesis. Here again, reference molecule A was considered as the active molecule for the hypothesis ARRR_1 (Fig. 8).

**Table 3.** Different Pharmacophore Hypotheses (Arranged on the Basis of BEDROC score) Generated by PHASE

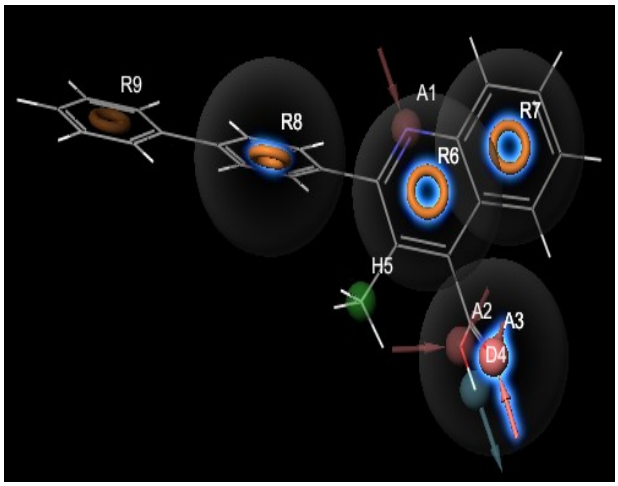
| Hypothesis ID | Hypothesis role | Survival score | Site score | Vector score | Volume score | Selectivity score | Num matched | Inactive score | Adjusted score | BEDROC score |
|---|---|---|---|---|---|---|---|---|---|---|
| ARRR_1 | Hypo | 5.726 | 0.905 | 0.892 | 0.738 | 1.232 | 91 | 0 | 5.726 | 0.996 |
| AARR_1 | Hypo | 5.672 | 0.922 | 0.893 | 0.736 | 1.163 | 91 | 0 | 5.672 | 0.996 |
| ARRR_2 | Hypo | 5.633 | 0.825 | 0.889 | 0.73 | 1.23 | 91 | 0 | 5.633 | 0.996 |
| AARR_2 | Hypo | 5.578 | 0.861 | 0.879 | 0.719 | 1.16 | 91 | 0 | 5.578 | 0.996 |
| AARRR_1 | Hypo | 5.887 | 0.787 | 0.888 | 0.747 | 1.541 | 84 | 0 | 5.887 | 0.964 |
| AARRR_2 | Hypo | 5.848 | 0.733 | 0.905 | 0.739 | 1.541 | 85 | 0 | 5.848 | 0.964 |
| AAARR_1 | Hypo | 5.779 | 0.786 | 0.888 | 0.748 | 1.433 | 84 | 0 | 5.779 | 0.964 |
| AAARR_2 | Hypo | 5.74 | 0.734 | 0.905 | 0.74 | 1.432 | 85 | 0 | 5.74 | 0.964 |
| AARRR_3 | Hypo | 5.687 | 0.602 | 0.905 | 0.715 | 1.541 | 84 | 0 | 5.687 | 0.964 |
| AAARR_3 | Hypo | 5.579 | 0.602 | 0.905 | 0.715 | 1.433 | 84 | 0 | 5.579 | 0.964 |
| AARR_3 | Hypo | 5.453 | 0.763 | 0.861 | 0.746 | 1.159 | 84 | 0 | 5.453 | 0.969 |
| AARR_4 | Hypo | 5.412 | 0.704 | 0.882 | 0.738 | 1.159 | 85 | 0 | 5.412 | 0.968 |
| AARRR_4 | Hypo | 5.477 | 0.469 | 0.815 | 0.693 | 1.56 | 87 | 0 | 5.477 | 0.963 |
| AARRR_5 | Hypo | 5.436 | 0.428 | 0.854 | 0.669 | 1.56 | 84 | 0 | 5.436 | 0.962 |
| AARR_5 | Hypo | 5.391 | 0.784 | 0.856 | 0.688 | 1.134 | 85 | 0 | 5.391 | 0.962 |
| AAARR_4 | Hypo | 5.376 | 0.468 | 0.816 | 0.695 | 1.457 | 87 | 0 | 5.376 | 0.963 |
| AAARR_5 | Hypo | 5.334 | 0.428 | 0.854 | 0.67 | 1.457 | 84 | 0 | 5.334 | 0.962 |
| AAAR_1 | Hypo | 5.304 | 0.774 | 0.859 | 0.706 | 1.036 | 85 | 0 | 5.304 | 0.964 |
| AARR_6 | Hypo | 5.349 | 0.806 | 0.927 | 0.593 | 1.098 | 84 | 0 | 5.349 | 0.961 |
| AAAR_2 | Hypo | 5.289 | 0.782 | 0.935 | 0.631 | 1.011 | 85 | 0 | 5.289 | 0.964 |




**A) Pharmacophore model**




**B) Pharmacophore model properties**




**C) Pharmacophore model with the ref. molecule A**

**Fig. 8.** A six-point pharmacophore model (ARRR_1) generated by PHASE.

## Molecular Docking and QikProp ADMET Predictions

**The relationship between DHODH enzyme inhibitors and COVID-19 targets.** Despite the strenuous effort made to develop a vaccine to combat SARS-CoV-2, no effective antiviral medication or vaccine has yet been developed to be able to prevent or cure COVID-19. Among coronaviruses, the primary protease (Mpro, also known as 3CLpro) is a desirable therapeutic target. Several effective inhibitors of the SARS-CoV-2 3CLpro and the association of their crystal structures with the protease have been identified and introduced. The production of DNA, RNA, phospholipids, and glycoproteins, which is necessary for both cell survival and proliferation, is made possible by the use of pyrimidines as key building components. Human DHODH is a flavin-dependent mitochondrial enzyme that catalyzes the oxidation of dihydroorotate to orotate, the fourth and rate-limiting step in the de novo production of pyrimidine-based nucleotides. HDHODH is a member of the class 2 DHODH family. As a result, DHODH is a desirable therapeutic target for a range of infections, such as cancer and autoimmune disorders. Recently, a study proposed the possible correlation between

COVID-19 and inhibitors of hDHODH [22]. Accordingly, in this study, we investigated the current dataset of DHODH inhibitors against COVID-19 targets.

There are about 178 protein structures of PDBs (PDB; http://www.rcsb.org/pdb/home/home.do) associated with DHODH enzymes, which, themselves, are associated with more than 16 different organisms to the best of our understanding. All these structures are determined by X-ray crystallography, either in complex with substrates, products, or inhibitors or in the hollow form for both native and mutant proteins. Selective DHODH inhibitors against infective microorganisms may be useful for the development of drugs against different infections.

In this study, dataset inhibitors were docked against SARS-CoV propane-like protease inhibitor, PLpro (PDB: 3MJ5), to evaluate the effectiveness of dataset molecules for any probable anti-viral activity against COVID-19 (Fig. 9). The strength of the 1:1 complex was determined using the docking score. The receptor parameters and docking scores of the top five ligands are provided in Tables S10 and S11. Moreover, in silico analysis of the best top hits retained good pharmacokinetic profiles.

All dataset DHODH inhibitors did not exhibit effective binding against SARS-CoV-2 propane-like protease target (3MJ5). While some of the DHODH inhibitors, including P19, P14, P39, P18, P45, P40, P35, P48, P37, P49, P30, and P42) were found to bind effectively to the target with weak interactions (lower docking scores), the remaining inhibitors docked with the absorption of energy (positive docking scores). It was concluded that the molecules found to have weak binding may not act as anti-viral agents against the receptors of SARS-CoV-2. However, with the help of the developed QSAR models, the top best-bound molecules were considered for designing and synthesizing newer sets of

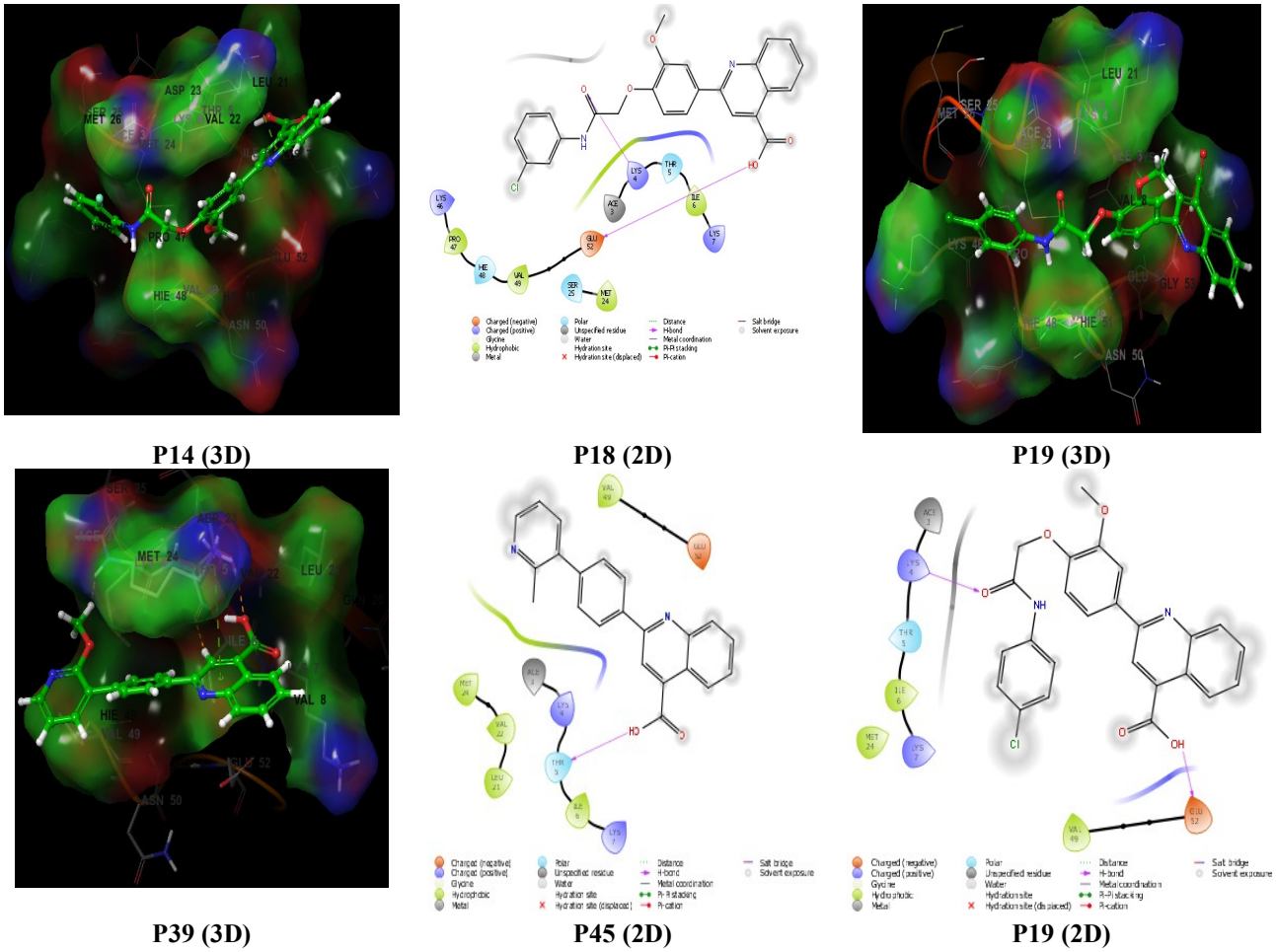



**Fig. 9.** 2D/3D binding images of the best selected docked ligands against PDB: 3MJ5.

molecules acting on both targets. These new sets of molecules may emerge as dual inhibitors of DHODH and SARS-Cov-2.

## DISCUSSION

The blue occlusions around the substituents at the $C_4$ ($R_1$-group), $C_7$, and $C_8$ positions of the quinoline ring and the aryl group (all positions except $C_2$) attached to the $R_4$-position indicate that electron-withdrawing functionalities may favor the activities while the H-bond donor group at the –OH group of carboxylic acid group, $C_8$-position of quinoline ring, and $C_2$-position of aryl ring attached to the $R_4$-position may enhance the activity. Moreover, the hydrophobic groups attached to the quinoline ring at N, $C_5$, and $C_6$-positions and the alkyl/aryl group attached to the $R_4$-position may also favor the activity. On the other hand, the electron-withdrawing groups at the hydrogen of the carboxylic group, $C_5$, and $C_6$-positions of the quinoline rings, and $C_2$-position of the aryl ring attached to the $R_4$-position, H-bond donor group at the carboxylic group and $C_7$ position of the quinoline ring and $C_3$-position of the aryl ring attached to the $R_4$-position, and hydrophobic substitutions at other positions of the quinoline ring were marked with the red color in the maps, indicating that further electron-withdrawing substituents may disfavor the activity.

In the field-based 3D-QSAR model, a sterically hindered group on the $R_4$-position was attached to the aryl rings, suggesting that the hydrophobic effect may marginally affect the activity at the maximum position and enhance it. However, H-bond donor and acceptor groups attached to the side chain at $R_2$-groups ($C_3$-position), CO-position at $R_1$, and aryl substitutions at $R_4$-position changed the value of activity (*i.e*., bioactivity), which depends on the nature of H-donor and acceptor atoms. On the other hand, the steric hindrance of other groups did not affect the DHODH inhibition activity. Only the balance of the other electronic substituent properties is required to be maintained. The electrostatic effect at the $R_2$, $R_1$, $R_5$, and 4-position of the aryl ring at $R_4$ showed a negative contribution while the C3 and C5 positions of the aryl ring at the $R_4$-position, $R_1$, and $R_3$ atoms showed a positive contribution to the bioactivity. KPLS-based analysis suggested the consideration of the dendritic fingerprint descriptor for the best performance.

## CONCLUSIONS

In summary, in the present, statistically robust and interpretable 2D- and 3D-QSAR models were developed based on a dataset of 92 molecules of biologically active 2-aryl-4-quinoline carboxylic acid analogs, reported as DHODH inhibitors in the literature. Based on the QSAR analysis, it was found that the generated 4-point pharmacophore model (ARRR_1), with one H-bond acceptor and three aromatic rings, was crucial for retaining the activities of 2-aryl-4-quinoline carboxylic acid analogs as DHODH inhibitors.

Thus, the developed QSAR models are hoped to help researchers to better design 2-aryl-4-quinoline carboxylic acid analogs as DHODH inhibitors. Furthermore, based on the correlation between DHODH inhibitors and some COVID-19 targets, we suggest that the dataset hits P19, P14, P39, P18, and P45 can be repurposed and be used as plausible SARS-CoV-2 inhibitors.

## ACKNOLWEDGMENTS

The authors would like to thank the Schrodinger team, Bangalore, India for providing the technical and software support to carry out this study (a trial license provided to SM).

## REFERENCES

[1] Palfey, B. A.; Björnberg, O.; Jensen, K. F., Insight into the chemistry of flavin reduction and oxidation in Escherichia coli dihydroorotate dehydrogenase obtained by rapid reaction studies, *Biochemistry*. **2001**, *40*, 4381-90. DOI: 10.1021/bi0025666.

[2] Björnberg, O.; Grüner, A. C.; Roepstorff, P.; Jensen, K. F., The activity of Escherichia coli dihydroorotate dehydrogenase is dependent on a conserved loop identified by sequence homology, mutagenesis, and limited proteolysis, *Biochemistry*. **1999**, *38*, 2899-908. DOI: 10.1021/bi982352c.

[3] Sorensen, P. G.; Dandanell, G., A new type of dihydroorotate dehydrogenase, type 1S, from the thermoacidophilic archaeon Sulfolobussolfataricus. *Extremophiles*. **2002**, *6*, 245-51. Ph.D. thesis.

University of Copenhagen, Copenhagen, Denmark. DOI: 10.1007/s00792-001-0249-0.

[4] Löffler, M.; Fairbanks, L. D.; Zameitat, E.; Marinaki, AM.; Simmonds, HA, Pyrimidine pathways in health and disease, *Trends Mol. Med*. **2005**, *11*, 430-7. DOI: 10.1016/j.molmed.2005.07.003.

[5] Lane, A. N.; Fan, T. W. M., Regulation of mammalian nucleotide metabolism and biosynthesis, *Nucleic Acids Res*. **2015**, *43*, 2466-85. DOI: 10.1093/nar/gkv047.

[6] Das, P.; Deng, X.; Zhang, L.; Roth, M. G.; Fontoura, B. M.; Phillips, M. A.; De Brabander, J. K., SAR based optimization of a 4-quinoline carboxylic acid analog with potent anti-viral activity, *ACS Med. Chem. Lett*. **2013**, *4*, 517-521. DOI: 10.1021/ml300464h.

[7] Petrović, M. M.; Roschger, C.; Chaudary, S.; Zierer, A.; Mladenović, M.; Marković, V., Low cytotoxic quinoline-4-carboxylic acids derived from vanillin precursors as potential human dihydroorotate dehydrogenase inhibitors, *Bioorg. Med. Chem. Lett.,* **2021**, *46*, 128194. DOI: 10.1016/j.bmcl.2021.128194

[8] Ekins, S.; Mestres, J.; Testa, B., *In silico* pharmacology for drug discovery: methods for virtual ligand screening and profiling, *Br. J. Pharmacol.* **2007**, *152*, 9-20, DOI: 10.1038/sj.bjp.0707306.

[9] Montgomery, D. C.; Peck, E. A.; Vining, G. G., Introduction to Linear Regression Analysis: John Wiley & Sons, **2021**.

[10] Topliss, J. G.; Edwards, R. P., Chance factors in studies of quantitative structure-activity relationships, *J. Med. Chem.* **1979**, *22*, 1238-1244, DOI: 10.1021/ jm00196a017.

[11] Wold, S., PLS for Multivariate Linear Modeling. QSAR: Chemometric Methods in Molecular Design Methods and Principles in Medicinal Chemistry). Weinheim, Germany: Verlag-Chemie, **1994**.

[12] Eriksson, L.; Verhaar, H. J.; Hermens, J. L., Multivariate characterization and modeling of the chemical reactivity of epoxides, *Environ. Toxicol. Chem.* **1994**, *13*, 683-691, DOI: 10.1002/etc.5620130502.

[13] Sandberg, M.; Eriksson, L.; Jonsson, J.; Sjöström, M.; Wold, S., New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids, *J. Med. Chem.* **1998**, *41*, 2481-2491, DOI: 10.1021/jm9700575.

[14] Mali, S. N.; Pandey, A., Balanced QSAR and molecular modeling to identify structural requirements of imidazopyridine analogues as anti-infective agents against trypanosomiases, *J. Comput. Biophys. Chem.* **2022**, *21*, 83-114. DOI: 10.1142/S2737416521410015.

[15] Thorat, B. R.; Mali, S. N.; Rani, D.; Yamgar, R. S., Synthesis, In silico and In vitro Analysis of Hydrazones as Potential Antituberculosis Agents, *Curr. Computer-Aided Drug Des.* **2021**, *17* (2), 294-306. DOI: https://doi.org/10.2174/1573409916666200302120942

[16] Kapale, S. S.; Mali, S. N.; Chaudhari, H. K., Molecular modelling studies for 4-oxo-1, 4-dihydroquinoline-3-carboxamide derivatives as anticancer agents, *Med. Drug Disco.* **2019**, *2*, 100008. DOI: https://doi.org/10.1016/j.medidd.2019.100008.

[17] Mali, S. N.; Pandey, A.; Bhandare, R. R.; Shaik, A. B, Identification of hydantoin based Decaprenylphosphoryl-β-d-Ribose Oxidase (DprE1) inhibitors as antimycobacterial agents using computational tools, *Sci. Rep.* **2022**, *12* (1), 1-21. DOI: https://doi.org/10.1038/s41598-022-20325-1

[18] Neto, R. D. A.; Santos, C. B.; Henriques, S. V.; Machado, L. D. O.; Cruz, J. N., da Silv; C. H. D. P., Federico, L. B.; Oliveira, E. H. D.; de Souza, M. P.; da Silva, P. N.; Taft, C. A., Novel chalcones derivatives with potential antineoplastic activity investigated by docking and molecular dynamics simulations, *J. Biomol. Struct. Dyn.*, **2022***, 40*(5), 2204-2216.https://doi.org/10.1080/07391102.2020.1839562.

[19] da Rocha Galucio, N. C.; de Araújo Moysés, D.; Pina, J. R. S.; Marinho, P. S. B.; Júnior, P. C. G.; Cruz, J. N.; Vale, V. V.; Khayat, A. S.; do Rosario Marinho, A. M., Antiproliferative, genotoxic activities and quantification of extracts and cucurbitacin B obtained from Luffa operculata (L.) Cogn. *Arab. J. Chem*, **2022**, *15* (2), 103589.https://doi.org/10.1016/j.arabjc.2021.103589.

[20] Rego, C. M. A.; Francisco, A. F.; Boeno, C. N., Inflammasome NLRP3 activation induced by Convulxin, a C-type lectin-like isolated from *Crotalus durissusterrificus* snake venom. *Sci. Rep.* **2022**, 12, 4706. https://doi.org/10.1038/s41598-022-08735-7.

[21] Almeida, V. M.; Dias, E. R.; Souza, B. C.; Cruz, J. N.; Santos, C. B.; Leite, F. H.; Queiroz, R. F.; Branco, A.,

Methoxylated flavonols from VelloziadasypusSeub ethyl acetate active myeloperoxidase extract: *in vitro* and *in silico* assays, *J. Biomol. Struct. Dyn.*, **2022**, *40* (16), 7574-7583.https://doi.org/10.1080/07391102.2021.1900916

[22] Xu, Y.; Jiang, H., Potential treatment of COVID-19 by inhibitors of human dihydroorotate dehydrogenase. *Protein & Cell*, **2020**, *11* (10), 699-702.