Biointerface Research in Applied Chemistry
Open-Access Journal (ISSN: 2069-5837)

Article
Volume 13, Issue 4, 2023, 370
https://doi.org/10.33263/BRIAC134.370

2D-QSAR Modeling of Chalcone Analogues as Angiotensin Converting Enzyme Inhibitor

Sapan Shah [1,*], Dinesh Chaple [1], Puja Badne [1], Samruddhi Khonde [1], Shivani Deshmukh [1], Sumit Arora [2], Subhash Yende [3]

[1] Department of Pharmaceutical Chemistry, Priyadarshini J. L. College of Pharmacy, Hingna Road, Nagpur-440016, Maharashtra, India; shah.sapan@rediffmail.com; (S.S.); d.chaple@rediffmail.com (D.C.); pujabadne5@gmail.com (P.B.); samruddhikhonde97@gmail.com (S.K.); deshmukhshivani38@gmail.com (S.D.);
[2] Department of Pharmacognosy and Phytochemistry, Gurunanak College of Pharmacy, Nari, Nagpur 440026, Maharashtra, India; sumitkishanarora@gmail.com (S.A.);
[3] Department of Pharmacology, Gurunanak College of Pharmacy, Nari, Nagpur 440026, Maharashtra, India; subhashyende@gmail.com (S.Y.);
* Correspondence: shah.sapan@rediffmail.com (S.S.);

Scopus Author ID 55574193515
Received: 27.05.2022; Accepted: 27.06.2022; Published: 7.10.2022

Abstract: Targeting angiotensin-converting enzyme (ACE) comes out to be an effective mechanism for controlling hypertension. Two-dimensional quantitative structural activity relationship models were generated to predict the ACE inhibitory activity of chalcone analogs. The genetic algorithm- multiple linear regression models (GA-MLR) approach was used to generate highly predictive models using straightforwardly interpretable Py, Estate, Alvadesc, and Padel descriptors. Application of Intelligent consensus modeling confirms that model-2 is statistically robust (R^2_{tr} = 0.66, Q^2_{LOO} = 0.5621) with good external predictivity (Concordance Correlation Coefficient, CCC_{ex} = 0.9109, Q^2-F^1 = 0.85818, Q^2-F^2 = 0.85782 and Q^2-F^3 = 0.88489). Novel analogs designed according to the synthetic route considering structural requirements indicated by the model were found to be satisfactory and could be considered for synthesis and subsequent screening.

Keywords: angiotensin-converting enzyme; chalcone; 2D-QSAR; consensus modeling; validation.

1. Introduction

Hypertension is a widespread illness in adults all over the world, and it is one of the leading causes of cardiovascular and renal disorders [1,2] Angiotensin-converting enzyme (ACE) inhibition has come out to be an effective mechanism in controlling hypertension [3]. Angiotensin-converting enzyme inhibitors (ACEIs) regulate the biosynthesis of a vital chemical known as angiotensin II thereby decreasing salts concentration, dilating arteries, and controlling hypertension. ACE and Angiotensin II are considered essential points of regulation in the Renin-Angiotensin-Aldosterone System (RAAS); deregulation of any of them is accountable for cardiovascular and renal diseases [4–7]. The RAAS is widely known for its importance in cardiovascular physiology, water-electrolyte balance, and cell function. Excessive activation of this system is a major contributor to hypertension. ACE is the essential regulation locus of RAAS to combat hypertension. ACE is a dipeptidyl carboxypeptidase that removes the carboxy-terminal dipeptide of angiotensin I, inactivating the vasodepressor bradykinin and activating the strong vasoconstrictor. Growth-promoting chemical angiotensin

II [8,9] There are many developments in which ACE-related exploration requisite like deficiencies attributed to the current generation of these drugs, overcoming side effects such as dry cough or angioedema, and the development of irreversible ACE inhibitors (ACEIs) [10–12].

Computer-aided drug design (CADD) is one of the most important ways to preclinical drug discovery today. It employs a variety of computational methodologies and software packages to reach the desired result [13–17]. QSAR is the final result of computer processes that begin with a sufficient description of molecular structure and conclude with inferences, hypotheses, and predictions about the behavior of molecules in the environment, physicochemical system, and biological system under investigation [18,19]. Moreover, QSAR is an accurate method that can be applied to predict the half-maximal inhibitory concentration(IC_{50}) value of novel designed candidates [20–22]. Many computational studies have been performed [23–28] so far for designing new inhibitors against ACE. However, we are still far from finding a potentially small non-peptide moiety as ACEIs.

Chalcone derivatives belong to a potent category and display a broad spectrum of pharmacological effects [29–33]. Also, the enormous number of computational studies performed on the activities like cytotoxic activity [34–36], antitubercular activity[37], antiviral activity [38], and antimalarial activity [39], etc. of chalcone analogs. In addition to these, synthetic and natural Chalcone analogs have shown significant potential to inhibit ACE [40–43]. However, no attempts have been made to develop a QSAR model for predicting the ACE inhibitory activity of chalcone analogs. Thus, in the present research work, a dataset of 45 chalcone analogs evaluated as ACEIs, was used to develop a statistically robust 2D-QSAR model using the GA-MLR approach with a high external predictive ability. New compounds were designed according to the synthetic route considering structural requirements, and the 2D-QSAR model investigated their inhibitory activity against ACE. Such a model could be useful to explore the binding mode of inhibitor with ACE to gain further insight into the structure-activity relationship.

2. Materials and Methods

2.1. Datasets selection.

A dataset of 45 substituted chalcones analogs tested for ACE inhibitory activity was selected from the ChEMBL database for the present work [40,41,44,45] The following criteria were adopted to refine the data sets to get homogenous records for reliable prediction: (1) Compound analysis on *Oryctolagus cuniculus* organism was only composed; (2) data with inhibitory assay executed by an *in vitro* experiment on single protein target Angiotensin-converting enzyme; (3) biological activity of each compound is expressed in IC_{50} (nM) values. Before the QSAR model building, the reported half-maximum inhibitory concentration, IC50 (nM) values for ACEIs were converted to molar units and then pIC_{50} ($-\log_{10}IC_{50}$). The dataset was put through chemical curation by using the Konstanz Information Miner (KNIME) workflow (https://www.knime.org/) and 3D optimized using MMFF94 force-field by OpenBabel [46–48].

2.2. Descriptor calculation and pre-treatment.

To get paramount knowledge about characteristics of molecular structures, we have calculated a variety of descriptors, namely, Py-Descriptor (Constitutional, geometric, circular

fingerprint, quantum chemical and topological, no. of descriptors = 16,359) [49], Estate Descriptors (for calculating of atom and bond types properties such as n-octanol/water partition coefficient, solubility, etc.) (no. of descriptors = 25) 50], Alvadesc-Descriptors (Constitutional, topological, connectivity indices, 2D matrix-based, ETA indices, Atom type E-state, Functional group count, 2D atom pair, Ring, Atom-centered fragments, Molecular properties and Drug like indices, no. of descriptors = 3,126) and PaDEL-Descriptor 2.20 (for extended topochemical atom indices, no. of descriptors = 99) [51].

Additional data pre-treatment for descriptors (Total no. of descriptors = 19,609) with constant, semi-constant values (>95%) and highly correlated (|R| > 95%) ones was done before particular feature selection using QSARINS-Chem 2.2.4 to reduce the repeated and inefficient information in the data [52]. Finally, a set of only 894 descriptors (Covering extensively different structural features, viz. constitutional, geometric, autocorrelation, topological, quantum chemical, and circular fingerprints) were used as input for developing the QSAR model for ACE inhibitory activity (see Table S1).

2.3. QSAR model development.

2.3.1. QSAR model.

One of the principal applications of QSAR modeling is to explore imperative structural characteristics to predict pIC_{50} of unknown chemical data before actual experimental evaluation. The genetic algorithm (GA) module of QSARINS-Chem 2.2.4 was applied to select an optimum number of sets of descriptors. Multiple QSAR models were developed using a divided dataset (75% training and 25% prediction set) by structure-based splitting (based on structural similarity computed from the principal component from the pool of input descriptors) to ensure the inclusion of all the relevant activity and structural knowledge. The strategy used in QSAR model development is given in Figure 1.

2.3.2. Statistical validation of the generated 2D-QSAR models.

Organization of Economic Cooperation and Development (OECD) guidelines were followed while performing internal and external validation of QSAR models to evaluate statistical robustness [52,53]. The developed QSAR models are validated according to fitting criteria (regression coefficient: R^2), internal (cross-validated correlation coefficient leave-one-out: Q^2_{LOO}, leave many out: Q^2_{LMO}), and external (Q^2-F^1, Q^2-F^2, Q^2-F^3, R^2_m, and CCC) parameters to assure the significant level of the generated models [54]. In addition, prediction accuracy root means square of errors ($RMSE_{CV}$ and $RMSE_{Ex}$), and mean absolute error (MAE_{CV} and MAE_{Ex}) were also calculated. Additionally, Y-scrambling [2000 iteration] [55] and applicability domain (AD) criteria were checked using the Williams plot—the plot of HAT i/i(h*) versus standardized cross-validated residuals [56]. Furthermore, to endorse consistency or to get reliable predictions "intelligent consensus modeling" technique based on several parameters was used [57].

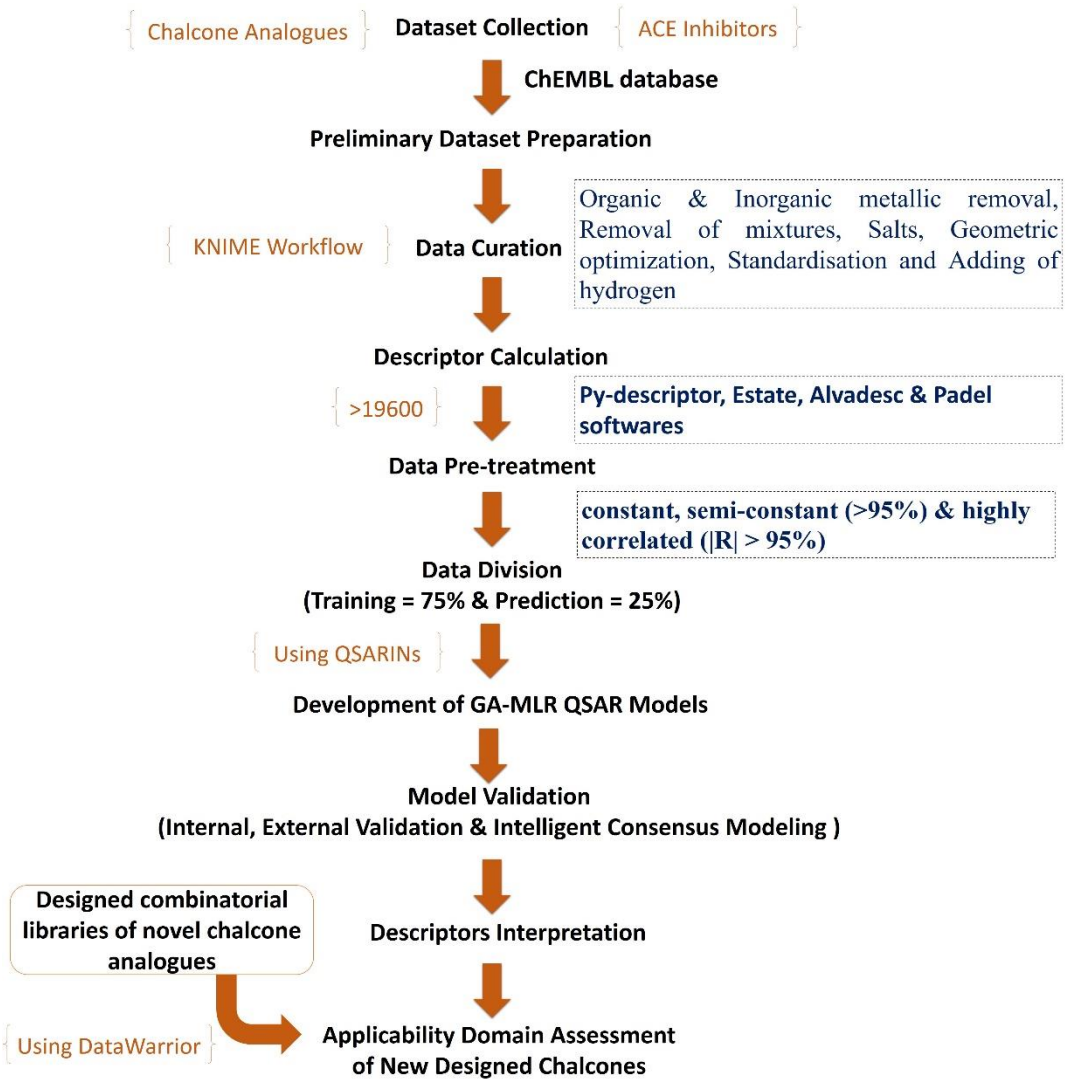


Figure 1. Stages intricate in the development of QSAR model and analysis of designed compounds.

2.4. Modeling of new compounds.

About 457 novel chalcones analog were designed using DataWarrior software according to the synthetic route considering structural requirements indicated by the developed QSAR model [58]. The chemical similarity analysis with the compounds database in ChEMBL reveals that designed analogs are novel and have not been experimented with before for ACE inhibitors. All compounds generated were 3D optimized using MMFF94 force-field by OpenBabel, and descriptors are calculated as discussed above. The best acceptable and predictable QSAR model was used to predict the applicability domain of the optimized compounds using QSARINs, and the activity (pIC_{50}) was predicted for the generated compounds exhibited HAT i/i (h*=0.4412) or less.

2.5. ADMET study of designed compounds.

The selected 113 designed chalcone analogs were subjected to evaluate drug-likeness properties according to Lipinski, Ghose, Veber, Egan, and Muegge filters. The pharmacokinetic profile (ADME) and toxicity predictions of ligands were conducted using SwissADME (http://www.swissadme.ch) and pkCSM (http://biosig.unimelb.edu.au/pkcsm/prediction). SMILES notations or SDF files are uploaded

to analyze the toxicological properties of ligands, followed by selection of required models for generating numerous information about structure-related effects [59,60].

3. Results and Discussion

3.1. QSAR models.

In the present research work, QSAR studies were performed to reveal concealed structural features responsible for ACE inhibition. To identify important structural features of chalcone analogs, multiple QSAR models were generated by the GA-MLR method [61]. Further, in developing the QSAR model, over-fitting is restricted by selecting an optimum number of molecular descriptors [62]. The plot between a variety of molecular descriptors and R^2_{tr} and Q^2_{LOO} values suggests that the presence of four variables (descriptors) represents the breaking point of over-fitting (Figure 2). The descriptors selected in building QSAR models are easy to access and simple to interpret, which leads to a relevant correlation with biological activity. Also, Organization for Economic Co-operation and Development (OECD) guidelines were taken into account for building GA-MLR-based QSAR models for ACE inhibition (Table 1).

Table 1. QSAR model developed for angiotensin-converting-enzyme inhibition.

Model 1 (Divided)	pIC_{50} =	4.817 + 2.0633 (± 0.6441) PW2 + 0.7776 (±0.2653) Eta_D_beta_A – 1.4265 (±0.3986) GGI5 – 3.1513 (±0.8378) C_N_1Bc
	Validation parameters (Fitting Criteria), R^2_{tr}: 0.8737, $R^2_{adj.}$: 0.8563, $R^2_{tr} – R^2_{adj}$: 0.0174, LOF: 0.1294, Kxx: 0.3884, DeltaK: 0.0165, $RMSE_{tr}$: 0.2751, MAE_{tr}: 0.1914, RSS_{tr}: 2.5736, CCC_{tr}: 0.9326, S: 0.2979, F: 50.1423	
Model 2 (Divided)	pIC_{50} =	5.9719 – 6.8898(± 0.4617) Eta_sh_p – 9.1831 (±1.1239) C_N_1Bc – 2.2493 (±0.2923) C_acc_1Bc + 2.4588 (±0.5874) N_C_9Ac
	Validation parameters (Fitting Criteria), R^2_{tr}: 0.6600, $R^2_{adj.}$: 0.6131, $R^2_{tr} – R^2_{adj}$: 0.0469 , LOF: 0.3599, Kxx: 0.4521, DeltaK: 0.0016, $RMSE_{tr}$: 0.4587, MAE_{tr}: 0.2978, RSS_{tr}: 7.1552, CCC_{tr}: 0.7952, S: 0.4967, F: 14.0757	
Model 3 (Divided)	pIC_{50} =	5.9832 + 32.2368 (± 0.3984) JGI2 – 106.1411 (±0.275) JGI5 – 0.0804 (±0.6383) sp3C_all_5A
	Validation parameters (Fitting Criteria), R^2_{tr}: 0.7105, $R^2_{adj.}$: 0.6805, $R^2_{tr} – R^2_{adj}$: 0.0300, LOF: 0.2744, Kxx: 0.0328, DeltaK: 0.3097, $RMSE_{tr}$: 0.4286, MAE_{tr}: 0.3642, RSS_{tr}: 6.0612, CCC_{tr}: 0.8307, S: 0.4572, F: 23.7200	
Model 4 (Divided)	pIC_{50} =	4.9584 – 64.2176 (±0.2245) Eta_B_A + 38.1096 (±0.4709) JGI2 – 4.7323 (±0.706) C_acc_2Bc + 2.3446 (±0.5608) C_ringC_2Bc – 0.0895 (±0.7104) sp3C_all_5A
	Validation parameters (Fitting Criteria), R^2_{tr}: 0.8324, $R^2_{adj.}$: 0.8014, $R^2_{tr} – R^2_{adj}$: 0.0310, LOF: 0.2188, Kxx: 0.3727, DeltaK: 0.0836, $RMSE_{tr}$: 0.3260, MAE_{tr}: 0.2521, RSS_{tr}: 3.5076, CCC_{tr}: 0.9086, S: 0.3604, F: 26.8280	
Model 5 (Divided)	pIC_{50} =	5.1865 + 0.7324 (±0.2259) D/Dtr06 + 2.0598 (±0.643) PW2 – 1.9848 (±0.5546) GGI5 – 3.3231 (±0.8835) C_N_1Bc
	Validation parameters (Fitting Criteria), R^2_{tr}: 0.8775, $R^2_{adj.}$: 0.8607, $R^2_{tr} – R^2_{adj}$: 0.0169, LOF: 0.1255, Kxx: 0.3108, DeltaK: 0.0431, $RMSE_{tr}$: 0.2709, MAE_{tr}: 0.2049, RSS_{tr}: 2.4948, CCC_{tr}: 0.9348, S: 0.2933, F: 51.9557	

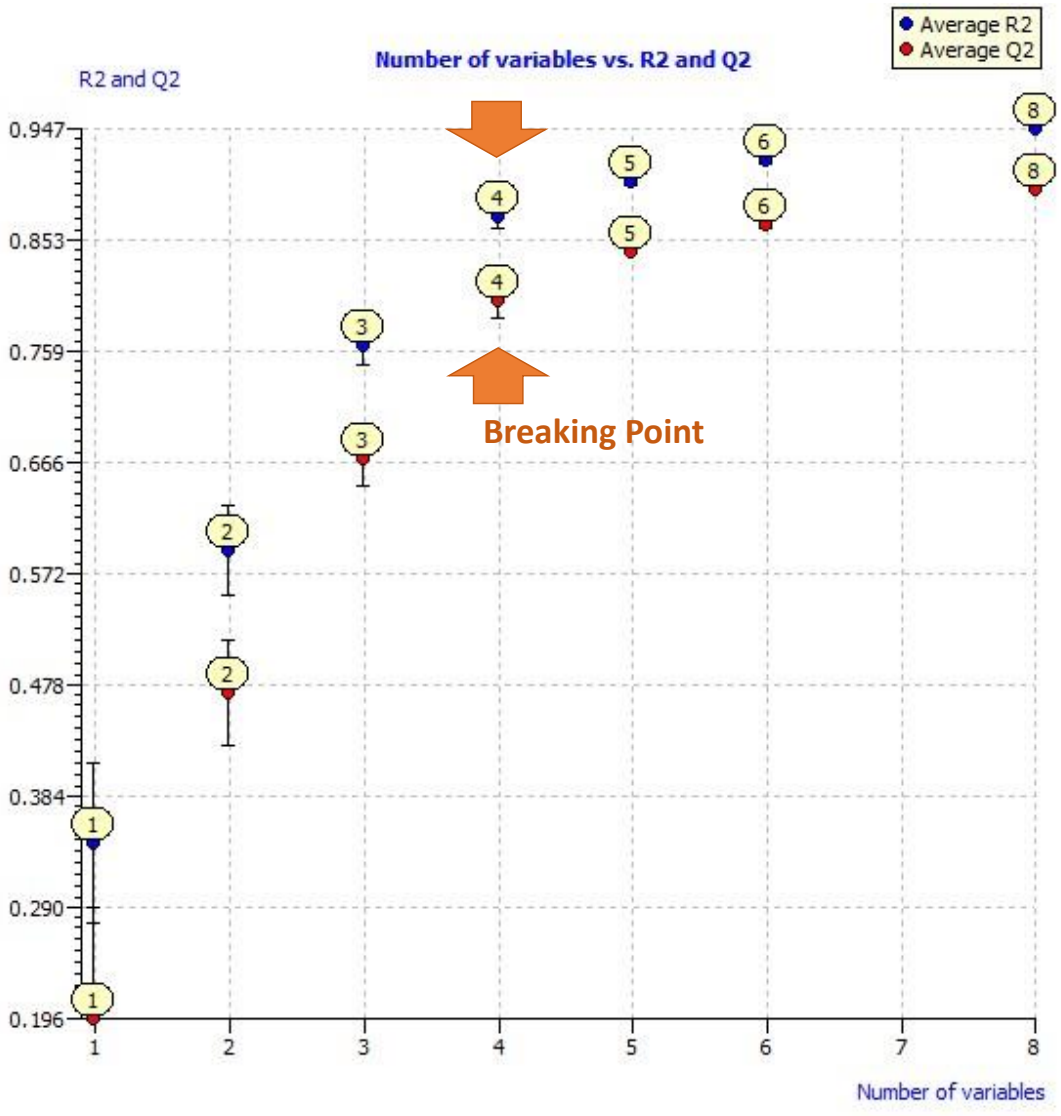


Figure 2. A plot between selected descriptors and R^2_{tr} and Q^2_{LOO} values suggests optimum descriptors requisite in the model.

3.2. Validation of QSAR models.

The statistical quality of the QSAR models was validated according to OECD guidelines taking into consideration of fitting criteria (R^2_{tr}, $RMSE_{tr}$, CCC_{tr}, and F values), Internal validation (Q^2_{LOO}, $RMSE_{cv}$, CCC_{cv}, Q^2_{LMO}), and External validation ($RMSE_{ex}$, Q^2-F^1, Q^2-F^2, Q^2-F^3, r^2m average) parameters as models will be applied in prediction of the biological activity of novel lead molecules or optimization through drug discovery.

Primarily, fitting criteria results suggest that model show linear relationship ($R^2 > 0.6$), no over-fitting ($R^2_{tr} - R^2_{adj} < 0.05$), Lack of fit (LOF < 0.5) and absence of inter-correlation among descriptors (k < 0.5) [58,59]. Further, Internal validation parameters indicate that developed QSAR models are statistically robust based on cross-validated (Q^2_{LOO}, Q^2_{LMO}, and $CCC_{CV} > 0.5$), predictive ($R^2 - Q^2_{LOO} < 0.3$), and not developed by chance correlation (low value of R^2_{Yscr} and Q^2_{Yscr}) [65,66]. Figure 3a is a plot of Q^2_{LMO} versus the correlation between descriptors and angiotensin-converting enzyme inhibition (Kxy) in which it is observed that the values of Q^2_{LMO} are very similar to each other, which confirms that the model has a good fit, robustness, and stability [67,68].

External prediction capacity of the models are indicated by the high value of external determination coefficient (R^2_{ex} > 0.5), variances of external prediction (Q^2-F^1, Q^2-F^2, Q^2-F^3 > 0.5), Concordance Correlation Coefficient (CCC_{ex} > 0.6) and squared correlation coefficient without intercept (r^2m aver. > 0.5)[67]. The fact is reinforced by the lower value of $RMSE_{ex}$, MAE_{ex}, and $PRESS_{ext}$. Also, the Williams plot of leverage (h) against standardized residuals by model and LOO for the selected dataset of ACEIs were created, which ascertained that compounds lie within the applicability domain [56]. Two compounds are considered as potential outliers- a) compound 25, with a value greater than the critical leverage (h = 0.441) but showing residual within limits, and b) compound 3, having residual off-limits (± 3.0) though this is within critical leverage (Figure 3b).

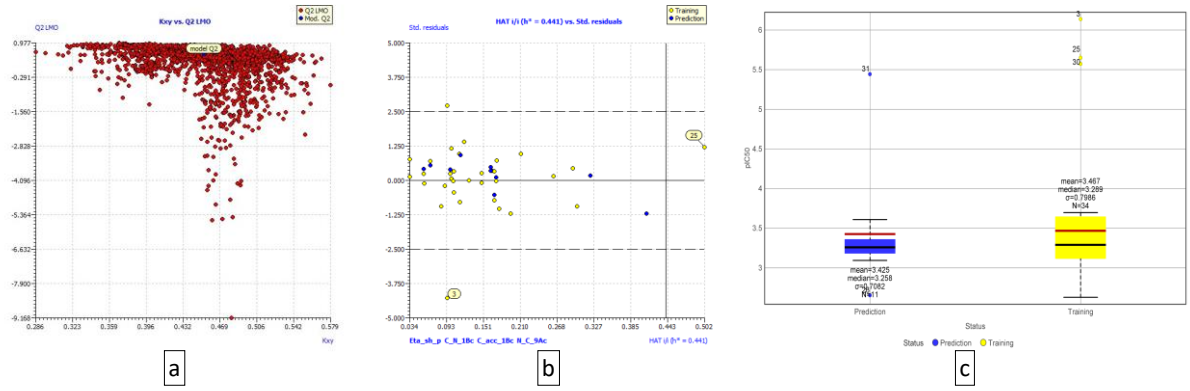


Figure 3. Validation graphs of developed QSAR a) plot of Q^2_{LMO} versus the correlation between descriptors and angiotensin-converting enzyme inhibition (Kxy) b) Williams plot of leverage (h) against standardized residuals by LOO c) box plot of a divided dataset in model-2 suggesting proper distribution.

Table 2. Statistical performance of developed QSAR models based on internal validation criteria (using QSARINs).

Sr. No.	Statistical Parameter	Model 1	Model 2	Model 3	Model 4	Model 5
Internal Validation Criteria						
	R^2_{cv} (Q^2_{LOO})	0.8207	0.5621	0.6190	0.7561	0.8245
	$R^2 - Q^2_{LOO}$	0.0529	0.0979	0.0915	0.0764	0.0530
	$RMSE_{cv}$	0.3277	0.5206	0.4916	0.3934	0.3243
	MAE_{cv}	0.2279	0.3483	0.4169	0.3072	0.2474
	$PRESS_{cv}$	3.6520	9.2165	7.9757	5.1061	3.5749
	CCC_{cv}	0.9020	0.7418	0.7710	0.8658	0.9053
	Q^2_{LMO}	0.7616	0.4661	0.5655	0.6979	0.7671
	R^2_{Yscr}	0.1221	0.1205	0.0937	0.1521	0.1219
	Q^2_{Yscr}	-0.2520	-0.2261	-0.1895	-0.2981	-0.2347
	RMSE AV $_{Yscr}$	0.7242	0.7371	0.7576	0.7323	0.7243

Likewise, the significantly low values of R^2 and Q^2 of the developed models from the values obtained for those parameters in the Y-scrambling at 2000 iterations experiment confirms the reliability and further that models not obtained by chance correlation. Moreover, after evaluating Golbraikh and Tropsha's criteria [69,70], we can observe that model 2 is statistically better than the rest of the models (Tables 2 and 3). The nearly same median value indicates an absence of any statistical leakage in the divided dataset training (median=3.289) and prediction (median=3.258) datasets with appropriate composition (Figure 3c). To advance the analysis of predictive efficiency of individual QSAR models, intelligent consensus modeling was previously utilized in our research article [47]. The results suggested that the individual Model-2 comes out to be the winning model, which strengthens overall prediction accuracy (Table 4). The Dixon Q-test, AD criteria, and Euclidean distance cutoff parameters

were used to check the prediction of each compound. Thus, the designed compounds' prediction quality might be more precise and acceptable by applying the Dixon-Q test.

Table 3. Statistical performance of developed QSAR models based on external validation criteria (using QSARINs).

Sr. No.	Statistical Parameter	Model 1	Model 2	Model 3	Model 4	Model 5
External Validation Criteria						
	θ*	-5.6539°	-8.9606°	-8.2600°	-7.8200°	-2.8373°
	$RMSE_{ex}$	0.4480	0.2551	0.4193	0.4531	0.4306
	MAE_{ex}	0.3310	0.2247	0.3523	0.3368	0.3342
	$PRESS_{ext}$	2.0074	0.7161	1.9335	2.2579	1.8538
	R^2_{ex}	0.7285	0.9188	0.7535	0.7398	0.7751
	Q^2-F^1	0.6415	0.8578	0.6168	0.5525	0.6689
	Q^2-F^2	0.6415	0.8572	0.6145	0.5498	0.6689
	Q^2-F^3	0.6650	0.8948	0.7229	0.6764	0.6906
	CCC_{ex}	0.8206	0.9107	0.7948	0.7712	0.8454
	r^2_m aver.	0.6241	0.6393	0.5880	0.5876	0.6815
	r^2_m delta	0.0297	0.1572	0.2313	0.2344	0.1402
External predictions by model equation						
	R^2	0.7285	0.9188	0.7535	0.7398	0.7751
	R'^2o:	0.7134	0.7669	0.6143	0.6072	0.7742
	k'	0.9353	1.0178	1.0612	1.0736	0.9360
	$1-(R^2/R'^2_o)$	0.0208	0.1653	0.1847	0.1793	0.0012
	r'^2_m	0.6389	0.5607	0.4724	0.4704	0.7516
	R^2_o	0.7017	0.8710	0.7491	0.7376	0.7305
	K	1.0550	0.9778	0.9335	0.9223	1.0556
	$1-(R^2-ExPy/R_o^2)$	0.0368	0.0520	0.0058	0.0030	0.0575
	r^2_m	0.6092	0.7179	0.7037	0.7048	0.6115

The comparative standing of the variables in the GA-MLR based developed QSAR model is in the following order: C_N_1Bc > Eta_sh_p > N_C_9Ac > C_acc_1Bc. Not only in model 2 but also in models 1 and 5, the descriptor 'C_N_1Bc' contributes utmost towards the ACE inhibitory activity that relates from its highest negative correlation coefficient with pIC_{50} (ACE inhibitory activity) [71]. The brief explanation of descriptors that are present in the winner model 2 is summarized in Table 5, and definitions of all statistical parameters are given in Table S2.

Table 4. Statistical performance of the developed individual and consensus QSAR models (using intelligent consensus predictor).

Statistical Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	CM0	CM1	CM2	CM3
NCompExt	10	10	10	10	10	10	10	10	10
Q^2-F^1(100%)	0.6415	0.8582	0.6208	0.5694	0.6689	-	-	-	-
Q^2-F^2 (100%)	0.6415	0.8578	0.6192	0.5676	0.6689	0.2139	0.6678	0.6603	0.5617
Q^2-F^3 (100%)	0.665	0.8849	0.6992	0.6584	0.6906	-	-	-	-
CCC (100%)	0.8206	0.911	0.7955	0.7782	0.8454	-	-	-	-
r^2_m aver. (100%)	0.6066	0.6233	0.4697	0.4776	0.6032	0.0910	0.4652	0.4712	0.4273
Delta r^2_m (100%)	0.1627	0.1367	0.2492	0.2434	0.1525	0.2424	0.2473	0.2427	0.2484
MAE (100%)	0.3310	0.2412	0.3716	0.3406	0.3342	0.4777	0.3699	0.3671	0.3785
MAE (95%)	0.2669	0.2172	0.3364	0.2606	0.2696	0.3513	0.3220	0.3180	0.3187
PRESS (100%)	2.0074	0.7126	1.9084	2.1670	1.8538	4.4014	1.860	1.9019	2.4541
PRESS (95%)	1.1821	0.5028	1.4339	1.0473	1.0147	1.79	1.2201	1.2470	1.6150

Statistical Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	CM0	CM1	CM2	CM3
SDEP (100%)	0.4480	0.2669	0.4369	0.4655	0.4306	0.6634	0.4313	0.4361	0.4954
SDEP (95%)	0.3624	0.2364	0.3992	0.3411	0.3358	0.446	0.3682	0.3722	0.4236
f1	0	0	0	0.1	0	0.2	0	0	0
f2(TrainMean)	0.1	0.1	0.1	0.1	0.1	0.1			
f2(TestMean)	0.1	0.1	0.1	0.1	0.1	0.1			
ED cutoff	0.3								
AD criteria	YES								
Dixon-Q Test	YES								

3.3 Descriptors' interpretations.

The winner model-2 equation is mainly governed by the quantum chemical and extended topochemical atom indices (ETA) descriptors. A similar outcome for ACE inhibitory activity was also found from the observations described in other developed models. Thus, highlighting the importance of these descriptors for ACE inhibitory activity. Accordingly, effective interpretation of these descriptors such as Py-descriptors (C_N_1Bc, C_acc_1Bc, and N_C_9Ac) and Alvadesc descriptor (Eta_sh_p) is requisite for inhibitory activity[72].

Table 5. Symbols and definitions for the descriptors selected in the QSAR model.

Name	Brief Description	Descriptor Type
Eta_sh_p	Eta p shape index	ETA indices
C_N_1Bc	Sum of partial charges of carbon atoms present within one bond from N atom	Quantum chemical
C_acc_1Bc	Sum of partial charges of carbon atoms present within one bond from acceptor atom	Quantum chemical
N_C_9Ac	Sum of partial charges of N atoms within 9 A.U. from C atom	Quantum chemical
PW2	path/walk 2 - Randic shape index	Topological indices
Eta_D_beta_A	eta average measure of electronic features	ETA indices
GGI5	Topological charge index of order 5	Topological Charge
JGI2	Mean topological charge index of order 2	Topological Charge
JGI5	Mean topological charge index of order 5	Topological Charge
sp3C_all_5A	Presence of sp3 hybridised C atom at a distance of 9 A.U. from all atoms	Topological
Eta_B_A	Eta average branching index	ETA indices
C_acc_2Bc	Sum of partial charges of carbon atoms present within two bond from acceptor atom	Quantum chemical
C_ringC_2Bc	Sum of partial charges of caronb atoms present withing two bonds from ring C atoms	Quantum chemical
D/Dtr06	distance/detour ring index of order 6	Ring descriptors

Eta_sh_p: The molecular descriptor Eta_sh_p is an extended topochemical atom descriptor representing the effect of branching in the cationic structure attached to three non-H-atom forming trigonal shape fragments[73]. The negative coefficient suggests that decreasing this descriptor's value could lead to better enzyme inhibition. On replacing this descriptor with other highly correlated descriptors such as D/Dtr06 (R= –0.52925), Eta_D_beta_A (R= –0.63038), GGI5 (R=0.52991), C_ringC_2Bc (R=0.57392), and JGI2 (R=0.73722) decreases model statically predictability.

C_N_1Bc: The quantum chemical descriptor C_N_1Bc points out that the decrease in partial charges on carbon atoms present within one bond from the N atom contributes to ACE inhibitory activity. It has a significant and negative correlation for ACE inhibitory activity with R = 0.638. It can be seen in the compounds 25 (pIC_{50} = 5.65) and 31 (pIC_{50} = 5.44) with the lower numerical value of this descriptor (Figure 4). In contrast, compounds like 42 (pIC_{50} = 3.34) and c43 (pIC_{50} = 3.69) having high numerical values show lower inhibitory activity.

Substituting the descriptor with other highly correlated descriptors such as JGI2 (R= –0.6545) decreases the model statically predictability.

C_acc_1Bc: Another quantum chemical descriptor with a negative coefficient is C_acc_1Bc, which stands for the sum of partial charges of carbon atoms present within one bond from the acceptor atom and stipulates the role of Carbon atoms in deciding the activity. Replacing this descriptor with other highly correlated descriptors such as C_acc_1Bc (R= 0.88579), C_ringC_2Bc (R=0.52724), and PW2 (R=–0.58081) decreases model statically predictability.

N_C_9Ac: The molecular descriptor N_C_9Ac indicates the sum of partial charges of N atoms within 9 A° from the C atom. This descriptor's positive contribution indicates that compounds' inhibitory activity is directly proportional to the numerical value of N_C_9Ac. Replacing this descriptor with other highly correlated descriptors such as C_acc_2Bc (R=0.56148) and C_ringC_2Bc (R=0.63069) decreases the model's statically predictability.

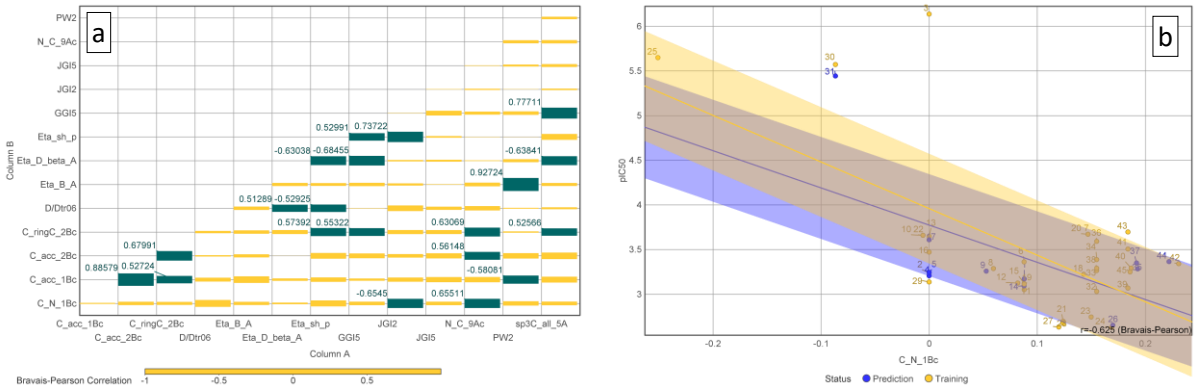


Figure 4. Graphical representation of correlation of **a)** all selected descriptors in developed QSAR models **b)** molecular descriptor C_N_1Bc with pIC50 for ACE inhibitory activity. The shaded colored portion around straight lines represents the standard deviation (±2%).

Firstly, Wang *et al*. 2011, used a novel set of descriptors G-scale (hydrophobic, stereo, and electrical characteristics) in their model [74]. Similarly, the descriptors Eta_sh_p (effect of branching in the cationic structure), C_N_1Bc, C_acc_1Bc, and N_C_9Ac represent electrical characteristics in our model-2. This descriptor indicates the importance of charges on atoms, but also, its position with respect to acceptor atoms gives an additional advantage over the published one. Secondly, Amaya *et al*. 2020, employed nominal descriptors [partition coefficient, molar refractivity, and dipole moment] in the reported model [23]. The descriptors sp3C_all_5A and logP in Model-2 (Table 1) and Amaya *et al*. model, respectively, strengthen the fact that the lipophilic carbon chain plays an important role in determining the angiotensin-converting enzyme inhibitory activity.

Consequently, the developed QSAR Model-2 can be considered to be able to predict the ACE inhibitory activity of chalcone analogs with the added advantage in comparison to the previously developed models.

3.4. AD Studies of designed compounds.

The AD screening of 457 novel chalcone analogs facilitated the identification of 216 potential leads with the potential to inhibit ACE (Figure 5). The predicted pIC_{50} values are based on rigorous validation with exhibited HAT i/i (h*=0.4412) or less. The structure and predictive activity of the recognized compounds were access through weblink

(https://sites.google.com/view/chalconeanalogues). It can be seen that about 130 compounds have better-predicted activity than the median compound (pIC50 = 3.28) and 113 compounds have better-predicted activity than the mean experimental activity (pIC50 = 3.46) of the training dataset. Thus, it is satisfying to see that the developed QSAR model revealed some novel compounds with more binding potential than the reported one.

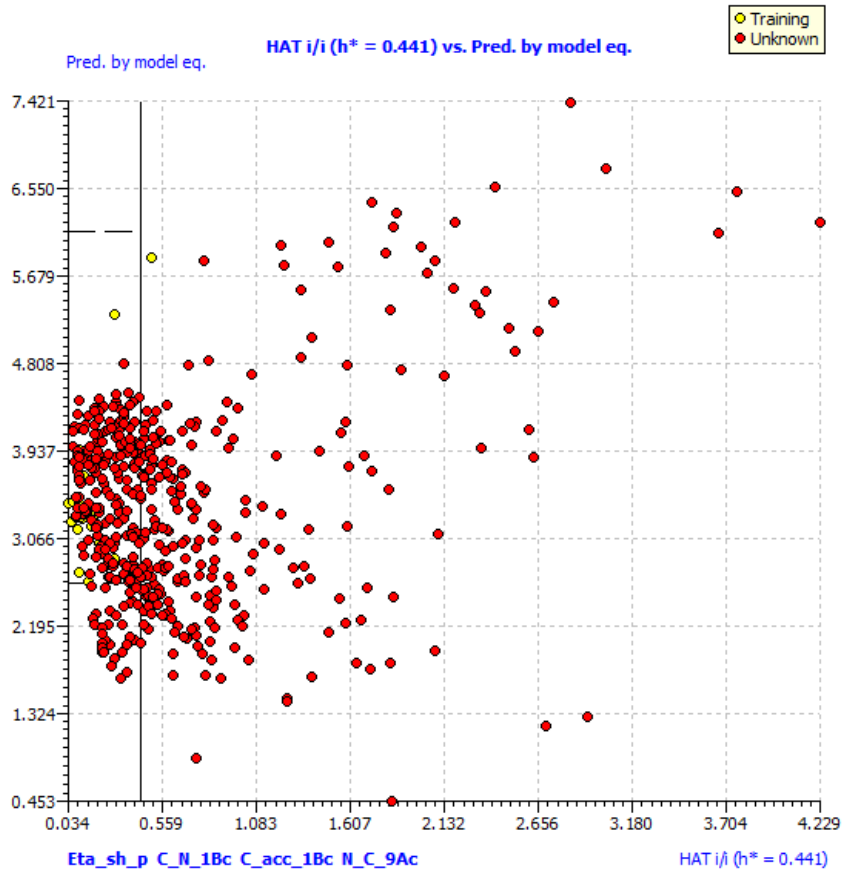
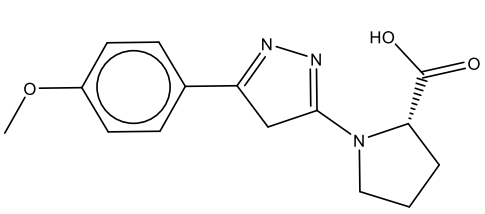


Figure 5. Applicability domain analysis of novel designed compounds using winner model-2.

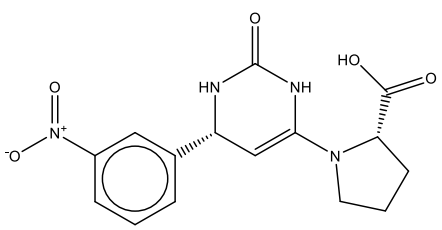
3.5 ADMET analysis of designed compounds.

In the study, 50 designed compounds with better-predicted activity than mean experimental activity follow drug-likeness parameters with no violations for Lipinski, Ghose, Veber, Egan, and Muegge filters. Additionally, all 50 designed compounds pass the pan assay interference compound (PAINS) test. Further, it is predicted that designed compounds 368 and 386 do not show AMES toxicity, hepatotoxicity, and skin sensitivity (Figure 6). Also, it does not inhibit hERG-I and II (low risk of cardiac toxicity). Moreover, T. pyriformis toxicity, Minnow toxicity, maximum tolerated dose, rat acute oral toxicity (LD50), and chronic toxicity are depicted in the Table S3 file.



(5-(4-methoxyphenyl)-4*H*-pyrazol-3-yl)-*L*-proline

(368)



((*R*)-6-(3-nitrophenyl)-2-oxo-1,2,3,6-tetrahydropyrimidin-4-yl)-*L*-proline

(386)

Figure 6. Structure of the two most promising novel-designed chalcone analogs.

4. Conclusions

The current research work describes easily interpretable, statistically robust QSAR models developed with masked structural features responsible for angiotensin-converting enzyme inhibitory activity. The developed 2D-QSAR model has a robust fitting and follows the criterion of internal validation. The external extrapolative ability of individual QSAR Model-2 is better compared to four intelligent consensus models. The results suggested that descriptors, including the Py-descriptors (quantum chemical) and Alvadesc descriptors (ETA indices), are the most significant descriptors to explain the inhibitory potential of chalcone analogs. Thus, the application of statistically robust QSAR modeling as a virtual screening tool helps discover novel compounds having the potential to inhibit ACE. In future modifications to improve the activity of chalcone analogs, emphasis on structural modifications leading to an augmented value of molecular descriptors, which have a positive coefficient in the model, must be given.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgments

The authors thank Dr. Satish Chaturvedi, Chairman, LTJSS, Nagpur, for giving moral support. We want to thank Dr. Vijay Masand, Vidya Bharati College, Amravati, for providing detailed descriptions of the Py-descriptors for interpreting models. We thank Dr. Pravin Ambure, Senior Researcher (Former MSCA-IF fellow), for his valuable support.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Abraham, H.M.A.; White, C.M.; White, W.B. The Comparative Efficacy and Safety of the Angiotensin Receptor Blockers in the Management of Hypertension and Other Cardiovascular Diseases. *Drug Saf.* **2015**, *38*, 33–54, https://doi.org/10.1007/s40264-014-0239-7.
2. Tackling, G.; Borhade, M.B. Hypertensive Heart Disease. *StatPearls* **2021**, https://www.ncbi.nlm.nih.gov/books/NBK539800/.
3. Kumar, H.; Devaraji, V.; Joshi, R.; Jadhao, M.; Ahirkar, P.; Prasath, R.; Bhavana, P.; Ghosh, S.K. Antihypertensive Activity of a Quinoline Appended Chalcone Derivative and Its Site Specific Binding Interaction with a Relevant Target Carrier Protein. *RSC Adv.* **2015**, *5*, 65496–65513, https://doi.org/10.1039/C5RA08778C.
4. Bhatti, M.S.; Asiri, Y.I.; Uddin, J.; El-Seedi, H.R.; Musharraf, S.G. Repurposing of Pharmaceutical Drugs by High-Throughput Approach for Antihypertensive Activity as Inhibitors of Angiotensin-Converting Enzyme (ACE) Using HPLC-ESI-MS/MS Method. *Arab. J. Chem.* **2021**, *14*, 103279, https://doi.org/10.1016/j.arabjc.2021.103279.
5. Ghatage, T.; Goyal, S.G.; Dhar, A.; Bhat, A. Novel Therapeutics for the Treatment of Hypertension and Its Associated Complications: Peptide- and Nonpeptide-Based Strategies. *Hypertens. Res.* **2021**, *44*, 740–755, https://doi.org/10.1038/s41440-021-00643-z.
6. Coto, E.; Avanzas, P.; Gómez, J. The Renin–Angiotensin–Aldosterone System and Coronavirus Disease 2019. **2021**, *16*, https://doi.org/10.15420/ecr.2020.30.
7. Capric, V.; Chandrakumar, H.P.; Celenza-Salvatore, J.; Makaryus, A.N. The Role of the Renin-Angiotensin-Aldosterone System in Cardiovascular Disease: Pathogenetic Insights and Clinical Implications. *Renin-*

Angiotensin Aldosterone Syst. **2021**, https://doi.org/10.5772/INTECHOPEN.96415.
8. Laragh, J.H.; Baer, L.; Brunner, H.R.; Buhler, F.R.; Sealey, J.E.; Vaughan, E.D.J. Renin, Angiotensin and Aldosterone System in Pathogenesis and Management of Hypertensive Vascular Disease. *Am. J. Med.* **1972**, *52*, 633–652, https://doi.org/10.1016/0002-9343(72)90054-x.
9. Urbizo-Reyes, U.; Liceaga, A.M.; Reddivari, L.; Kim, K.-H.; Anderson, J.M. Enzyme Kinetics, Molecular Docking, and in Silico Characterization of Canary Seed (Phalaris Canariensis L.) Peptides with ACE and Pancreatic Lipase Inhibitory Activity. *J. Funct. Foods* **2022**, *88*, 104892, https://doi.org/10.1016/j.jff.2021.104892.
10. Regulska, K.; Stanisz, B.; Regulski, M.; Murias, M. How to Design a Potent, Specific, and Stable Angiotensin-Converting Enzyme Inhibitor. *Drug Discov. Today* **2014**, *19*, 1731–1743, https://doi.org/10.1016/j.drudis.2014.06.026.
11. He, Z.; Liu, G.; Qiao, Z.; Cao, Y.; Song, M. Novel Angiotensin-I Converting Enzyme Inhibitory Peptides Isolated From Rice Wine Lees: Purification, Characterization, and Structure-Activity Relationship. *Front. Nutr.***2021**, *8*, https://www.frontiersin.org/articles/10.3389/fnut.2021.746113/full.
12. Kavanagh, R. Chapter 18 - Antihypertensive Drugs. In *A Worldwide Survey of New Data in Adverse Drug Reactions*; Ray, S.D.B.T.-S.E. of D.A. **2020**; *42*, 215–226. https://www.sciencedirect.com/science/article/abs/pii/S0378608020300106
13. Sabe, V.T.; Ntombela, T.; Jhamba, L.A.; Maguire, G.E.M.; Govender, T.; Naicker, T.; Kruger, H.G. Current Trends in Computer Aided Drug Design and a Highlight of Drugs Discovered via Computational Techniques: A Review. *Eur. J. Med. Chem.* **2021**, *224*, 113705, https://doi.org/10.1016/j.ejmech.2021.113705.
14. Sofi, M.Y.; Shafi, A.; Masoodi, K.Z. Chapter 22 - Introduction to Computer-Aided Drug Design. In; Sofi, M.Y., Shafi, A., Masoodi, K.Z.B.T.-B. 2022; pp. 215–229. https://www.sciencedirect.com/science/article/pii/B9780323911283000021?via%3Dihub
15. Llorach-Pares, L.; Nonell-Canals, A.; Avila, C.; Sanchez-Martinez, M. Computer-Aided Drug Design (CADD) to De-Orphanize Marine Molecules: Finding Potential Therapeutic Agents for Neurodegenerative and Cardiovascular Diseases. *Mar. Drugs* **2022**, *20*, https://doi.org/10.3390/md20010053.
16. Tiwari, A.; Singh, S. Chapter 13 - Computational Approaches in Drug Designing. In; Singh, D.B., Pathak, R.K.B.T.-B. *Academic Press* **2022**; 207–217 ISBN 978-0-323-89775-4, https://doi.org/10.1016/B978-0-323-89775-4.00010-9.
17. Bharatam, P. V Computer-Aided Drug Design BT - Drug Discovery and Development: From Targets and Molecules to Medicines. In; Poduri, R. *Springer Singapore: Singapore* **2021**, 137–210. https://link.springer.com/chapter/10.1007/978-981-15-5534-3_6
18. Eriksson L, Jaworska J, A. P. Worth, M.T. Cronin, R. MMcDowell, P.G. Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification- and Regression-Based QSARs. *Environ. Health Perspect.* **2003**, *22*, 1361–1376, https://doi.org/10.1289/ehp.5758.
19. Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.-D.; Lee, K.-H.; Tropsha, A. Rational Selection of Training and Test Sets for the Development of Validated QSAR Models. *J. Comput. Aided. Mol. Des.* **2003**, *17*, 241–253, https://doi.org/10.1023/A:1025386326946.
20. Aykul, S.; Martinez-Hackert, E. Determination of Half-Maximal Inhibitory Concentration Using Biosensor-Based Protein Interaction Analysis. *Anal. Biochem.* **2016**, *508*, 97–103, https://doi.org/10.1016/j.ab.2016.06.025.
21. Ghamri, M.; Harkati, D.; Belaidi, S.; Boudergua, S.; Said, R. Ben; Linguerri, R.; Chambaud, G.; Hochlaf, M. Carbazole Derivatives Containing Chalcone Analogues Targeting Topoisomerase II Inhibition: First Principles Characterization and QSAR Modelling. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *242*, 118724, https://doi.org/10.1016/j.saa.2020.118724.
22. Matsumoto, K.; Miyao, T.; Funatsu, K. Ranking-Oriented Quantitative Structure–Activity Relationship Modeling Combined with Assay-Wise Data Integration. *ACS Omega* **2021**, *6*, 11964–11973, https://doi.org/10.1021/acsomega.1c00463.
23. Gonzalez Amaya, J.A.; Cabrera, D.Z.; Matallana, A.M.; Arevalo, K.G.; Guevara-Pulido, J. In-Silico Design of New Enalapril Analogs (ACE Inhibitors) Using QSAR and Molecular Docking Models. *Informatics Med. Unlocked* **2020**, *19*, 100336, https://doi.org/10.1016/j.imu.2020.100336.
24. Yan, W.; Lin, G.; Zhang, R.; Liang, Z.; Wu, L.; Wu, W. Studies on Molecular Mechanism between ACE and Inhibitory Peptides in Different Bioactivities by 3D-QSAR and MD Simulations. *J. Mol. Liq.* **2020**, *304*, 112702, https://doi.org/10.1016/j.molliq.2020.112702.
25. Sagardia, I.; Roa-Ureta, R.H.; Bald, C. A New QSAR Model, for Angiotensin I-Converting Enzyme

Inhibitory Oligopeptides. *Food Chem.* **2013**, *136*, 1370–1376, https://doi.org/10.1016/j.foodchem.2012.09.092.
26. Politi, A.; Durdagi, S.; Moutevelis-Minakakis, P.; Kokotos, G.; Papadopoulos, M.G.; Mavromoustakos, T. Application of 3D QSAR CoMFA/CoMSIA and in Silico Docking Studies on Novel Renin Inhibitors against Cardiovascular Diseases. *Eur. J. Med. Chem.* **2009**, *44*, 3703–3711, https://doi.org/10.1016/j.ejmech.2009.03.040.
27. Carli, N.; Massarelli, I.; Bianucci, A.M. A New Neural Network (Tiling-Contextual Neural Network for Structures, TC-NNfS) Enabling the Treatment of Relatively Small Datasets of Therapeutic Interest: An Application to a Small Dataset of ACE Inhibitors. *Chemom. Intell. Lab. Syst.* **2014**, *137*, 1–9, https://doi.org/10.1016/j.chemolab.2014.06.001.
28. Stoičkov, V.; Šarić, S.; Golubović, M.; Zlatanović, D.; Krtinić, D.; Dinić, L.; Mladenović, B.; Sokolović, D.; Veselinović, A.M. Development of Non-Peptide ACE Inhibitors as Novel and Potent Cardiovascular Therapeutics: An in Silico Modelling Approach. *SAR QSAR Environ. Res.* **2018**, *29*, 503–515, https://doi.org/10.1080/1062936X.2018.1485737.
29. Rahman, M.A. Chalcone: A Valuable Insight into the Recent Advances and Potential Pharmacological Activities. *Chem. Sci. J.* **2011**, https://www.semanticscholar.org/paper/Chalcone%3A-A-Valuable-Insight-into-the-Recent-and-Rahman/64be4e62f02054b129600cf5e4b98226bb6a6581.
30. Vadivoo, V.S.; Mythili, C.V.; Balachander, R.; Vijayalakshmi, N.; Vijaya, P. Computational and Spectral Discussion of Some Substituted Chalcone Derivatives. *Biointerface Res. Appl. Chem.* **2022**, *12*, 7159–7176, https://doi.org/10.33263/BRIAC126.71597176.
31. Al-Ostoot, F.H.; Zabiulla; Salah, S.; Khanum, S.A. Recent Investigations into Synthesis and Pharmacological Activities of Phenoxy Acetamide and Its Derivatives (Chalcone, Indole and Quinoline) as Possible Therapeutic Candidates. *J. Iran. Chem. Soc.* **2021**, *18*, 1839–1875, https://doi.org/10.1007/s13738-021-02172-5.
32. Salehi, B.; Quispe, C.; Chamkhi, I.; El Omari, N.; Balahbib, A.; Sharifi-Rad, J.; Bouyahya, A.; Akram, M.; Iqbal, M.; Docea, A.O.; *et al*. Pharmacological Properties of Chalcones: A Review of Preclinical Including Molecular Mechanisms and Clinical Evidence . *Front. Pharmacol.* **2021**, *11*, https://doi.org/10.3389/fphar.2020.592654.
33. Liu, W.; He, M.; Li, Y.; Peng, Z.; Wang, G. A Review on Synthetic Chalcone Derivatives as Tubulin Polymerisation Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2022**, *37*, 9–38, https://doi.org/10.1080/14756366.2021.1976772.
34. Marquina, S.; Maldonado-Santiago, M.; Sánchez-Carranza, J.N.; Antúnez-Mojica, M.; González-Maya, L.; Razo-Hernández, R.S.; Alvarez, L. Design, Synthesis and QSAR Study of 2′-Hydroxy-4′-Alkoxy Chalcone Derivatives That Exert Cytotoxic Activity by the Mitochondrial Apoptotic Pathway. *Bioorg. Med. Chem.* **2019**, *27*, 43–54, https://doi.org/10.1016/j.bmc.2018.10.045.
35. Dimić, D.; Mercader, A.G.; Castro, E.A. Chalcone Derivative Cytotoxicity Activity against MCF-7 Human Breast Cancer Cell QSAR Study. *Chemom. Intell. Lab. Syst.* **2015**, *146*, 378–384, https://doi.org/10.1016/j.chemolab.2015.06.011.
36. Rybka, M.; Mercader, A.G.; Castro, E.A. Predictive QSAR Study of Chalcone Derivatives Cytotoxicity Activity against HT-29 Human Colon Adenocarcinoma Cell Lines. *Chemom. Intell. Lab. Syst.* **2014**, *132*, 18–29, https://doi.org/10.1016/j.chemolab.2013.12.005.
37. Gomes, M.N.; Braga, R.C.; Grzelak, E.M.; Neves, B.J.; Muratov, E.; Ma, R.; Klein, L.L.; Cho, S.; Oliveira, G.R.; Franzblau, S.G.; *et al*. QSAR-Driven Design, Synthesis and Discovery of Potent Chalcone Derivatives with Antitubercular Activity. *Eur. J. Med. Chem.* **2017**, *137*, 126–138, https://doi.org/10.1016/j.ejmech.2017.05.026.
38. Chen, Z.; Li, P.; Hu, D.; Dong, L.; Pan, J.; Luo, L.; Zhang, W.; Xue, W.; Jin, L.; Song, B. Synthesis, Antiviral Activity, and 3D-QSAR Study of Novel Chalcone Derivatives Containing Malonate and Pyridine Moieties. *Arab. J. Chem.* **2019**, *12*, 2685–2696, https://doi.org/10.1016/j.arabjc.2015.05.003.
39. Xue, C.X.; Cui, S.Y.; Liu, M.C.; Hu, Z.D.; Fan, B.T. 3D QSAR Studies on Antimalarial Alkoxylated and Hydroxylated Chalcones by CoMFA and CoMSIA. *Eur. J. Med. Chem.* **2004**, *39*, 745–753, https://doi.org/10.1016/j.ejmech.2004.05.009.
40. Bonesi, M.; Loizzo, M.R.; Statti, G.A.; Michel, S.; Tillequin, F.; Menichini, F. The Synthesis and Angiotensin Converting Enzyme (ACE) Inhibitory Activity of Chalcones and Their Pyrazole Derivatives. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1990–1993, https://doi.org/10.1016/j.bmcl.2010.01.113.
41. Kantevari, S.; Addla, D.; Bagul, P.K.; Sridhar, B.; Banerjee, S.K. Synthesis and Evaluation of Novel 2-Butyl-

4-Chloro-1-Methylimidazole Embedded Chalcones and Pyrazoles as Angiotensin Converting Enzyme (ACE) Inhibitors. *Bioorg. Med. Chem.* **2011**, *19*, 4772–4781, https://doi.org/10.1016/j.bmc.2011.06.085.
42. Bukhari, S.N.A.; Butt, A.M.; Amjad, M.W.B.; Ahmad, W.; Shah, V.H.; Trivedi, A.R. Synthesis and Evaluation of Chalcone Analogues Based Pyrimidines as Angiotensin Converting Enzyme Inhibitors. *Pakistan J. Biol. Sci. PJBS* **2013**, *16*, 1368–1372, https://doi.org/10.3923/pjbs.2013.1368.1372.
43. Kang, D.G.; Kim, Y.C.; Sohn, E.J.; Lee, Y.M.; Lee, A.S.; Yin, M.H.; Lee, H.S. Hypotensive Effect of Butein *via* the Inhibition of Angiotensin Converting Enzyme. *Biol. Pharm. Bull.* **2003**, *26*, 1345–1347, https://doi.org/10.1248/bpb.26.1345.
44. Davies, M.; Nowotka, M.; Papadatos, G.; Dedman, N.; Gaulton, A.; Atkinson, F.; Bellis, L.; Overington, J.P. ChEMBL Web Services: Streamlining Access to Drug Discovery Data and Utilities. *Nucleic Acids Res.* **2015**, *43*, W612–W620, https://doi.org/10.1093/nar/gkv352.
45. Bukhari, A.P.D.S.N.A.; Butt, A.M.; Amjad, M.; Waqas, A.; Shah, V.H.; Trivedi, A. Synthesis and Evaluation of Chalcone Analogues Based Pyrimidines as Angiotensin Converting Enzyme Inhibitors. *Pakistan J. Biol. Sci.* **2013**, *Online Fir*, https://doi.org/10.3923/pjbs.2013.1368.1372.
46. Berthold, M.R.; Cebron, N.; Dill, F.; Gabriel, T.R.; Kötter, T.; Meinl, T.; Ohl, P.; Thiel, K.; Wiswedel, B. KNIME - the Konstanz Information Miner: Version 2.0 and Beyond. *SIGKDD Explor. Newsl.* **2009**, *11*, 26–31, https://doi.org/10.1145/1656274.1656280.
47. Shah, S.K.; Chaple, D.R. 2D-QSAR Modeling of Quinazolinone Derivatives as Angiotensin II Type 1a Receptor Blockers. *Int. J. Quant. Struct. Relationships* **2021**, *7*, 1–20, https://doi.org/10.4018/ijqspr.290012.
48. O'Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3*, 33, https://doi.org/10.1186/1758-2946-3-33.
49. Masand, V.H.; Rastija, V. PyDescriptor: A New PyMOL Plugin for Calculating Thousands of Easily Understandable Molecular Descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18, https://doi.org/10.1016/j.chemolab.2017.08.003.
50. Tetko, I. V; Tanchuk, V.Y.; Villa, A.E.P. Prediction of N-Octanol/Water Partition Coefficients from PHYSPROP Database Using Artificial Neural Networks and E-State Indices. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1407–1421, https://doi.org/10.1021/ci010368v.
51. Yap, C.W. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *J. Comput. Chem.* **2011**, *32*, 1466–1474, https://doi.org/10.1002/jcc.21707.
52. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. *J. Comput. Chem.* **2013**, *34*, 2121–2132, https://doi.org/10.1002/jcc.23361.
53. Papa, E.; Gramatica, P. QSPR as a Support for the EU REACH Regulation and Rational Design of Environmentally Safer Chemicals: PBT Identification from Molecular Structure. *Green Chem.* **2010**, *12*, 836–843, https://doi.org/10.1039/B923843C.
54. Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inform.* **2010**, *29*, 476–488, https://doi.org/10.1002/minf.201000061.
55. Kumar, V.; De, P.; Ojha, P.K.; Saha, A.; Roy, K. A Multi-Layered Variable Selection Strategy for QSAR Modeling of Butyrylcholinesterase Inhibitors. *Curr. Top. Med. Chem.* **2020**, *20*, 1601–1627, https://doi.org/10.2174/1568026620666200616142753.
56. Roy, K.; Kar, S.; Ambure, P. On a Simple Approach for Determining Applicability Domain of QSAR Models. *Chemom. Intell. Lab. Syst.* **2015**, *145*, 22–29, https://doi.org/10.1016/j.chemolab.2015.04.013.
57. Roy, K.; Ambure, P.; Kar, S.; Ojha, P.K. Is It Possible to Improve the Quality of Predictions from an "Intelligent" Use of Multiple QSAR/QSPR/QSTR Models? *J. Chemom.* **2018**, *32*, e2992, https://doi.org/10.1002/cem.2992.
58. Sander, T.; Freyss, J.; von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. *J. Chem. Inf. Model.* **2015**, *55*, 460–473, https://doi.org/10.1021/ci500588j.
59. Shah, S.; Chaple, D.; Arora, S.; Yende, S.; Moharir, K.; Lohiya, G. Exploring the Active Constituents of Oroxylum Indicum in Intervention of Novel Coronavirus (COVID-19) Based on Molecular Docking Method. *Netw. Model. Anal. Heal. Informatics Bioinforma.* **2021**, *10*, 8, https://doi.org/10.1007/s13721-020-00279-y.
60. Arora, S.; Lohiya, G.; Moharir, K.; Shah, S.; Yende, S. Identification of Potential Flavonoid Inhibitors of the SARS-CoV-2 Main Protease 6YNQ: A Molecular Docking Study. *Digit. Chinese Med.* **2020**, *3*, 239–248, https://doi.org/10.1016/j.dcmed.2020.12.003.
61. Masand, V.H.; El-Sayed, N.N.E.; Bambole, M.U.; Patil, V.R.; Thakur, S.D. Multiple Quantitative Structure-

Activity Relationships (QSARs) Analysis for Orally Active Trypanocidal N-Myristoyltransferase Inhibitors. *J. Mol. Struct.* **2019**, *1175*, 481–487, https://doi.org/10.1016/j.molstruc.2018.07.080.
62. Jawarkar, R.D.; Bakal, R.L.; Zaki, M.E.A.; Al-Hussain, S.; Ghosh, A.; Gandhi, A.; Mukerjee, N.; Samad, A.; Masand, V.H.; Lewaa, I. QSAR Based Virtual Screening Derived Identification of a Novel Hit as a SARS CoV-229E 3CLpro Inhibitor: GA-MLR QSAR Modeling Supported by Molecular Docking, Molecular Dynamics Simulation and MMGBSA Calculation Approaches. *Arab. J. Chem.* **2022**, *15*, 103499, https://doi.org/10.1016/j.arabjc.2021.103499.
63. Dearden, J.C.; Cronin, M.T.D.; Kaiser, K.L.E. How Not to Develop a Quantitative Structure–Activity or Structure–Property Relationship (QSAR/QSPR). *SAR QSAR Environ. Res.* **2009**, *20*, 241–266, https://doi.org/10.1080/10629360902949567.
64. Gramatica, P. External Evaluation of QSAR Models, in Addition to Cross-Validation: Verification of Predictive Capability on Totally New Chemicals. *Mol. Inform.* **2014**, *33*, 311–314, https://doi.org/10.1002/minf.201400030.
65. Hajalsiddig, T.T.H.; Osman, A.B.M.; Saeed, A.E.M. 2D-QSAR Modeling and Molecular Docking Studies on 1H-Pyrazole-1-Carbothioamide Derivatives as EGFR Kinase Inhibitors. *ACS Omega* **2020**, *5*, 18662–18674, https://doi.org/10.1021/acsomega.0c01323.
66. Jawarkar, R.D.; Bakal, R.L.; Khatale, P.N.; Lewaa, I.; Jain, C.M.; Manwar, J. V; Jaiswal, M.S. QSAR, Pharmacophore Modeling and Molecular Docking Studies to Identify Structural Alerts for Some Nitrogen Heterocycles as Dual Inhibitor of Telomerase Reverse Transcriptase and Human Telomeric G-Quadruplex DNA. *Futur. J. Pharm. Sci.* **2021**, *7*, 231, https://doi.org/10.1186/s43094-021-00380-7.
67. Cañizares-Carmenate, Y.; Mena-Ulecia, K.; Perera-Sardiña, Y.; Torrens, F.; Castillo-Garit, J.A. An Approach to Identify New Antihypertensive Agents Using Thermolysin as Model: In Silico Study Based on QSARINS and Docking. *Arab. J. Chem.* **2019**, *12*, 4861–4877, https://doi.org/10.1016/j.arabjc.2016.10.003.
68. Bakal, R.L.; Jawarkar, R.D.; Manwar, J. V; Jaiswal, M.S.; Ghosh, A.; Gandhi, A.; Zaki, M.E.A.; Al-Hussain, S.; Samad, A.; Masand, V.H.; *et al*. Identification of Potent Aldose Reductase Inhibitors as Antidiabetic (Anti-Hyperglycemic) Agents Using QSAR Based Virtual Screening, Molecular Docking, MD Simulation and MMGBSA Approaches. *Saudi Pharm. J.* **2022**, https://doi.org/10.1016/j.jsps.2022.04.003.
69. Chtita, S.; Belhassan, A.; Bakhouch, M.; Taourati, A.I.; Aouidate, A.; Belaidi, S.; Moutaabbid, M.; Belaaouad, S.; Bouachrine, M.; Lakhlifi, T. QSAR Study of Unsymmetrical Aromatic Disulfides as Potent Avian SARS-CoV Main Protease Inhibitors Using Quantum Chemical Descriptors and Statistical Methods. *Chemom. Intell. Lab. Syst.* **2021**, *210*, 104266, https://doi.org/10.1016/j.chemolab.2021.104266.
70. Tabti, K.; Elmchichi, L.; Sbai, A.; Maghat, H.; Bouachrine, M.; Lakhlifi, T.; Ghosh, A. In Silico Design of Novel PIN1 Inhibitors by Combined of 3D-QSAR, Molecular Docking, Molecular Dynamic Simulation and ADMET Studies. *J. Mol. Struct.* **2022**, *1253*, 132291, https://doi.org/10.1016/j.molstruc.2021.132291.
71. Gramatica, P. Principles of QSAR Modeling: Comments and Suggestions From Personal Experience. *Int. J. Quant. Struct. Relationships* **2020**, *5*, https://doi.org/10.4018/IJQSPR.20200701.oa1.
72. Mauri, A. AlvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints. In *Methods in Pharmacology and Toxicology*; **2020**, https://link.springer.com/protocol/10.1007/978-1-0716-0150-1_32.
73. Kumar, V.; Saha, A.; Roy, K. In Silico Modeling for Dual Inhibition of Acetylcholinesterase (AChE) and Butyrylcholinesterase (BuChE) Enzymes in Alzheimer's Disease. *Comput. Biol. Chem.* **2020**, *88*, 107355, https://doi.org/10.1016/j.compbiolchem.2020.107355.
74. Wang, X.; Wang, J.; Lin, Y.; Ding, Y.; Wang, Y.; Cheng, X.; Lin, Z. QSAR Study on Angiotensin-Converting Enzyme Inhibitor Oligopeptides Based on a Novel Set of Sequence Information Descriptors. *J. Mol. Model.* **2011**, *17*, 1599–1606, https://doi.org/10.1007/s00894-010-0862-x.

Supplementary materials

Table S1. List of descriptors used in the QSAR models development.

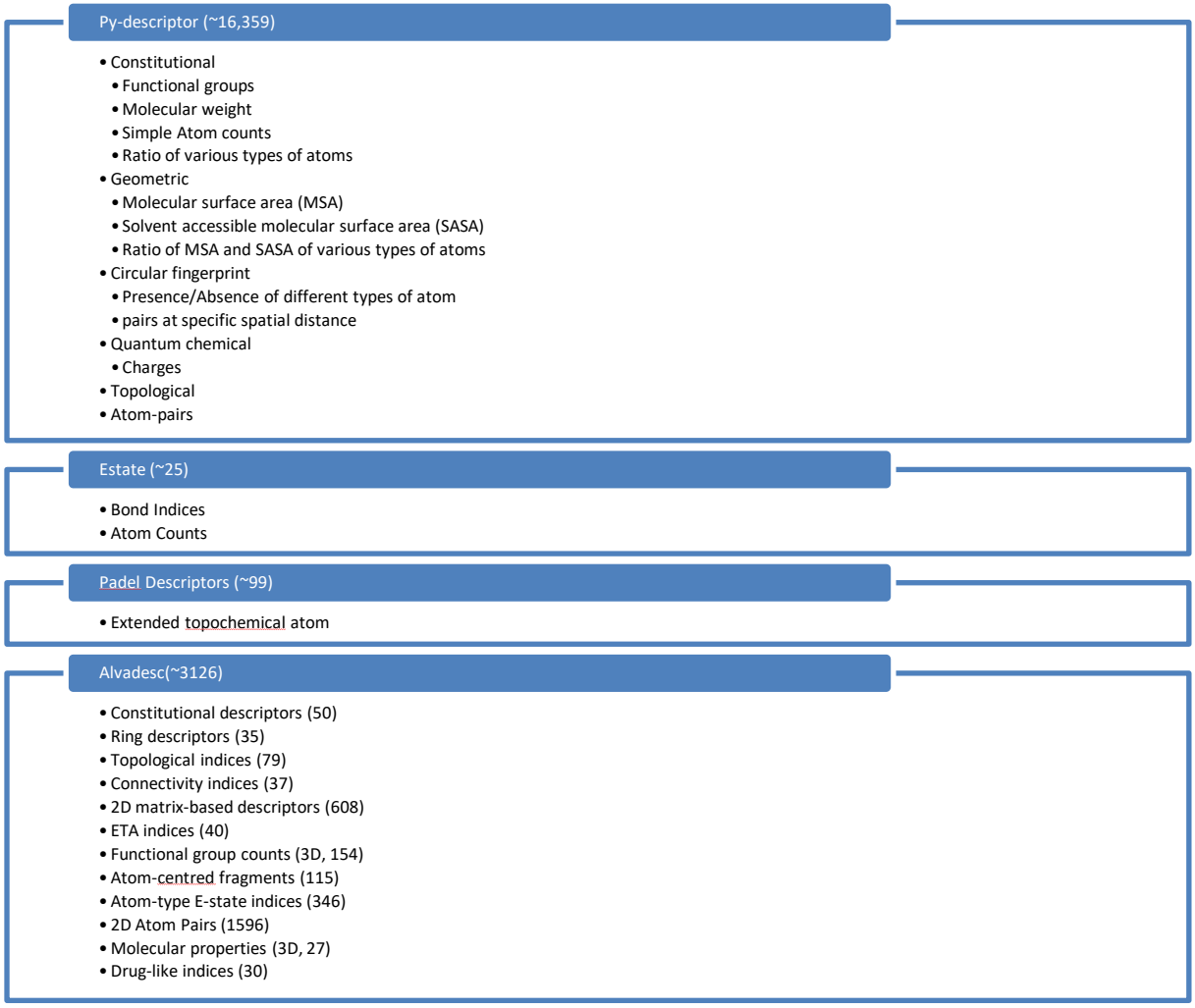


Table S2. Description of Important Statistical Parameters

Performance parameter	**Calculated during**	**Formula**	**Description**
R^2, R^2_{ext}	Training, External, validation	$R^2 = 1-\frac{\sum_{i=1}^{n}(y_i-\check{y}_i)^2}{\sum_{i=1}^{n}(y_i-y)^2} = 1-\frac{RSS}{ISS}$	Explained variance; coefficient of determination, squareof the multiple correlation coefficient
$R^2_{adj.}$	training	$R^2\ adj = R^2 - (1-R^2) \times \frac{p}{n-p-1}$	R^2 corrected with the degree of freedom
$R^2-R^2_{adj}$	training	See above	Difference of the two
K_x	training	$K = \frac{\sum_{i=1}^{n}(y_i-\check{y}_i)}{\sum_{i=1}^{n}(\check{y}_i)^2}$	Inter-correlation among descriptors

Performance parameter	Calculated during	Formula	Description
ΔK	training	Based on PCA,	Difference of correlation among descriptors (K_x) and the descriptors plus responses (K_{xy})
RMSE	training, int. val., ext. val.	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i-y)^2}{n}}$	Root mean square error
MAE	training, int. val., ext. val.	$MAE = \frac{\sum_{i=1}^{n}\lvert yi-\hat{y}\rvert}{n}$	Mean absolute error
RSS	training	$RSS = \sum_{i-1}^{n}(y_i - y_i)^2$	Residual sum of squares
CCC	training, int. val., ext. val.	$CCC = \frac{2\sum_{i=1}^{n}(y_i-y)(y_i-y)}{2\sum_{i=1}^{n}(y_i-y)^2+\sum_{i=1}^{n}(y_i-y)^2+n(y_i-y)^2}$	Coefficient of concordance, concordance correlationcoefficient
S	training	$S = \sqrt{\frac{1}{N-1}}\sum_{i=1}^{N}(y_i - y_i)^2$	Standard error of the estimate
F	training	$F = (\frac{\sum_{1=i}^{n}(\bar{y}-y_i)^2}{p-1})/(\frac{\sum_{i=1}^{n}(\bar{y}-y_i)^2}{n-p})$	Fisher value
Q^2_{LOO}	Internal validation	$Q^2_{LOO} = 1-\frac{\sum_{i=1}^{n}(y_i-y_{i/i})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} -1 - \frac{PRESS}{TSS}$	Leave-one-out cross-validated square of the (multiple) correlation coefficient
$R^2 - Q^2_{LOO}$	Internal validation	See above	Difference of the two
PRESS	internal, external validation	$PRESS= \sum_{i=1}^{n}(y_i - y_{i/i})^2$	Predicted residual sum of squares (either cross-validatedor calculated on the external set)
Q^2_{LMO}	internal validation	$Q^2_{LMO} = 1-\frac{\sum_{j=1}^{m}\sum_{i=1}^{n}(y_i-\hat{y}_{i/j})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$	Leave-many-out cross-validated square of the (multiple) correlation coefficient
$R^2_{Y\text{-}SCRAMBLE}$	Internal validation	See above	R^2 of the training set with Y-scrambling
$RMSE_{Avg,\ Y\text{-}SCRAMBLE}$	Internal validation	See above	Average RMSE with Y-scrambling
$Q^2_{Y\text{-}SCRAMBLE}$	Internal validation	See above	Q^2_{LOO} of the training set with Y-scrambling
$R^2_{RND\text{-}DESCR}$	Internal validation	See above	R^2 of the training set with randomized descriptors
$Q^2_{RND\text{-}DESCR}$	Internal Validation	See above	Q^2_{LOO} of the training set with randomized descriptors
$R^2_{RND\text{-}RESP}$	Internal validation	See above	R^2 of the training set with randomized responses
$Q^2_{RND\text{-}RESP}$	Internal validation	See above	Q^2_{LOO} of the training set with randomized responses

Performance parameter	**Calculated during**	**Formula**	**Description**
Q^2_{F1}	External validation	$Q^2_{F1} = 1 - \frac{\sum_{i=1}^{nEXT}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{nEXT}(y_i - y_{TR})^2}$	Definition 1 in for Q^2 of the external test set TR: training set, EXT: external test set
Q^2_{F2}	External validation	$Q^2_{F2} = 1 - \frac{\sum_{i=1}^{nEXT}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{nEXT}(y_i - y_{EXT})^2}$	Definition 2 in for Q^2 of the external test set, EXT: external test set
Q^2_{F3}	External validation	$Q^2_{F3} = 1 - \frac{\sum_{i=1}^{nEXT}(y_i - \hat{y}_i)^2/n_{EXT}}{\sum_{i=1}^{nEXT}(y_i - \bar{y}_{TR})^2/n_{TR}}$	Definition 3 in for Q^2 of the external test set TR: training set, EXT: external test set
r^2_m	External validation	$r^2_m = \frac{r^2_m + r'^2_m}{2}$	Here, r2m = R2× $(1 - \sqrt{R^2 - R_0^2})$ were R_0^2 is the squared correlation coefficient without intercept r'^2_m is the same as r^2_m with x and y axis exchanged.
Δr^2_m	External validation	$\Delta r^2_m = r^2_m - r'^2_m$	See above.

Table S3. Toxicity Screening of novel designed compounds

Compound ID	Smiles	AMES Toxicity	Max tolerated dose	hERG I inhibitor	hERG II inhibitor	Oral Rat acute toxicity (LD50)	Oral Rat Chronic toxicity (LOAEL)	Hepatotoxicity	Skin Sensitization	T. Pyriformis	Minnow Toxicity
11	O=C(c1c(F)c(F)c(c(c1F)F)F)/C=C/c1ccc(s1)B(O)O	No	0.622	No	No	2.113	2.688	Yes	No	0.319	0.367
278	O=C(c1cc(cc(c1)[N+](=O)[O-])[N+](=O)[O-])/C=C/c1ccccc1	Yes	-0.094	No	No	2.508	1.544	No	No	0.729	-0.464
280	COc1ccc(cc1)/C=C/C(=O)c1cc(cc(c1)[N+](=O)[O-])[N+](=O)[O-]	Yes	-0.071	No	No	2.28	2.067	No	No	0.365	-1.176
281	O=C(c1ccccc1)/C=C/c1cc(cc(c1)[N+](=O)[O-])[N+](=O)[O-]	Yes	-0.005	No	No	2.505	1.449	No	No	0.669	-0.789
283	CN(c1ccc(cc1)C1=NN=C(C1)c1ccccc1)C	Yes	0.276	No	No	2.107	0.598	No	No	2.737	0.366
285	[O-][N+](=O)c1cccc(c1)C1=NN=C(C1)c1ccccc1	Yes	0.131	No	No	2.384	1.403	No	No	1.825	0.201
286	[O-][N+](=O)c1cccc(c1)C1=NN=C(C1)c1cccc(c1)[N+](=O)[O-]	Yes	-0.232	No	No	2.565	1.129	No	No	1.38	-0.29
287	CN(c1ccc(cc1)C1=NN=C(C1)c1ccccc(c1)[N+](=O)[O-])C	Yes	0.112	No	No	2.497	1.28	No	No	1.72	-0.181
288	COc1ccc(cc1)C1=NN=C(C1)c1ccccc(c1)[N+](=O)[O-]	Yes	0.067	No	No	2.477	1.338	No	No	1.407	-1.02
305	COc1ccc(cc1)C1=NN=C(C1)c1cc(cc(c1)[N+](=O)[O-])[N+](=O)[O-]	Yes	-0.384	No	No	2.572	1.05	No	No	0.819	-1.821
333	COc1cc(ccc1O)[C@@H]1NC(=O)NC(=C1)c1ccccc1	No	-0.313	No	Yes	2.064	1.454	No	No	0.7	1.136
335	O=C1N[C@H](C=C(N1)c1cccc(c1)[N+](=O)[O-])c1ccc(cc1)O	No	0.292	No	Yes	2.951	2.223	No	No	0.381	0.663
337	COc1ccc(cc1)[C@@H]1NC(=O)NC(=C1)c1cccc(c1)[N+](=O)[O-]	Yes	-0.15	No	Yes	2.692	2.001	Yes	No	0.651	-0.063
338	COc1cc(ccc1O)[C@@H]1NC(=O)NC(=C1)c1cccc(c1)[N+](=O)[O-]	No	-0.197	No	Yes	3.356	1.974	Yes	No	0.412	1.158
339	O=C1N[C@H](C=C(N1)c1ccc(cc1)Br)c1cccc(c1)[N+](=O)[O-]	Yes	0.208	No	Yes	2.244	1.728	No	No	0.51	-0.157

Compound ID	Smiles	AMES Toxicity	Max tolerated dose	hERG I inhibitor	hERG II inhibitor	Oral Rat acute toxicity (LD50)	Oral Rat Chronic toxicity (LOAEL)	Hepatotoxicity	Skin Sensitization	T. Pyriformis	Minnow Toxicity
340	COc1ccc(cc1)[C@@H]1NC(=O)NC(=C1)c1ccc(cc1)Br	No	0.124	No	Yes	2.055	1.395	No	No	1.215	-0.297
341	COc1cc(ccc1O)[C@@H]1NC(=O)NC(=C1)c1ccc(cc1)Br	No	-0.323	No	Yes	2.106	1.367	No	No	0.688	0.676
342	O=C1N[C@H](C=C(N1)c1ccc(cc1)O)c1cccc(c1)[N+](=O)[O-]	No	0.321	No	Yes	2.697	1.925	Yes	No	0.353	1.053
343	Oc1ccc(cc1)[C@@H]1NC(=O)NC(=C1)c1ccc(cc1)O	No	-0.082	No	Yes	2.04	1.973	No	No	0.712	1.606
344	COc1ccc(cc1)[C@@H]1NC(=O)NC(=C1)c1ccc(cc1)O	Yes	-0.234	No	Yes	2.031	1.819	No	No	0.735	0.704
345	COc1cc(ccc1O)[C@@H]1NC(=O)NC(=C1)c1ccc(cc1)O	**No**	**-0.054**	**No**	**Yes**	**2.318**	**1.746**	**Yes**	**No**	**0.488**	**1.775**
349	COc1cc(ccc1O)C1=C[C@@H](NC(=N1)O)c1ccccc1	**No**	**-0.101**	**No**	**Yes**	**2.121**	**0.758**	**No**	**No**	**0.637**	**1.636**
350	OC1=NC(=C[C@@H](N1)c1cccc(c1)[N+](=O)[O-])c1ccc(cc1)O	**Yes**	**0.075**	**No**	**Yes**	**3.528**	**2.246**	**No**	**No**	**0.79**	**0.723**
352	COc1cc(ccc1O)C1=C[C@@H](NC(=N1)O)c1cccc(c1)[N+](=O)[O-]	**No**	**0.067**	**No**	**Yes**	**3.247**	**1.288**	**Yes**	**No**	**0.448**	**1.29**
353	COc1ccc(cc1)C1=C[C@@H](NC(=N1)O)c1ccc(cc1)Br	**No**	**0.209**	**No**	**Yes**	**2.409**	**1.252**	**No**	**No**	**1.181**	**-0.046**
354	OC1=NC(=C[C@@H](N1)c1ccc(cc1)O)c1cccc(c1)[N+](=O)[O-]	**Yes**	**0.041**	**No**	**Yes**	**3.317**	**1.948**	**No**	**No**	**0.738**	**1.113**
355	Oc1ccc(cc1)[C@@H]1NC(=NC(=C1)c1ccc(cc1)O)O	**No**	**-0.151**	**No**	**Yes**	**2.459**	**2.097**	**No**	**No**	**0.702**	**1.494**
356	COc1ccc(cc1)C1=C[C@@H](NC(=N1)O)c1ccc(cc1)O	**No**	**0.004**	**No**	**Yes**	**2.281**	**1.619**	**No**	**No**	**0.672**	**0.955**
357	COc1cc(ccc1C)C1=C[C@@H](NC(=N1)O)c1ccc(cc1)O	**No**	**0.042**	**No**	**Yes**	**2.305**	**1.215**	**No**	**No**	**0.644**	**1.015**
364	OC(=O)[C@H]1CCCN1C1=NN=C(C1)c1ccccc1	**No**	**-0.181**	**No**	**No**	**2.714**	**1.405**	**Yes**	**No**	**1.033**	**2.009**
365	OC(=O)[C@H]1CCCN1C1=NN=C(C1)c1cccc(c1)[N+](=O)[O-]	**Yes**	**-0.668**	**No**	**No**	**2.272**	**1.476**	**No**	**No**	**0.965**	**1.798**
367	OC(=O)[C@H]1CCCN1C1=NN=C(C1)c1ccc(cc1)N(C)C	**No**	**-0.403**	**No**	**No**	**2.775**	**1.399**	**Yes**	**No**	**1.139**	**1.906**

Compound ID	Smiles	AMES Toxicity	Max tolerated dose	hERG I inhibitor	hERG II inhibitor	Oral Rat acute toxicity (LD50)	Oral Rat Chronic toxicity (LOAEL)	Hepatotoxicity	Skin Sensitization	T. Pyriformis	Minnow Toxicity
368	COc1ccc(cc1)C1=NN=C(C1)N1CCC[C@@H]1C(=O)O	No	0.44	No	No	2.783	1.678	No	No	0.287	1.945
370	COc1cc(ccc1O)[C@H]1NN=C(C1)N1CCC[C@@H]1C(=O)O	No	-0.055	No	No	2.907	1.558	Yes	No	0.285	2.839
385	O=C1NC(=C[C@@H](N1)c1ccccc1)N1CCC[C@@H]1C(=O)O	No	0.284	No	No	2.26	1.81	Yes	No	0.285	1.815
386	O=C1NC(=C[C@@H](N1)c1cccc(c1)[N+](=O)[O-])N1CCC[C@@H]1C(=O)O	No	-0.328	No	No	2.063	2.083	No	No	0.285	1.507
387	O=C1NC(=C[C@@H](N1)c1ccc(cc1)O)N1CCC[C@@H]1C(=O)O	No	0.162	No	No	2.295	1.993	Yes	No	0.285	2.356
389	COc1ccc(cc1)[C@@H]1NC(=O)NC(=C1)N1CCC[C@@H]1C(=O)O	No	1.183	No	No	2.617	1.834	Yes	No	0.285	2.758
391	COc1cc(ccc1O)[C@@H]1NC(=O)NC(=C1)N1CCC[C@@H]1C(=O)O	No	0.419	No	No	2.621	2.596	Yes	No	0.285	3.495
399	Nc1ccc(cc1)/C=C/C(=O)c1cccc(c1)[N+](=O)[O-]	Yes	0.166	No	No	2.079	1.285	Yes	No	0.586	0.583
401	Nc1ccc(cc1)C1=NN=C(C1)c1ccccc1	Yes	0.23	No	No	2.409	0.775	No	No	1.204	0.825
402	Nc1ccc(cc1)C1=NN=C(C1)c1cccc(c1)[N+](=O)[O-]	Yes	0.009	No	No	2.358	1.267	No	No	0.849	0.279
406	OC(=O)[C@H]1CCCN1C1=NN=C(C1)c1ccc(cc1)N	No	-0.542	No	No	2.246	1.129	Yes	No	0.589	2.366
419	O=C1N[C@H](C=C(N1)c1cccc(c1)[N+](=O)[O-])c1ccc(cc1)N	Yes	0.301	No	Yes	2.761	2.235	Yes	No	0.329	0.676
425	OC1=NC(=C[C@@H](N1)c1cccc(c1)[N+](=O)[O-])c1ccc(cc1)N	Yes	0.081	No	Yes	2.62	2.258	No	No	0.522	0.736
430	O=C(c1cccc(c1)[N+](=O)[O-])/C=C/c1ccccc1O	Yes	0.336	No	No	2.989	1.204	No	No	1.19	0.293
452	OC1=NC(=C[C@@H](N1)c1ccccc1)c1ccccc1O	Yes	-0.011	No	Yes	2.338	1.264	No	No	1.164	1.183
453	OC1=NC(=C[C@@H](N1)c1cccc(c1)N(=O)=O)c1ccccc1O	Yes	0.075	No	Yes	3.528	2.246	No	No	0.79	0.723

Compound ID	Smiles	AMES Toxicity	Max tolerated dose	hERG I inhibitor	hERG II inhibitor	Oral Rat acute toxicity (LD50)	Oral Rat Chronic toxicity (LOAEL)	Hepatotoxicity	Skin Sensitization	T. Pyriformis	Minnow Toxicity
454	Brc1ccc(cc1)[C@@H]1NC(=NC(=C1)c1ccccc1O)O	No	0.088	No	Yes	2.414	2.169	No	No	0.767	0.406
455	Oc1ccc(cc1)[C@@H]1NC(=NC(=C1)c1ccccc1O)O	No	0.09	No	Yes	2.272	2.356	No	No	0.657	1.211