



# HHS Public Access

Author manuscript

*Curr Top Med Chem.* Author manuscript; available in PMC 2015 October 05.

Published in final edited form as:

*Curr Top Med Chem.* 2014 ; 14(11): 1399–1415.

## Tuning hERG out: Antitarget QSAR Models for Drug Development

Rodolpho C. Braga<sup>1,2,†</sup>, Vinícius M. Alves<sup>1,†</sup>, Meryck F. B. Silva<sup>1</sup>, Eugene Muratov<sup>3</sup>, Denis Fourches<sup>3</sup>, Alexander Tropsha<sup>3</sup>, and Carolina H. Andrade<sup>1,\*</sup>

<sup>1</sup>Laboratory for Molecular Modeling and Drug Design, Faculty of Pharmacy, Federal University of Goiás, Rua 240, Qd. 87, Setor Leste Universitário, Goiânia, Goiás 74605-170, Brazil

<sup>2</sup>Laboratory of Nuclear Magnetic Resonance, Chemistry Institute, Federal University of Goiás, P.O. Box 131, Goiânia, Goiás 74001-970, Brazil

<sup>3</sup>Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA

### Abstract

Several non-cardiovascular drugs have been withdrawn from the market due to their inhibition of hERG K<sup>+</sup> channels that can potentially lead to severe heart arrhythmia and death. As hERG safety testing is a mandatory FDA-required procedure, there is a considerable interest for developing predictive computational tools to identify and filter out potential hERG blockers early in the drug discovery process. In this study, we aimed to generate predictive and well-characterized quantitative structure–activity relationship (QSAR) models for hERG blockage using the largest publicly available dataset of 11,958 compounds from the ChEMBL database. The models have been developed and validated according to OECD guidelines using four types of descriptors and four different machine-learning techniques. The classification accuracies discriminating blockers from non-blockers were as high as 0.83–0.93 on external set. Model interpretation revealed several SAR rules, which can guide structural optimization of some hERG blockers into non-blockers. We have also applied the generated models for screening the World Drug Index (WDI) database and identify putative hERG blockers and non-blockers among currently marketed drugs. The developed models can reliably identify blockers and non-blockers, which could be useful for the scientific community. A freely accessible web server has been developed allowing users to identify putative hERG blockers and non-blockers in chemical libraries of their interest (<http://labmol.farmacia.ufg.br/predherg>).

\* Address correspondence to this author at LabMol, Faculdade de Farmácia, Universidade Federal de Goiás, Rua 240, Qd. 87, Setor Leste Universitário, Goiânia, Goiás 74605-170, Brazil. Tel: + 55 62 3209-6451; Fax: +55 62 3209 6037; carolina@ufg.br.

#### †AUTHOR CONTRIBUTIONS

These authors contributed equally to this work.

#### CONFLICT OF INTERESTS

The authors confirm that this article content has no conflicts of interest.

#### SUPPLEMENTARY INFORMATION

Supplementary tables and curated datasets are available as supplementary material (<http://labmol.farmacia.ufg.br/predherg>).

## Keywords

antitarget; drug development; hERG; QSAR modeling; virtual screening

---

## INTRODUCTION

The human *ether-à-go-go* related gene (hERG) encodes the subunit of a delayed rectifier voltage-gated K<sup>+</sup> channel. These channels are expressed in a wide array of tissues, but their physiological function is best characterized in cardiac cells, where they play a critical role in the repolarization of the cardiac action potential [1]. The QT interval is the measure of the duration of ventricular depolarization and repolarization in an electrocardiogram. hERG inhibition is associated with QT interval prolongation and can evolve into a lethal cardiac arrhythmia [2]. Many drugs covering a broad spectrum of pharmaceutical classes have been withdrawn from the market or have had their usage limited due to blockage of the hERG, e.g., astemizole, terfenadine [3], cisapride [4], sertindole [5], terolidine, droperidol, lidoflazine, and grepafloxacin [6]. Those diverse chemicals can indeed interact with hERG because of its notorious ligand promiscuity, mainly due to its large hydrophobic intracellular binding pocket and its multiple states (open, inactive, and closed) [7].

The increased incidence of sudden death related to non-cardiac drugs earmarked the hERG as one of the most important antitargets to be considered in the early stages of drug development processes, elicited the interest of academia and industry, and has become a major concern for regulatory agencies [8]. Currently, hERG safety is required by the US Food and Drug Administration (FDA) in clinical trials [9, 10]. To identify the potential of a substance to delay ventricular repolarization and to estimate the risk for QT interval prolongation in humans, several experiments are conducted. Although new techniques have appeared in the last few years [11–13], conventional patch-clamp electrophysiology remains the “gold-standard.” In this test, the electric current passing through hERG channels expressed in cells, is measured [14]. Often, the cell lines utilized as an expression system for the hERG in studies include human embryonic kidney (HEK) 293 cells, Chinese hamster ovary (CHO) cells, or *Xenopus laevis* oocytes (XO) cells. No proven advantage of one cell line over another is known, but discrepancies of outcomes are clearly related to the lack of standardized protocols. Unfortunately, these assays are time-consuming, expensive, labor-intensive, and suffer from low throughput [15, 16]. High-throughput ion-channel screening data exhibit significant levels of variability, which affects the reliability of the results and the applicability for regulatory agencies [17].

Meanwhile, computational approaches are emerging as a practical solution for the quick and inexpensive evaluation of chemical substances [18]. The main question relates to their inherent lack of reliability (due to the high complexity of the underlying biological mechanisms and the non-negligible experimental variability of toxicity endpoints) [19]. Many Quantitative Structure-Activity Relationship (QSAR) models discriminating hERG blockers from non-blockers were reported in the scientific literature (see Table 1 for the overview).

Although these models appear to be well-fitted, a critical analysis reveals that the vast majority of the published QSAR models do not comply with the standard validation procedures and the different statistical criteria described in the best practices of QSAR modeling [80, 81]. Most of those models are indeed not compliant with the OECD guidance on QSAR model development and validation [82]. More specifically, the primary drawbacks of the majority of published QSAR studies are: (i) most models do not have proof of passing the Y-randomization test [21, 23, 26, 28, 29, 31–35, 38, 40, 41, 45–49, 51–56, 58, 59, 63–65, 68–70, 75, 79]; (ii) no proof of applicability domain (AD) estimation is provided [21, 23, 26–29, 31–36, 40, 45, 48–53, 56, 58, 63–65, 68–71, 75, 79]; and (iii) model predictivity is not acceptable [39, 61, 66]. As a consequence, despite the large number of QSAR models for hERG blockage available in the literature, only very few models can actually be employed to predict hERG blockage [60, 61, 74, 78]. Most of the models and associated datasets used to build them are not available online for the scientific community. These major drawbacks compromise the practical use of prior models for reliable assessment of drug-induced QT syndrome.

Given the risks associated with hERG inhibition and the lack of reliable models freely available for the research community, we aimed to build predictive and well-characterized QSAR models for hERG blockage using the largest publicly available dataset for hERG blockage. In this study, we developed several consensus QSAR models combining different descriptor types and machine learning techniques (Combi-QSAR), all validated using a modeling workflow fully compliant with OECD guidelines. Moreover, we have applied these models to the World Drug Index (WDI) database for assessing whether some putative hERG blockers and non-blockers among marketed drugs and drug candidates could be identified.

## MATERIALS AND METHODS

### Data preparation

**hERG modeling set**—We retrieved 11,958 chemical records containing affinity and inhibition data for the hERG channel from ChEMBL [83] v13 database (March, 2013).

Only the records related to the potency and the affinity values reported in activity as  $IC_{50}$ ,  $K_i$ , and  $EC_{50}$ , were retained. Subsequently, all concentrations were converted to  $-\log(M)$  values. Compounds with multiple hERG measurements were identified during analyses of duplicates (see Data curation section). Because this dataset was composed from measurements done by multiple laboratories and different types of assays, the binary hERG blockade potential for duplicated records was analyzed to verify the dataset consistency as well as inter- and intra-laboratory assay variability. Different threshold levels have been proposed in the literature; for this reason we have used three binary classification thresholds (1  $\mu M$ , 10  $\mu M$ , and 20  $\mu M$ ) to discriminate between hERG blockers and non-blockers. Importantly, we have found an overall concordance between duplicates, considering multiple assays, as high as 93.61%, 90.73%, and 90.19% for the three aforementioned thresholds respectively. Given the high concordance between multiple assays for the same compound, we decided to merge the data. Original references were verified to guarantee that biological activities were correct in ChEMBL database and adjusted if needed. We have

noted that compounds from five publications [73, 84–87] had their potencies wrongly transcribed from original sources to the ChEMBL database. Moreover, compounds with undetermined activities (e.g., >20  $\mu\text{M}$ ; <1  $\mu\text{M}$ ; etc.) were kept only if they fit the class discrimination threshold. Finally, the datasets were divided into modeling sets (80%) and test sets (20%) using the modified Kennard-Stone algorithm (<http://labmol.farmacia.ufg.br/qsar>).

**External validation set**—Additional chemical data for 561 compounds were retrieved from the hERG study published by Li et al. [48]. After curating the data, 553 compounds were retained. Subsequently, the overlap between this collection and the hERG dataset generated from the ChEMBL database was determined. There were 174 compounds that were present in both datasets, and only nine divergent binary hERG annotations were identified (94.8% of agreement), demonstrating the strong consistency for this dataset. The remaining 379 compounds that were absent from the modeling set were thus utilized for external validation of the QSAR models (see Results).

**WOMBAT-PK dataset**—As an additional external validation set, the performance of the models developed in this study were compared to the ones from Li et al. [48]. The authors used 66 compounds with reported hERG activity from the WOMBAT-PK database [88]. These data originated from different sources with experimental binding activities evaluated in mammalian and non-mammalian cell lines, and were expressed in  $\text{IC}_{50}$ ,  $\text{K}_i$ , or percentage of current inhibition [89–93].

**WDI dataset**—The WDI dataset (version 2010, <http://thomsonreuters.com/world-drug-index/>) involved almost 53,965 chemical compounds and pharmacologically active compounds, including all marketed drugs and compounds that entered clinical trials.

## Data curation

All aforementioned chemical datasets were carefully curated and standardized according to the protocol proposed by Fourches et al. [94]. Structural normalization of specific chemotypes, such as aromatic and nitro groups, was performed using ChemAxon Standardizer (v. 6.1, ChemAxon, Budapest, Hungary, <http://www.chemaxon.com>). Inorganic salts, organometallic compounds, polymers, and mixtures were removed. Duplicates, i.e., identical compounds reported several times in the dataset, were identified using ISIDA/Duplicates software [95] and analyzed. If the experimental hERG data varied from different sources for a given compound, it was removed.

## Cheminformatics approaches

**Dataset diversity analysis**—The Sequential Agglomerative Hierarchical Non-overlapping (SAHN) method implemented in the ISIDA/Cluster software (<http://fourches.web.unc.edu>) was applied to check the dataset structure diversity [95]. In this method, sub-structural molecular fragments (SMF) [96] are used as input for Euclidean distance calculation. Each compound is initially treated as one cluster. The algorithm proceeds by merging the  $n$  compounds sequentially into clusters using their pairwise Euclidean distances. New clusters are formed by the merger of existent clusters with the

most similar clusters at each stage, whereas the distance matrix is updated with the distances between the newly formed cluster and all the other ones, according to the type of linkage specified by the user (complete linkage was used in this study). The process continues until one cluster remains. The software generates a dendrogram of the parent-child relationships between the clusters and a heat map of the proximity matrix colored according to the pairwise chemical similarity between compounds.

**Molecular descriptors**—Four different types of molecular fingerprints reflecting the absence (0) or presence (1) of substructural fragment for each compound [97] were utilized in this study.

**The Molecular ACCess System (MACCS) structural keys:** The MACCS structural keys were calculated using the RDKit (<http://www.rdkit.org>) in the KNIME platform [98]. The MACCS structural keys [99] are a collection of 166 predefined substructures associated with a SMARTS pattern and belonging to the dictionary-based fingerprint class. They were first planned for substructure searches and typically show a low performance level for virtual screening; thus, they are often used as a baseline fingerprint for benchmarking studies.

**FeatMorgan:** FeatMorgan fingerprints are circular fingerprints based on the Morgan algorithm and feature invariants (FCFP-like) [100, 101]. They combine the RDKit Morgan fingerprint algorithm with pharmacophoric features calculated using “better” feature definitions. A pharmacophore is the ensemble of steric and electronic features essential for interaction with the biological target and responsible for biological activity [102], FCFPs are circular topological fingerprints where each pharmacophore represents a bit at the start. A number of iterations are performed to combine the initial pharmacophore identifiers with identifiers of neighboring pharmacophores until a specified diameter is reached and counted. The FCFP rule is derived from pharmacophore feature definitions (e.g., donor, acceptor, aromatic, halogen, basic, acidic, etc.) of the atoms in a molecule (<http://www.rdkit.org/docs/index.html>).

**Pharmacophore fingerprints:** The pharmacophore fingerprints were calculated using the JChem suite from ChemAxon (v. 6.1.3), ChemAxon, Budapest, Hungary, (<http://www.chemaxon.com>) in the KNIME platform. The 2D pharmacophore fingerprints account for the pharmacophore properties of each atom, and the collection of all atom-atom pharmacophore feature pairs, along with their topological distances. More details are available at [www.chemaxon.com/jchem/doc/user/PFp2D.html](http://www.chemaxon.com/jchem/doc/user/PFp2D.html).

**PubChem fingerprints:** The PubChem fingerprints were calculated using the Chemistry Development Kit (CDK) [103] in the KNIME platform. PubChem fingerprints consist of an 881-dimensional vector of bits that accounts for the absence (0) or presence (1) of a substructure (fragment) for each compound. The 2D chemical representation of compounds is based on specific elements, types of ring systems, atom pairing, or atomic environment (nearest neighbors), etc. A detailed description of this fingerprint system is available at [ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem\\_fingerprints.txt](ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem_fingerprints.txt).

**QSAR modeling**—The QSAR modeling workflow was carefully conducted in three major steps [81, 104]: (i) data curation/preparation/analysis (selection of compounds and descriptors), (ii) model building, and (iii) model validation/selection. First, each dataset was randomly divided into a modeling set (80%) and a test set (20%) using the modified Kennard-Stone algorithm implemented in qsaR package v. 0.7 made for R, available at our lab group webpage (<http://labmol.farmacia.ufg.br/qsar>). Five-fold cross-validation procedure was implemented for model generation. The modeling set with known experimental activity was randomly divided into five subsets; subsequently, one subset (20% of the compounds) was selected as a test set, while the other four subsets (80% of the compounds) were merged as a training set. This procedure was repeated with the other subsets, allowing each of the five subsets to be used once as a test set. The 5-fold external cross-validation procedure was repeated three times and the predictions were averaged. Although models were generated only using the training set, model selection depended on the performance of both the training and test sets, because training set accuracy alone is insufficient to establish robust and externally predictive models [80]. After model selection, the external test set was screened in order to evaluate the actual predictivity of the model. In addition, 10 rounds of Y-randomization were performed for each dataset to assure that the accuracy of the models was not obtained due to chance correlations. The applicability domain (AD) for each descriptor type was estimated based on the Euclidean distances among the training sets of each model generated in the 5-fold cross-validation procedure. The distance of a test compound to its nearest neighbor in the training set was compared to the predefined applicability domain threshold level. If the distance was greater than this threshold level, the prediction was considered to be less trustworthy [105]. Four different machine learning methods, including the support vector machine (SVM) method with a radial basis kernel function (SVMradii) [106], the random forest (RF) method [107], the tree bagging method, and the gradient boosting method (GBM) were used for model building. The models were built using the qsaR package and its integration workflow plan for KNIME 2.9. All these procedures were united in publicly available KSAR workflow (<http://labmol.farmacia.ufg.br/ksar>) used in this study. KSAR workflow is tightly integrated with R and KNIME and includes many modules, such as the module for curating the data (e.g., removal of duplicates), the rational module (Kennard-Stone and modified Kennard-Stone algorithm), and the random dataset splits module, multiple machine learning methods, performance metrics to evaluate 5-fold cross-validation and external evaluation, the applicability domain (AD), the Y-randomization test, and many other utilities.

**SVM method:** The SVM method is a general data modeling methodology first developed by Vapnik [106]. Briefly, a hyperplane in a high-dimensional feature space is built based on molecular descriptors using kernel functions; subsequently, a linear or non-linear model is constructed in this feature space to segregate compounds with different activities. In this study, a radial basis kernel function (SVMradii) was chosen to seek the optimal pair of the penalty parameter C and the kernel parameter  $\gamma$ .

**RF method:** The complete description of the original RF algorithm can be found elsewhere [107]. The RF method is an ensemble learning method in which single decision trees are built, and the final prediction is defined by all tree outputs. In each tree, 1/3 of the training

set is randomly extracted (i.e., bootstrap sample) and used as an out-of-bag (OOB) set, while the remaining 2/3 of the training set is used for model building. The best split generated by the CART algorithm [108], among the  $m$  randomly selected descriptors from the entire pool in each node, is chosen. Then, each tree is grown to the largest possible extent without pruning. The OOB set is used as a test set for the current tree. The predicted classification values are defined by majority voting for one of the classes.

**Tree bagging method:** The tree bagging method averages the decision tree over many samples extracted from the modeling set by the bootstrap replicate. The same compound may appear multiple times in the bootstrap replicate, or it may not appear at all. Thus, on each of  $n$  rounds of bagging, a bootstrap replicate is created from the original training set. A base classifier is then trained on this replicate, and the process continues. After  $n$  rounds, a final combined classifier is formed by the majority vote of all of the base classifiers [109].

**GBM:** The GBM generates models by computing a sequence of trees, in which each successive tree is built from the prediction residuals of the preceding tree. A simple (best) partitioning of the data is determined at each step in the boosting tree algorithm, and the deviations of the observed values from the respective residuals for each partition are computed. Given the preceding sequence of trees, the next 3-node tree will then be fitted to the residuals in order to find another partition that will further reduce the residual (error) variance for the data [110].

**Evaluation of prediction performance—**The following metrics (Equations 1–8) were used to assess different aspects of model performance:

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad \text{Equation (1)}$$

$$BAC = \frac{Sensitivity+Specificity}{2} \quad \text{Equation (2)}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad \text{Equation (3)}$$

$$Sensitivity (recall) = \frac{TP}{TP+FN} \quad \text{Equation (4)}$$

$$Specificity = \frac{TN}{TN+FP} \quad \text{Equation (5)}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{Equation (6)}$$

$$F1 = 2 \frac{Precision \times Recall}{Precision + Recall} \quad \text{Equation (7)}$$

$$AUC = \sum_i [(Sensitivity_{i+1})(Specificity_{i+1} - Specificity_i)] \quad \text{Equation (8)}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

## RESULTS AND DISCUSSION

The largest publicly available dataset for hERG liability retrieved from the ChEMBL database (<https://www.ebi.ac.uk/chembl/>) contained 11,958 associated bioactivity records for the hERG K<sup>+</sup> channel. Once curated, this dataset only contained 4,980 compounds for modeling. Threshold values for the blocker/non-blocker classification vary in the literature from 1  $\mu\text{M}$  to 40  $\mu\text{M}$  [56, 66, 69, 111]. For this reason, binary classification models were built for three different thresholds: 1  $\mu\text{M}$ , 10  $\mu\text{M}$ , and 20  $\mu\text{M}$ . Therefore, three datasets were derived from the original dataset and divided accordingly: 4,938 compounds met the threshold level of 1  $\mu\text{M}$ , 4,833 compounds met the threshold level of 10  $\mu\text{M}$ , and 4,544 compounds met the threshold level of 20  $\mu\text{M}$  (Table 2). Models were generated separately for each of the three different threshold levels to define the most suitable cutoff for discriminating between hERG blockers and non-blockers.

The examination of the 4,980 compounds suggested a high level of structural dissimilarity (dendrogram and heat map are shown on line at <http://labmol.farmacia.ufg.br/predherg>).

The 5-fold cross-validation procedure was used to estimate the robustness of the models developed. The test set was applied to validate and to estimate the predictive power of the models. In this work, we have chosen the models generated for the threshold level of 10  $\mu\text{M}$ , which showed the best performance and were validated internally and externally. The statistical results of generated QSAR models for the modeling set of the 10  $\mu\text{M}$  threshold level are summarized in Table 3. The detailed results for this threshold level, as well as the full results for the threshold levels of 1  $\mu\text{M}$  and 20  $\mu\text{M}$ , are available in Tables S1–S9 (Supplementary Material). The combination of different descriptors and machine learning methods led to robust and predictive QSAR models, with balanced accuracy (BAC) values ranging between 0.74–0.87 and a coverage of 0.77–0.93 (Table 3).

The best individual model was generated using the combination of featMorgan fingerprints with SVM (BAC = 78%; sensitivity = 86%; specificity = 69%; see Table 3 for more details).

To assure that the accuracy of the models was not due to chance correlation, 10 rounds of Y-randomization were performed for each dataset. The results are shown in Tables S3, S5 and S9 (Supplementary Material).

Several QSAR models were generated using multiple machine learning algorithms and descriptors. Consensus QSAR modeling, *i.e.*, parallel development of multiple QSAR models using all pairwise combinations of different types of chemical descriptors and various machine learning techniques over single QSAR modes, has been shown to be advantageous [112, 113]. Nevertheless, no need exists of the overabundance of models in the consensus ensemble [94]. Therefore, a verification procedure was conducted to indicate whether a consensus model, based on models from Table 3, would offer additional advantages compared to the individual models. The consensus model was built by averaging the predicted values from the individual model for each machine learning technique that



yielded the best performance with higher coverage. Consensus model considered only compounds that were predicted identically (with AD taken into account) by all the models. For example, if models 1, 2, and 3 predicted a compound to be a blocker, and this compound is inside the AD for these models, but model 4 predicted this compound to be a non-blocker, and the compound is outside the AD for this model, in the consensus model this compound was still classified as a blocker. However, if the compound is predicted to be a blocker, and it is inside the AD for models 1, 2, and 3, but the compound is predicted to be a non-blocker by model 4 and is inside the AD, the final prediction of the consensus is specified as inconclusive. In another situation, if all four models, independently of the outcome, yielded predictions outside the AD, the result for this compound was classified as unreliable.

Thus, the consensus model was built by combining the SVM model with MACCS fingerprints (MACCS-SVM), the Tree Bagging method with PubChem Fingerprints (PubChem-TreeBag), the RF model with the ChemAxon Pharmacophore Fingerprint (Pharm. FP-RF), and the gradient boosting model (GBM) with the featMorgan fingerprint (featMorgan-GBM). The generated consensus model demonstrated BAC of 83%, sensitivity of 85%, specificity of 81%, and coverage of 74% (see Table 3 for more details). Therefore, the consensus model discriminates well between hERG blockers and nonblockers—better than any of the individual models.

More rigorous consensus model was also developed (consensus rigor), by combining the same models as in the consensus model with more restrictive conditions. The consensus rigor model only considered the outcome to be reliable when a compound was inside the AD for the four models, and all of the predictions were equal. Any non-concordant prediction was specified as inconclusive. If the compound was outside the AD for any model, then the outcome was specified as unreliable. Expectedly, the increase prediction performance of consensus rigor model (BAC = 87%, sensitivity = 89%, specificity = 84%; see Table 3) was achieved at the expense of coverage (34%). Although consensus rigor model is very accurate predictor, its applicability is limited only for certain chemical classes.

In summary, the consensus model demonstrated better results for 5-fold external CV, with 5% accuracy and 20% sensitivity increase when compared with the best individual model (featMorgan-SVM). The statistical results for the external test set at the 10  $\mu\text{M}$  threshold level are summarized in Table 4. The complete results are shown in Table S6 (Supplementary Material). Consensus model demonstrated the best performance among all other individual models (BAC of 91%, sensitivity of 89%, specificity of 93%, and coverage of 78%).

Li et al. [48] compiled a dataset of 561 compounds with chemical data for hERG activity. After exclusion of duplicates with our modeling set, 377 unique compounds were retained at the threshold level of 10  $\mu\text{M}$  and used as an additional external validation set. Our consensus model reached BAC of 95%, sensitivity of 91%, specificity of 99%, and 84% of coverage for this additional external validation set.

Moreover, Li et al. [48] used an additional evaluation set comprising 66 compounds from the WOMBAT-PK database with reported hERG activity to validate the performance of

their models. Therefore, this external set of 66 compounds from the WOMBAT-PK database was used to examine validation of the consensus model by comparing with the results from Li et al. [48] under the same conditions. The comparison of the statistical values is shown in Table 5. The consensus model outperformed the Li et al. [48] models in specificity, sensitivity, and in other performance metrics, which reflected a higher BAC (~ 0.27) for the models developed in this research. Moreover, the consensus model for the 10  $\mu$ M threshold level presented a BAC of 98% (which represents 1 compound misclassified out of 59), sensitivity of 98%, specificity 99%, and coverage of 89% (Table 5).

Czodrowski [71] performed an analysis of the hERG dataset retrieved from ChEMBL in which four classification models were built with different divisions of the dataset using the RDKit descriptors and the RF model. To the best of our knowledge, this is the only study found to use the hERG dataset content in the ChEMBL database. Initially we wanted to compare the models obtained by Czodrowski with the ones developed in our study. But then we have found that Czodrowski [71] did not calculate the AD for the models that allowed prediction of 100% of compounds but compromises their practical use. Consensus models developed in this study has AD estimation that reduced the coverage. We wanted to use the same set of compounds for fair comparison, but unfortunately predicted values were not reported in the study [71] that does not allow us to make direct comparison.

QSAR models were developed as virtual screening tools for revealing putative hERG blockers among marketed drugs and those in development using the WDI database for a case study. A total of 179 compounds were present in both the hERG and the WDI datasets: 103 blockers and 76 nonblockers. After the data curation 44,486 remaining unique compounds were predicted by consensus model developed in this research for revealing putative hERG blockers and non-blockers. 4,945 compounds were predicted to be blockers and 20,871 – to be non-blockers (the remaining compounds 18,670 were outside of the AD). All the compounds and corresponding predictions are available in the supplementary material and on-line (<http://labmol.farmacina.ufg.br/predherg/vs-wdi.pdf>).

Model interpretation revealed several SAR rules, which can guide structural optimization of hERG blockers into non-blockers. Figures 1 – 2 show some revealed SAR rules, involving changes in the amine nitrogen environment, adding oxygen atom, removing carbon atoms, aromatic substitutions, transformations involving some descriptors such as the difference between the topological polar surface area (TPSA) [114] of the two molecules in the transformation and the Labute's approximate surface area descriptor (Labute ASA) [115].

The general transformations in Fig. 1 show some changes in the environment of the amine nitrogen can reduce hERG inhibition. In Fig. 1A, we can see that removing carbons and/or changing the electronic environment around the basic nitrogen can result in a reduction in hERG inhibition. In this example, the modification of the pyrrolidine moiety by removing carbon atoms or changing it to another functionalized ring (in this example a morpholine ring), yielded in the reduction of the hERG binding. Furthermore, the next two transformations (Fig. 1B and 1C) show the same SAR rules that remove carbon atoms, reduce lipophilicity and/or change the electronic and steric environment around the basic

nitrogen can transform a potent hERG blocker to less potent blocker or even to a non-blocker compound. Some of those observations were also found previously [48, 116].

We can also observe that transformations that add a hydroxyl group reduce hERG inhibition (Fig. 2A). As already mentioned, removing carbon atoms, as well as reducing the lipophilicity, can result in a reduction in hERG binding (Fig. 2B).

We also noticed some SAR rules revealing specific structural changes through descriptors, like the topological polar surface area (TPSA) of the compounds [114]. If the difference between the descriptor TPSA ( - TPSA) of two compounds involved in the transformation is equal or greater to 60, this can result in reduction in hERG inhibition (Fig. 2C). We have found in our modeling set that 50 compounds follow this SAR rule and only 3 compounds do not follow this rule. Another descriptor observed and related with changes in hERG binding potency is the Labute's approximate surface area (Labute ASA) [115]. If the calculated Labute ASA descriptor is between 309 and 337, then the compound is frequently a hERG blocker (Fig. 2D). 130 compounds in our modeling set followed this SAR rule and 3 compounds do not follow this rule.

Importantly, our QSAR models were also capable of recognizing modifications that do not follow the general SAR rules, as shown in Fig. 3. As we can see, some bioisosteric replacements have resulted in dramatic changes in activity. For example, the replacement of a furane ring by a tetrazole ring, which is a bioisosteric replacement and therefore should preserve the activity, resulted in a substantial alteration in hERG binding, changing the compound from a blocker to non-blocker (Fig. 3A, left). The same is observed with the substitution of benzene to pyridine ring (Fig. 3A, right). The bioisosteric replacement of aromatic rings in our modeling set had 169 examples that follow the SAR rule, as the bioisosteric replacement did not alter the activity. However, there were 21 examples in which this modification had altered dramatically the activity, changing from a blocker to non-blocker compound, and our model could capture such modifications. These cases represent the activity cliffs, i.e., structurally similar compounds with large differences in potency [117]. The modification of a chlorine to hydroxyl group in an aromatic ring also reduced dramatically the binding to hERG (Fig. 3B, left). Although these groups are classic bioisosteres, this transformation involved the introduction of a hydroxyl group in an aromatic ring that alters the electronic environment in the aromatic group. The same is observed with the substitution of a chlorine by a nitrile group (Fig. 3B, right), leading to a notable change in hERG binding, changing from a blocker compound to a non-blocker compound. The following examples in Fig. 3C and 3D are also activity cliffs.

In general, our SAR rules showed that to decrease toxicity of hERG blockers one should consider decreasing their lipophilicity, removing carbons and/or changing the electronic environment around the basic nitrogen, and increasing the topological polar surface area. It is important to note that our observations also indicated that hERG inhibition has complex structure-activity relationship as subtle changes in the structure often result in small changes in activity. Moreover, we observed a considerable number of activity cliffs in hERG dataset, most of them span a potency difference of at least 2 orders of magnitude. The final model

showed the ability to predict such transformations and the threshold effect for compounds near to the border regions.

## CONCLUSIONS

Poor pharmacokinetics and toxicity are important causes of costly late-stage failures in drug development. Our laboratory has been working to overcome or reduce these failures, developing *in silico* tools to early predict and optimize some properties, such as metabolism [118–125], mutagenicity (Ames test), Caco-2 permeability, blood-brain barrier penetration (BBBP), and water solubility [126], skin sensitization, skin permeability, among others [127–131].

We have developed statistically significant and externally predictive QSAR models of hERG blockage. The best model was obtained for the 10  $\mu\text{M}$  threshold using the largest publicly available dataset of structurally diverse compounds including variety of drug classes. Consensus modeling by merging models developed with different sets of descriptors increased the balanced accuracy, sensitivity, and specificity of the models up to 81–85% with the coverage of ~75%. The models developed in this study can be used by the research community and regulatory scientists for the rapid evaluation of cardiac toxicity liability via hERG inhibition in chemical inventories. For instance, we applied our models for the virtual screening of the WDI dataset and identified 4,945 potential hERG blockers that may be candidates for targeted testing to determine hERG liability. As a result of our study, all curated datasets and developed models that can be used for the rapid identification of hERG blockers and non-blockers in the context of virtual screening for drug development, have been made publicly available at the LabMol (<http://labmol.farmacia.ufg.br/predherg>) and Chembench web-portals.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to thank Brazilian Funding Agencies CAPES, CNPq and FAPEG for financial support and fellowships. We are also grateful to ChemAxon for providing us with academic licenses for their software. DF, EM, and AT thank NIH (grants GM66940 and GM096967) and EPA (grant RD 83499901) for financial support.

## References

1. Vandenberg JI, Perry MD, Perrin MJ, Mann SA, Ke Y, Hill AP. hERG K<sup>+</sup> Channels: Structure, Function, and Clinical Significance. *Physiol Rev.* 2012; 92:1393–1478. [PubMed: 22988594]
2. Brown AM. Drugs, hERG and sudden death. *Cell Calcium.* 2004; 35:543–547. [PubMed: 15110144]
3. Woosley RL. Cardiac actions of antihistamines. *Annu Rev Pharmacol Toxicol.* 1996; 36:233–252. [PubMed: 8725389]
4. Rampe D, Roy ML, Dennis A, Brown AM. A mechanism for the proarrhythmic effects of cisapride (Propulsid): high affinity blockade of the human cardiac potassium channel HERG. *FEBS Lett.* 1997; 417:28–32. [PubMed: 9395068]
5. Alvarez PA, Pahissa J. QT alterations in psychopharmacology: proven candidates and suspects. *Curr Drug Saf.* 2010; 5:97–104. [PubMed: 20210726]

6. Roden DM. Drug-induced prolongation of the QT interval. *N Engl J Med*. 2004; 350:1013–1022. [PubMed: 14999113]
7. Mitcheson JS, Chen J, Lin M, Culberson C, Sanguinetti MC. A structural basis for drug-induced long QT syndrome. *Proc Natl Acad Sci U S A*. 2000; 97:12329–12333. [PubMed: 11005845]
8. Picard S, Goineau S, Guillaume P, Henry J, Hanouz JL, Rouet R. Supplemental studies for cardiovascular risk assessment in safety pharmacology: a critical overview. *Cardiovasc Toxicol*. 2011; 11:285–307. [PubMed: 21805209]
9. FDA. Guidance for industry. S7B nonclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals. Rockville, MD: 2005. p. 1-13.
10. FDA. E14 clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. Rockville, MD: 2005. p. 1-20.
11. Polonchuk L. Toward a New Gold Standard for Early Safety: Automated Temperature-Controlled hERG Test on the PatchLiner. *Front Pharmacol*. 2012; 3:3. [PubMed: 22303293]
12. Kiss L, Bennett PB, Uebele VN, Koblan KS, Kane SA, Neagle B, Schroeder K. High throughput ion-channel pharmacology: planar-array-based voltage clamp. *Assay Drug Dev Technol*. 2003; 1:127–135. [PubMed: 15090139]
13. Wen D, Liu A, Chen F, Yang J, Dai R. Validation of visualized transgenic zebrafish as a high throughput model to assay bradycardia related cardio toxicity risk candidates. *J Appl Toxicol*. 2012; 32:834–842. [PubMed: 22744888]
14. Hamill OP, Marty A, Neher E, Sakmann B, Sigworth FJ. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflugers Arch*. 1981; 391:85–100. [PubMed: 6270629]
15. Wi niowska B, Polak S. hERG in vitro interchange factors--development and verification. *Toxicol Mech Methods*. 2009; 19:278–284. [PubMed: 19778217]
16. Witchel HJ, Milnes JT, Mitcheson JS, Hancox JC. Troubleshooting problems with in vitro screening of drugs for QT interval prolongation using HERG K<sup>+</sup> channels expressed in mammalian cell lines and *Xenopus* oocytes. *J Pharmacol Toxicol Methods*. 2003; 48:65–80. [PubMed: 14565563]
17. Elkins RC, Davies MR, Brough SJ, Gavaghan DJ, Cui Y, Abi-Gerges N, Mirams GR. Variability in high-throughput ion-channel screening data and consequences for cardiac safety assessment. *J Pharmacol Toxicol Methods*. 2013; 68:112–122. [PubMed: 23651875]
18. Raunio H. In silico toxicology - non-testing methods. *Front Pharmacol*. 2011; 2:33. [PubMed: 21772821]
19. Gleeson MP, Modi S, Bender A, Robinson RLM, Kirchmair J, Promkatkaew M, Hannongbua S, Glen RC. The challenges involved in modeling toxicity data in silico: a review. *Curr Pharm Des*. 2012; 18:1266–1291. [PubMed: 22316153]
20. De Ponti F, Poluzzi E, Montanaro N. Organising evidence on QT prolongation and occurrence of Torsades de Pointes with non-antiarrhythmic drugs: a call for consensus. *Eur J Clin Pharmacol*. 2001; 57:185–209. [PubMed: 11497335]
21. Cavalli A, Poluzzi E, De Ponti F, Recanatini M, De Ponti F. Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K(+) channel blockers. *J Med Chem*. 2002; 45:3844–3853. [PubMed: 12190308]
22. Keating MT, Sanguinetti MC. Molecular and cellular mechanisms of cardiac arrhythmias. *Cell*. 2001; 104:569–580. [PubMed: 11239413]
23. Pearlstein RA, Vaz RJ, Kang J, Chen XL, Preobrazhenskaya M, Shchekotikhin AE, Korolev AM, Lysenkova LN, Miroshnikova OV, Hendrix J, Rampe D. Characterization of HERG potassium channel inhibition using CoMSiA 3D QSAR and homology modeling approaches. *Bioorg Med Chem Lett*. 2003; 13:1829–1835. [PubMed: 12729675]
24. Roche O, Trube G, Zuegge J, Pflimlin P, Alanine A, Schneider G. A virtual screening method for prediction of the HERG potassium channel liability of compound libraries. *Chembiochem*. 2002; 3:455–459. [PubMed: 12007180]
25. Keserü GM. Prediction of hERG potassium channel affinity by traditional and hologram qSAR methods. *Bioorg Med Chem Lett*. 2003; 13:2773–2775. [PubMed: 12873512]

26. Bains W, Basman A, White C. HERG binding specificity and binding site structure: evidence from a fragment-based evolutionary computing SAR study. *Prog Biophys Mol Biol.* 2004; 86:205–233. [PubMed: 15288759]
27. Yap CW, Cai CZ, Xue Y, Chen YZ. Prediction of torsade-causing potential of drugs by support vector machine approach. *Toxicol Sci.* 2004; 79:170–177. [PubMed: 14976348]
28. Tobita M, Nishikawa T, Nagashima R. A discriminant model constructed by the support vector machine method for HERG potassium channel inhibitors. *Bioorg Med Chem Lett.* 2005; 15:2886–2890. [PubMed: 15911273]
29. O'Brien SE, de Groot MJ. Greater than the sum of its parts: combining models for useful ADMET prediction. *J Med Chem.* 2005; 48:1287–1291. [PubMed: 15715500]
30. Kang J, Wang L, Cai F, Rampe D. High affinity blockade of the HERG cardiac K(+) channel by the neuroleptic pimozone. *Eur J Pharmacol.* 2000; 392:137–140. [PubMed: 10762666]
31. Cianchetta G, Li Y, Kang J, Rampe D, Fravolini A, Cruciani G, Vaz RJ. Predictive models for hERG potassium channel blockers. *Bioorg Med Chem Lett.* 2005; 15:3637–3642. [PubMed: 15978804]
32. Coi A, Massarelli I, Murgia L, Saraceno M, Calderone V, Bianucci AM. Prediction of hERG potassium channel affinity by the CODESSA approach. *Bioorg Med Chem.* 2006; 14:3153–3159. [PubMed: 16426850]
33. Ekins S, Balakin KV, Savchuk N, Ivanenkov Y. Insights for human ether-a-go-go-related gene potassium channel inhibition using recursive partitioning and Kohonen and Sammon mapping techniques. *J Med Chem.* 2006; 49:5059–5071. [PubMed: 16913696]
34. Gepp MM, Hutter MC. Determination of hERG channel blockers using a decision tree. *Bioorg Med Chem.* 2006; 14:5325–5332. [PubMed: 16616507]
35. Seierstad M, Agrafiotis DK. A QSAR model of HERG binding using a large, diverse, and internally consistent training set. *Chem Biol Drug Des.* 2006; 67:284–296. [PubMed: 16629826]
36. Song M, Clark M. Development and evaluation of an in silico model for hERG binding. *J Chem Inf Model.* 2006; 46:392–400. [PubMed: 16426073]
37. Sun H. An accurate and interpretable bayesian classification model for prediction of HERG liability. *Chem Med Chem.* 2006; 1:315–322. [PubMed: 16892366]
38. Dubus E, Ijjaali I, Petitot F, Michel A. In silico classification of HERG channel blockers: a knowledge-based strategy. *Chem Med Chem.* 2006; 1:622–630. [PubMed: 16892402]
39. Gavaghan CL, Arnby CH, Blomberg N, Strandlund G, Boyer S. Development, interpretation and temporal evaluation of a global QSAR of hERG electrophysiology screening data. *J Comput Aided Mol Des.* 2007; 21:189–206. [PubMed: 17384921]
40. Leong MK. A novel approach using pharmacophore ensemble/support vector machine (PhE/SVM) for prediction of hERG liability. *Chem Res Toxicol.* 2007; 20:217–226. [PubMed: 17261034]
41. Obrezanova O, Csanyi G, Gola JMR, Segall MD. Gaussian processes: a method for automatic QSAR modeling of ADME properties. *J Chem Inf Model.* 2007; 47:1847–1857. [PubMed: 17602549]
42. Kramer C, Beck B, Kriegl JM, Clark T. A composite model for HERG blockade. *Chem Med Chem.* 2008; 3:254–265. [PubMed: 18061919]
43. Filz O, Lagunin A, Filimonov D, Poroikov V. Computer-aided prediction of QT-prolongation. *SAR QSAR Environ Res.* 2008; 19:81–90. [PubMed: 18311636]
44. Garg D, Gandhi T, Gopi Mohan C. Exploring QSTR and toxicophore of hERG K+ channel blockers using GFA and HypoGen techniques. *J Mol Graph Model.* 2008; 26:966–976. [PubMed: 17928249]
45. Inanobe A, Kamiya N, Murakami S, Fukunishi Y, Nakamura H, Kurachi Y. In Silico Prediction of the Chemical Block of Human Ether-a-Go-Go-Related Gene (hERG) K+ Current. *J Physiol Sci.* 2008; 58:459–470. [PubMed: 19032804]
46. Thai KM, Ecker GF. A binary QSAR model for classification of hERG potassium channel blockers. *Bioorg Med Chem.* 2008; 16:4107–4119. [PubMed: 18243713]
47. Chekmarev DS, Kholodovych V, Balakin KV, Ivanenkov Y, Ekins S, Welsh WJ. Shape signatures: new descriptors for predicting cardiotoxicity in silico. *Chem Res Toxicol.* 2008; 21:1304–1314. [PubMed: 18461975]

48. Li Q, Jørgensen FS, Oprea T, Brunak S, Taboureau O. hERG classification model based on a combination of support vector machine method and GRIND descriptors. *Mol Pharm.* 2008; 5:117–127. [PubMed: 18197627]
49. Jia L, Sun H. Support vector machines classification of hERG liabilities based on atom types. *Bioorg Med Chem.* 2008; 16:6252–6260. [PubMed: 18448342]
50. Gunturi SB, Archana K, Khandelwal A, Narayanan R. Prediction of hERG Potassium Channel Blockade Using kNN-QSAR and Local Lazy Regression Methods. *QSAR Comb Sci.* 2008; 27:1305–1317.
51. Thai KM, Ecker GF. Similarity-based SIBAR descriptors for classification of chemically diverse hERG blockers. *Mol Divers.* 2009; 13:321–336. [PubMed: 19219559]
52. Fenu LA, Teisman A, De Buck SS, Sinha VK, Gilissen RAHJ, Nijssen MJMA, Mackie CE, Sanderson WE. Cardio-vascular safety beyond hERG: in silico modelling of a guinea pig right atrium assay. *J Comput Aided Mol Des.* 2009; 23:883–895. [PubMed: 19890608]
53. Ermondi G, Visentin S, Caron G. GRIND-based 3D-QSAR and CoMFA to investigate topics dominated by hydrophobic interactions: the case of hERG K<sup>+</sup> channel blockers. *Eur J Med Chem.* 2009; 44:1926–1932. [PubMed: 19110341]
54. Hansen K, Rathke F, Schroeter T, Rast G, Fox T, Kriegl JM, Mika S. Bias-correction of regression models: a case study on hERG inhibition. *J Chem Inf Model.* 2009; 49:1486–1496. [PubMed: 19435326]
55. Nisius B, Göller AH. Similarity-based classifier using topomers to provide a knowledge base for hERG channel inhibition. *J Chem Inf Model.* 2009; 49:247–256. [PubMed: 19434826]
56. Su BH, Shen M, Esposito EX, Hopfinger AJ, Tseng YJ. In silico binary classification QSAR models based on 4D-fingerprints and MOE descriptors for prediction of hERG blockage. *J Chem Inf Model.* 2010; 50:1304–1318. [PubMed: 20565102]
57. Doddareddy MR, Klaasse EC, Shagufta, Ijzerman AP, Bender A. Prospective validation of a comprehensive in silico hERG model and its applications to commercial compound and drug databases. *Chem Med Chem.* 2010; 5:716–729. [PubMed: 20349498]
58. Obiol-Pardo C, Gomis-Tena J, Sanz F, Saiz J, Pastor M. A multiscale simulation system for the prediction of drug-induced cardiotoxicity. *J Chem Inf Model.* 2011; 51:483–492. [PubMed: 21250697]
59. Robinson RLM, Glen RC, Mitchell JBO. Development and Comparison of hERG Blocker Classifiers: Assessment on Different Datasets Yields Markedly Different Results. *Mol Inform.* 2011; 30:443–458.
60. Sinha N, Sen S. Predicting hERG activities of compounds from their 3D structures: development and evaluation of a global descriptors based QSAR model. *Eur J Med Chem.* 2011; 46:618–630. [PubMed: 21185626]
61. Du-Cuny L, Chen L, Zhang S. A critical assessment of combined ligand- and structure-based approaches to hERG channel blocker modeling. *J Chem Inf Model.* 2011; 51:2948–2960. [PubMed: 21902220]
62. Thomson Reuters. IntegritySM. Barcelona: Prous Science, S.A.U a Thomson Reuters business; 2001. Available from: <http://integrity.prous.com>
63. Kim JH, Chae CH, Kang SM, Lee JY, Lee GN, Hwang SH, Kang NS. The Predictive QSAR Model for hERG Inhibitors Using Bayesian and Random Forest Classification Method. *Bull Korean Chem Soc.* 2011; 32:1237–1240.
64. Su BH, Tu Y, Esposito EX, Tseng YJ. Predictive toxicology modeling: protocols for exploring hERG classification and *Tetrahymena pyriformis* end point predictions. *J Chem Inf Model.* 2012; 52:1660–1673. [PubMed: 22642982]
65. Broccatelli F, Mannhold R, Moriconi A, Giuli S, Carosati E. QSAR Modeling and Data Mining Link Torsades de Pointes Risk to the Interplay of Extent of Metabolism, Active Transport, and hERG Liability. *Mol Pharm.* 2012; 9:2290–2301. [PubMed: 22742658]
66. Kar S, Roy K. Prediction of hERG Potassium Channel Blocking Actions Using Combination of Classification and Regression Based Models: A Mixed Descriptors Approach. *Mol Inform.* 2012; 31:879–894.

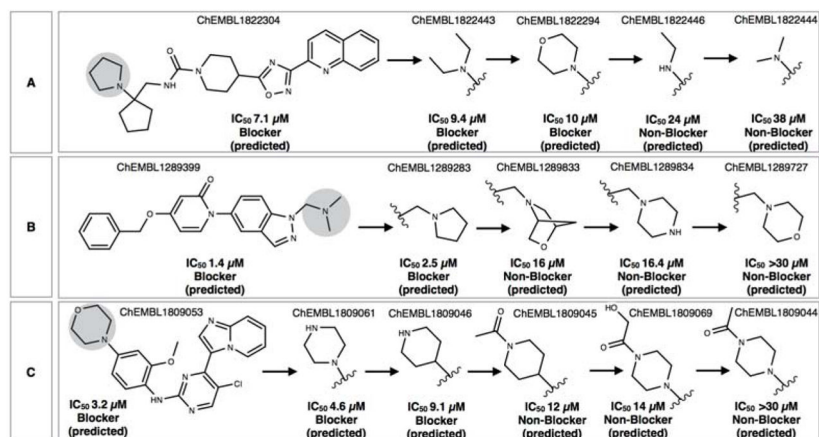
67. Polak S, Wi niowska B, Brandys J. Collation, assessment and analysis of literature in vitro data on hERG receptor blocking potency for subsequent modeling of drugs' cardiotoxic properties. *J Appl Toxicol.* 2009; 29:183–206. [PubMed: 18988205]
68. Tan Y, Chen Y, You Q, Sun H, Li M. Predicting the potency of hERG K<sup>+</sup> channel inhibition by combining 3D-QSAR pharmacophore and 2D-QSAR models. *J Mol Model.* 2012; 18:1023–1036. [PubMed: 21660488]
69. Wang S, Li Y, Wang J, Chen L, Zhang L, Yu H, Hou T. ADMET evaluation in drug discovery. 12. Development of binary classification models for prediction of hERG potassium channel blockage. *Mol Pharm.* 2012; 9:996–1010. [PubMed: 22380484]
70. Wang Z, Mussa HY, Lowe R, Glen RC, Yan A. Probability Based hERG Blocker Classifiers. *Mol Inform.* 2012; 31:679–685.
71. Czodrowski P. hERG Me Out. *J Chem Inf Model.* 2013; 53:2240–2251. [PubMed: 23944269]
72. Brugel TA, Smith RW, Balestra M, Becker C, Daniels T, Koether GM, Throner SR, Panko LM, Brown DG, Liu R, Gordon J, Peters MF. SAR development of a series of 8-azabicyclo[3.2.1]octan-3-yloxy-benzamides as kappa opioid receptor antagonists. Part 2. *Bioorg Med Chem Lett.* 2010; 20:5405–5410. [PubMed: 20719509]
73. Brugel TA, Smith RW, Balestra M, Becker C, Daniels T, Hoerter TN, Koether GM, Throner SR, Panko LM, Folmer JJ, Cacciola J, Hunter AM, Liu R, Edwards PD, Brown DG, Gordon J, Ledonne NC, Pietras M, Schroeder P, Sygowski LA, Hirata LT, Zacco A, Peters MF. Discovery of 8-azabicyclo[3.2.1]octan-3-yloxy-benzamides as selective antagonists of the kappa opioid receptor. Part 1. *Bioorg Med Chem Lett.* 2010; 20:5847–5852. [PubMed: 20727752]
74. Pourbasheer E, Beheshti A, Khajehsharifi H, Ganjali MR, Norouzi P. QSAR study on hERG inhibitory effect of kappa opioid receptor antagonists by linear and non-linear methods. *Med Chem Res.* 2013; 22:4047–4058.
75. Coi A, Bianucci AM. Combining structure- and ligand-based approaches for studies of interactions between different conformations of the hERG K(+) channel pore and known ligands. *J Mol Graph Model.* 2013; 46:93–104. [PubMed: 24185260]
76. Bilodeau MT, Balitza AE, Koester TJ, Manley PJ, Rodman LD, Buser-Doepner C, Coll KE, Fernandes C, Gibbs JB, Heimbrook DC, Huckle WR, Kohl N, Lynch JJ, Mao X, McFall RC, McLoughlin D, Miller-Stein CM, Rickert KW, Sepp-Lorenzino L, Shipman JM, Subramanian R, Thomas KA, Wong BK, Yu S, Hartman GD. Potent N-(1,3-thiazol-2-yl)pyridin-2-amine vascular endothelial growth factor receptor tyrosine kinase inhibitors with excellent pharmacokinetics and low affinity for the hERG ion channel. *J Med Chem.* 2004; 47:6363–6372. [PubMed: 15566305]
77. Berlin M, Lee YJ, Boyce CW, Wang Y, Aslanian R, McCormick KD, Sorota S, Williams SM, West RE, Korfmacher W. Reduction of hERG inhibitory activity in the 4-piperidinyl urea series of H3 antagonists. *Bioorg Med Chem Lett.* 2010; 20:2359–2364. [PubMed: 20188550]
78. Moorthy NSHN, Ramos MJ, Fernandes P. a QSAR and pharmacophore analysis of a series of piperidinyl urea derivatives as HERG blockers and H3 antagonists. *Curr Drug Discov Technol.* 2013; 10:47–58. [PubMed: 22564166]
79. Polak S, Wi niowska B, Glinka A, Fijorek K, Mendyk A. Slow delayed rectifying potassium current (IKs) - analysis of the in vitro inhibition data and predictive model development. *J Appl Toxicol.* 2013; 33:723–739. [PubMed: 22334483]
80. Golbraikh A, Tropsha A. Beware of q<sup>2</sup>! *J Mol Graph Model.* 2002; 20:269–276. [PubMed: 11858635]
81. Tropsha A. Best practices for QSAR model development, validation, and exploitation. *Mol Inform.* 2010; 29:476–488.
82. [accessed April 11, 2013] OECD OECD principles for the validation, for regulatory purposes, of (Quantitative) Structure-Activity Relationship models. <http://www.oecd.org/chemicalsafety/risk-assessment/37849783.pdf>
83. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 2012; 40:D1100–7. [PubMed: 21948594]
84. Haga Y, Mizutani S, Naya A, Kishino H, Iwaasa H, Ito M, Ito J, Moriya M, Sato N, Takenaga N, Ishihara A, Tokita S, Kanatani A, Ohtake N. Discovery of novel phenylpyridone derivatives as



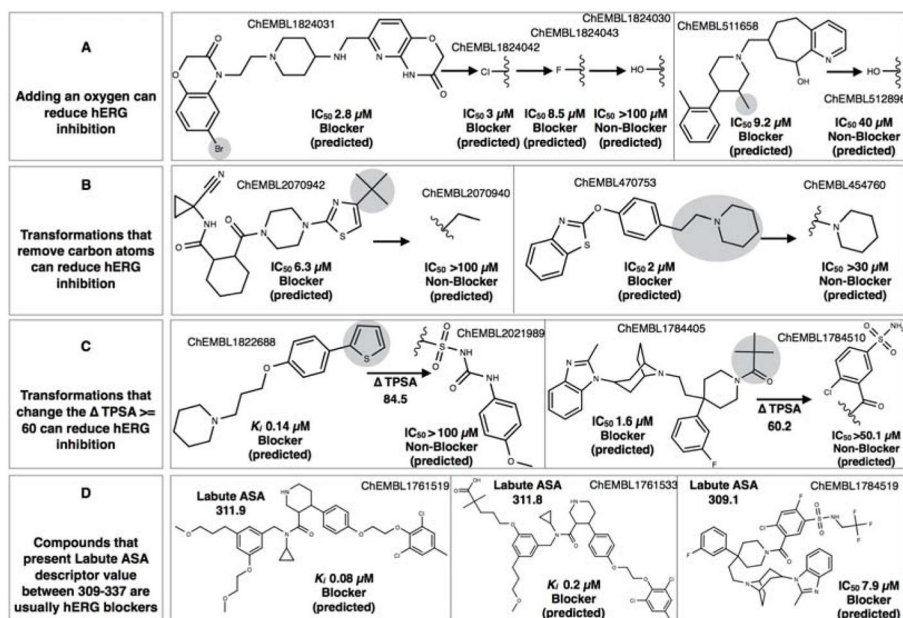
- potent and selective MCH1R antagonists. *Bioorg Med Chem*. 2011; 19:883–893. [PubMed: 21190859]
85. Marquis RW, Lago AM, Callahan JF, Rahman A, Dong X, Stroup GB, Hoffman S, Gowen M, DelMar EG, Van Wagenen BC, Logan S, Shimizu S, Fox J, Nemeth EF, Roethke T, Smith BR, Ward KW, Bhatnagar P. Antagonists of the calcium receptor. 2. Amino alcohol-based parathyroid hormone secretagogues. *J Med Chem*. 2009; 52:6599–6605. [PubMed: 19821575]
86. Xue CB, Feng H, Cao G, Huang T, Glenn J, Anand R, Meloni D, Zhang K, Kong L, Wang A, Zhang Y, Zheng C, Xia M, Chen L, Tanaka H, Han Q, Robinson DJ, Modi D, Storace L, Shao L, Sharief V, Li M, Galya LG, Covington M, Scherle P, Diamond S, Emm T, Yeleswaram S, Contel N, Vaddi K, Newton R, Hollis G, Friedman S, Metcalf B. Discovery of INCB3284, a Potent, Selective, and Orally Bioavailable hCCR2 Antagonist. *ACS Med Chem Lett*. 2011; 2:450–454. [PubMed: 24900329]
87. Lynch JK, Freeman JC, Judd AS, Iyengar R, Mulhern M, Zhao G, Napier JJ, Wodka D, Brodjian S, Dayton BD, Falls D, Ogiela C, Reilly RM, Campbell TJ, Polakowski JS, Hernandez L, Marsh KC, Shapiro R, Knourek-Segel V, Droz B, Bush E, Brune M, Preusser LC, Fryer RM, Reinhart Ga, Houseman K, Diaz G, Mikhail A, Limberis JT, Sham HL, Collins Ca, Kym PR. Optimization of chromone-2-carboxamide melanin concentrating hormone receptor 1 antagonists: assessment of potency, efficacy, and cardiovascular safety. *J Med Chem*. 2006; 49:6569–6584. [PubMed: 17064075]
88. Olah, M.; Rad, R.; Ostopovici, L.; Bora, A.; Hadaruga, N.; Hadaruga, D.; Moldovan, R.; Fuias, A.; Mractc, M.; Oprea, TI. *Chemical Biology*. Wiley-VCH Verlag GmbH; 2008. WOMBAT and WOMBAT-PK: Bioactivity Databases for Lead and Drug Discovery; p. 760-786.
89. Rowley M, Hallett DJ, Goodacre S, Moyes C, Crawford J, Sparey TJ, Patel S, Marwood R, Thomas S, Hitzel L, O'Connor D, Szeto N, Castro JL, Hutson PH, MacLeod AM. 3-(4-Fluoropiperidin-3-yl)-2-phenylindoles as high affinity, selective, and orally bioavailable h5-HT(2A) receptor antagonists. *J Med Chem*. 2001; 44:1603–1614. [PubMed: 11334570]
90. Bell IM, Gallicchio SN, Abrams M, Beshore DC, Buser CA, Culberson JC, Davide J, Ellis-Hutchings M, Fernandes C, Gibbs JB, Graham SL, Hartman GD, Heimbrook DC, Homnick CF, Huff JR, Kassahun K, Koblan KS, Kohl NE, Lobell RB, Lynch JJ, Miller PA, Omer CA, Rodrigues AD, Walsh ES, Williams TM. Design and biological activity of (S)-4-(5-((1-(3-chlorobenzyl)-2-oxopyrrolidin-3-ylamino)methyl)imidazol-1-ylmethyl)benzotrile, a 3-aminopyrrolidinone farnesyltransferase inhibitor with excellent cell potency. *J Med Chem*. 2001; 44:2933–2949. [PubMed: 11520202]
91. Bell IM, Gallicchio SN, Abrams M, Beese LS, Beshore DC, Bhimnathwala H, Bogusky MJ, Buser CA, Culberson JC, Davide J, Ellis-Hutchings M, Fernandes C, Gibbs JB, Graham SL, Hamilton KA, Hartman GD, Heimbrook DC, Homnick CF, Huber HE, Huff JR, Kassahun K, Koblan KS, Kohl NE, Lobell RB, Lynch JJ, Robinson R, Rodrigues AD, Taylor JS, Walsh ES, Williams TM, Zartman CB. 3-Aminopyrrolidinone farnesyltransferase inhibitors: design of macrocyclic compounds with improved pharmacokinetics and excellent cell potency. *J Med Chem*. 2002; 45:2388–2409. [PubMed: 12036349]
92. Peukert S, Brendel J, Pirard B, Brüggemann A, Below P, Kleemann HW, Hemmerle H, Schmidt W. Identification, synthesis, and activity of novel blockers of the voltage-gated potassium channel Kv1.5. *J Med Chem*. 2003; 46:486–498. [PubMed: 12570371]
93. Blum CA, Zheng X, De Lombaert S. Design, synthesis, and biological evaluation of substituted 2-cyclohexyl-4-phenyl-1H-imidazoles: potent and selective neuropeptide Y Y5-receptor antagonists. *J Med Chem*. 2004; 47:2318–2325. [PubMed: 15084130]
94. Fourches D, Muratov E, Tropsha A. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J Chem Inf Model*. 2010; 50:1189–1204. [PubMed: 20572635]
95. Varnek A, Fourches D, Horvath D, Klimchuk O, Gaudin C, Vayer P, Solov'ev V, Hoonakker F, Tetko IV, Marcou G. ISIDA - Platform for virtual screening based on fragment and pharmacophoric descriptors. *Curr Comput Aided Drug Des*. 2008; 4:191–198.
96. Solov'ev V, Varnek A, Wipff G. Modeling of ion complexation and extraction using substructural molecular fragments. *J Chem Inf Comput Sci*. 2000; 40:847–858. [PubMed: 10850791]

97. Duan J, Dixon SL, Lowrie JF, Sherman W. Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. *J Mol Graph Model*. 2010; 29:157–170. [PubMed: 20579912]
98. Mazanetz MP, Marmon RJ, Reisser CBT, Morao I. Drug discovery applications for KNIME: an open source data mining platform. *Curr Top Med Chem*. 2012; 12:1965–1979. [PubMed: 23110532]
99. MACCS structural keys. Accelrys; San Diego, CA: 2013.
100. Rogers D, Hahn M. Extended-connectivity fingerprints. *J Chem Inf Model*. 2010; 50:742–754. [PubMed: 20426451]
101. Morgan HL. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *J Chem Doc*. 1965; 5:107–113.
102. Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today*. 2010; 15:444–450. [PubMed: 20362693]
103. Steinbeck C, Han Y, Kuhn S, Horlacher O, Luttmann E, Willighagen E. The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics. *J Chem Inf Comput Sci*. 43:493–500. [PubMed: 12653513]
104. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R, Consonni V, Kuz'min VE, Cramer R, Benigni R, Yang C, Rathman J, Terfloth L, Gasteiger J, Richard A, Tropsha A. QSAR Modeling: Where Have You Been? Where Are You Going To? *J Med Chem*. 2014 Epub ahead.
105. Zhang S, Golbraikh A, Oloff S, Kohn H, Tropsha A. A novel automated lazy learning QSAR (ALL-QSAR) approach: method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models. *J Chem Inf Model*. 2006; 46:1984–1995. [PubMed: 16995729]
106. Vapnik, V. *The Nature of Statistical Learning Theory*. 2. Springer; New York: 2000. p. 314
107. Breiman LEO. *Random Forests*. *Mach Learn*. 2001; 45:5–32.
108. Breiman, L.; Friedman, JH.; Olshen, RA.; Stone, CJ. *Classification and Regression Trees*. In: Hall Crc, C., editor. *Statistics/Probability Series*. Vol. 19. Wadsworth; Belmont: 1984. p. 368
109. Breiman L. Bagging predictors. *Mach Learn*. 1996; 24:123–140.
110. Berk, RA. *Springer Series in Statistics*. Springer; New York, NY: 2008. *Statistical Learning from a Regression Perspective*; p. 360
111. Aronov AM, Goldman BB. A model for identifying HERG K+ channel blockers. *Bioorg Med Chem*. 2004; 12:2307–2315. [PubMed: 15080928]
112. Kuz'min VE, Muratov EN, Artemenko AG, Varlamova EV, Gorb L, Wang J, Leszczynski J. Consensus QSAR modeling of phosphor-containing chiral AChE inhibitors. *QSAR Comb Sci*. 2009; 28:664–677.
113. Zhu H, Tropsha A, Fourches D, Varnek A, Papa E, Gramatica P, Oberg T, Dao P, Cherkasov A, Tetko I. V Combinatorial QSAR modeling of chemical toxicants tested against *Tetrahymena pyriformis*. *J Chem Inf Model*. 2008; 48:766–784. [PubMed: 18311912]
114. Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem*. 2000; 43:3714–3717. [PubMed: 11020286]
115. Labute P. A widely applicable set of descriptors. *J Mol Graph Model*. 2000; 18:464–477. [PubMed: 11143563]
116. Springer C, Sokolnicki KL. A fingerprint pair analysis of hERG inhibition data. *Chem Cent J*. 2013; 7:167. [PubMed: 24144230]
117. Stumpfe D, Bajorath J. Exploring activity cliffs in medicinal chemistry. *J Med Chem*. 2012; 55:2932–2942. [PubMed: 22236250]
118. Melo-Filho CC, Braga RC, Andrade CH. Advances in Methods for Predicting Phase I Metabolism of Polyphenols. *Curr Drug Metab*. 2014; 15:120–126. [PubMed: 24479689]
119. Braga RC, Alves VM, Fraga CaM, Barreiro EJ, de Oliveira V, Andrade CH. Combination of docking, molecular dynamics and quantum mechanical calculations for metabolism prediction of

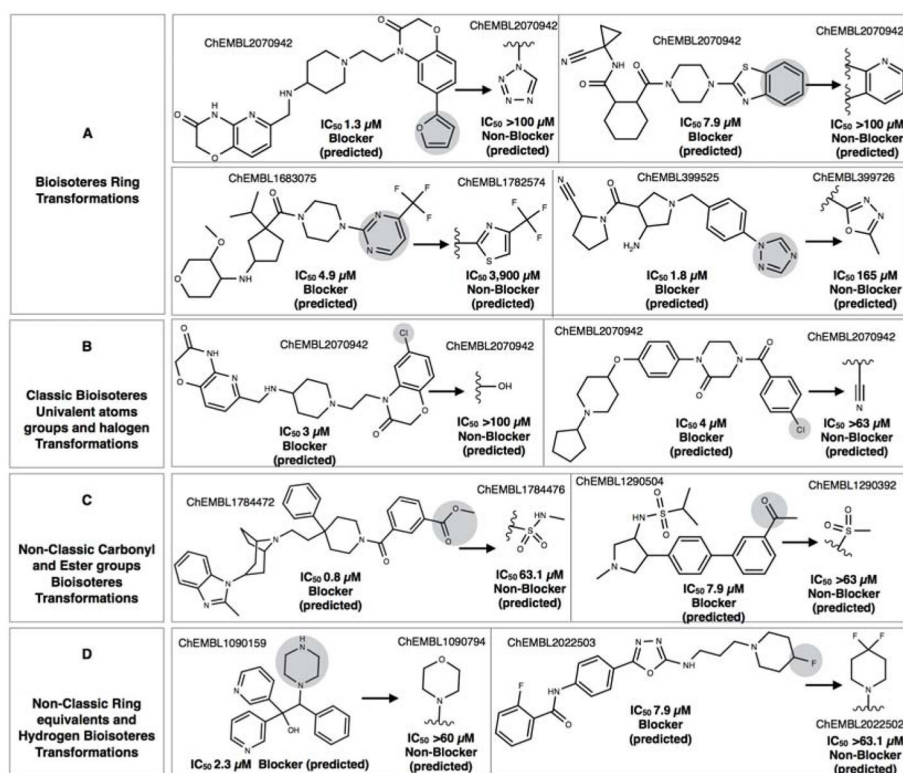
- 3,4-methylenedioxybenzoyl-2-thienylhydrazine. *J Mol Model*. 2012; 18:2065–2078. [PubMed: 21901409]
120. Braga RC, Andrade CH. QSAR and QM/MM approaches applied to drug metabolism prediction. *Mini Rev Med Chem*. 2012; 12:573–582. [PubMed: 22587770]
121. Andrade CH, Silva DC, Braga RC. In silico Prediction of Drug Metabolism by P450. *Curr Drug Metab*. 2014
122. Carneiro EO, Andrade CH, Braga RC, Tôrres ACB, Alves RO, Lião LM, Fraga CaM, Barreiro EJ, de Oliveira V. Structure-based prediction and biosynthesis of the major mammalian metabolite of the cardioactive prototype LASSBio-294. *Bioorg Med Chem Lett*. 2010; 20:3734–3736. [PubMed: 20488703]
123. Sousa MCM, Braga RRC, Cintra BASB, de Oliveira V, Andrade CH. In silico metabolism studies of dietary flavonoids by CYP1A2 and CYP2C9. *Food Res Int*. 2013; 50:102–110.
124. Pazini F, Menegatti R, Sabino JR, Andrade CH, Neves G, Rates SMK, Noël F, Fraga CaM, Barreiro EJ, de Oliveira V. Design of new dopamine D2 receptor ligands: biosynthesis and pharmacological evaluation of the hydroxylated metabolite of LASSBio-581. *Bioorg Med Chem Lett*. 2010; 20:2888–2891. [PubMed: 20363131]
125. Andrade CH, de Freitas LM, de Oliveira V. Twenty-six years of HIV science: an overview of anti-HIV drugs metabolism. *Brazilian J Pharm Sci*. 2011; 47:209–230.
126. Braga RC, Alves VM, Silva AC, Liao LM, Andrade CH. Virtual Screening Strategies in Medicinal Chemistry: The state of the art and current challenges. *Curr Top Med Chem*. 2014
127. Neves BJ, Bueno RV, Braga RC, Andrade CH. Discovery of new potential hits of *Plasmodium falciparum* enoyl-ACP reductase through ligand- and structure-based drug design approaches. *Bioorg Med Chem Lett*. 2013; 23:2436–2441. [PubMed: 23499236]
128. Bueno RV, Braga RC, Segretti ND, Ferreir EI, Trossini GHG, Andrade CH. New Tuberculostatic Agents Targeting Nucleic Acid Biosynthesis: Drug Design using QSAR Approaches. *Curr Pharm Des*. 2013
129. de Gil ES, Andrade CH, Barbosa NL, Braga RC, Serrano SHP. Cyclic Voltammetry and Computational Chemistry Studies on the Evaluation of the Redox Behavior of Parabens and other Analogues. *J Braz Chem Soc*. 2012; 23:565–572.
130. Braga, RC.; Sabino, JR.; de Valeria, O.; Andrade, CH. Discovery of novel hit compounds for *Trypanosoma cruzi* sterol 14 $\alpha$ -demethylase through structure-based virtual screening. Abstracts of Papers, 240th American Chemical Society National Meeting & Exposition; Boston, MA, United States. August 22–26; 2010. p. MEDI–379
131. Braga, RC.; Lião, LM.; Bezerra, JCB.; Vinaud, MCB.; Andrade, CH. Integrated chemoinformatics approaches to virtual screening in the search of new lead compounds against *Leishmania*. Abstracts of Papers, 244th American Chemical Society National Meeting & Exposition; Philadelphia, PA, United States. August 19–23, 2012; 2012. p. CINF–46



**Figure 1.** SAR rules involving structural transformations that change the environment of the amine nitrogen. For each compound, we show the experimental potency (IC<sub>50</sub>) available on ChEMBL database and the corresponding prediction by consensus model for 10 μM threshold.



**Figure 2.** SAR rules involving miscellaneous transformations. For each compound, we show the experimental potency ( $IC_{50}$ ) available on ChEMBL database and the corresponding prediction by consensus model for 10  $\mu$ M threshold. (C) TPSA is the difference between the topological polar surface area of the two molecules involved in the transformation. (D) Labute ASA is the Labute's approximate surface area descriptor.



**Figure 3.** Transformations involving bioisosteric replacements, showing activity cliffs. For each compound, we show the experimental potency ( $IC_{50}$ ) available on ChEMBL database and the corresponding prediction by consensus model for 10  $\mu$ M threshold.

Table 1

QSAR studies published for predicting hERG blockage.

Type of assay/Endpoint	Number of compounds	Dataset reference	Descriptors	Machine Learning	Performance	AD	Y-rand	Ref.
Inhibition of hERG K+ channels on mammalian cells evaluated by patch clamp.	37	[20]	CoMFA	PLS	$r^2=0.952$ $q^2=0.767$ $r^2_{\text{pred}}=0.744$	No	No	[21]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch clamp.	32	[7, 22]	CoMSIA	PLS	$q^2=0.571$	No	No	[23]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch clamp.	68	Fenichel's database, [24]	ClogP, molar refractivity, partial negative surface area, polarizability, and hydrophobicity.	SLR, PLS	$r^2=0.97-0.98$ $q^2=0.75-0.90$	No	No	[25]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	124	Various	Molecular, fragments, and experimental parameters	GP	ROC=0.89	No	No	[26]
Human studies	106	Various	LSER	SVM	SE=0.97 SP=0.84	No	Yes	[27]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	73	Various	MOE	SVM	Acc=90-95%	No	No	[28]
Dofetilide displacement assay	58,963	In-house	E-state keys, Banard fingerprints	NN NB	Se=0.86 Sp=0.83 Se=0.84 Sp=0.80	No	No	[29]
Inhibition of hERG K+ channels on CHO cells evaluated by patch clamp.	882	[30]	GRIND, hologram fingerprints	PLS	$r^2=0.76-0.77$ $q^2=0.72-0.74$	No	No	[31]
Inhibition of hERG K+ channels on HEK cells evaluated by patch clamp.	82	Various	CODESSA	Heuristic	$q^2=0.82$	No	No	[32]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	99	Various	Electronic, topological, spatial, and structural	RP	$q^2=0.83$	No	No	[33]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	339	Various	Molecular, topological, quantum mechanical	DT	Acc=0.76-0.80	No	No	[34]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	439	Johnson & Johnson	Topologic, atom types, keys, electrotopologic, drug-like profiling, MOE	Feature selection, NN	$q^2=0.52$	No	No	[35]

Type of assay/Endpoint	Number of compounds	Dataset reference	Descriptors	Machine Learning	Performance	AD	Y-rand	Ref.
Inhibition of hERG K <sup>+</sup> channels on various cells evaluated by patch-clamp.	90	Not specified	Fragment	SVM, PLS, RF	$r^2=0.912$ $q^2=0.848$	No	Yes	[36]
Inhibition of hERG K <sup>+</sup> channels on various cells evaluated by patch-clamp.	2045	Roche corporation, [25]	Atom-type, FCPF_6	Atom-type CT, NB	ROC=0.879	No	No	[37]
Inhibition of hERG K <sup>+</sup> channels on mammalian cells evaluated by patch-clamp.	203	Aureus Pharma	MOE	RP	Acc=0.81–0.96	Yes	No	[38]
IonWorks™ HT high throughput electrophysiology assay.	8,832	HTS	SELMA, VolSurf, fragments, DRONE	HPLS	$q^2=0.23-0.59$	Yes	Yes	[39]
Inhibition of hERG K <sup>+</sup> channels on various cells evaluated by patch-clamp.	39	Various	Pharmacophore ensemble	SVM	$q^2=0.89$ $r^2_{\text{pred}}=0.94$	No	No	[40]
Inhibition of hERG K <sup>+</sup> channels on mammalian cells evaluated by patch-clamp.	137	Various	SMARTS-based, logP, MW, McGowan's volume, Vx.	PLS, GP, GP-FVS, GP-Opt, GP-Nest	$r^2=0.63-0.82$ $q^2=0.74-0.81$	Yes	No	[41]
Inhibition of hERG K <sup>+</sup> channels on mammalian cells evaluated by patch-clamp.	113	Fenichel's database, PubChem	CORINA, ParaSurf'06, ClogP	PLS, SVR	$r^2=0.41-0.84$ $q^2=0.39-0.79$	No	No	[42]
Multiple assays	163	Various	MNA	PASS	Acc=0.87	No	No	[43]
Inhibition of hERG K <sup>+</sup> channels on mammalian cells evaluated by patch-clamp.	68	Various	Cerius	GFA	$q^2=0.776$ $r^2_{\text{pred}}=0.701$	Yes	Yes	[44]
Inhibition of hERG K <sup>+</sup> channels on mammalian cells evaluated by patch-clamp.	32	Various	3D pharmacophore	HipHop	$r^2=0.90$	No	No	[45]
Inhibition of hERG K <sup>+</sup> channels on various cells evaluated by patch-clamp.	313	Various	MOE	NB	Acc=0.70–0.89 Se=0.53–1.00 Sp=0.40–1.00	Yes	No	[46]
Inhibition of hERG K <sup>+</sup> channels on various cells evaluated by patch-clamp.	83	[33]	Shape, charges	SVM, UFS-SVM, kNN	Acc=0.67–0.74 Se=0.70–0.79 Sp=0.53–0.74	Yes	No	[47]
Inhibition of hERG K <sup>+</sup> channels on various cells evaluated by patch-clamp.	561	Various	GRIND	SVM	Acc=0.72 Se=0.85 Sp=0.36	No	No	[48]
Inhibition of hERG K <sup>+</sup> channels on CHO cells evaluated by patch-clamp.	1043	In-house and ref.[25]	Atom type	SVM	Acc=0.74–0.94 ROC=0.76–0.90	No	No	[49]



Type of assay/Endpoint	Number of compounds	Dataset reference	Descriptors	Machine Learning	Performance	AD	Y-rand	Ref.
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	166	Various	Bio-Suite VolSurf	KNN LLR	$q^2=0.59-0.70$ $q^2=0.76-0.81$	No	Yes	[50]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	285	Various	MOE, P_VSA, VolSurf, Inductive	SIBAR CPG-NN	Acc=0.70-0.92 Acc=0.49-0.89	No	No	[51]
Not specified	3916	In house	Fingerprints	NB	Ac=0.53	No	No	[52]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	37	[21]	GRIND 3D CoMFA	PLS PLS	$r^2=0.93$ $q^2_{LOO}=0.69$ $r^2=0.77$ $q^2_{LOO}=0.95$	No	No	[53]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	676	In house, [42]	JChem pKa, Corina, MMFF94, CATS, VolSurf	RR, GP, SVR, RF, BM	RMSE=0.60 $q^2_{LOO}=0.72$	Yes	No	[54]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	275	Various	Pharmacophore, CoMFA-like	GA Topomer search	Acc=0.75-0.92 Se=0.60-0.87 Sp=0.80-0.94	Yes	No	[55]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	1970	Various	4D-fingerprints, MOE	MOE, 4D fingerprints	Acc=0.83	No	No	[56]
Multiple assays	2644	Various	ECFP, FCFP	LDA, SVM	AUC=0.89-0.94	No	Yes	[57]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	400	Various	3D-GRIND-2	PLS	$r^2=0.52$ $q^2=0.46$	No	No	[58]
Multiple assays	368	Various	Structural fingerprints, P_VSA	SVM, RF, Winnow	MCC=0.10-0.83	Yes	No	[59]
Inhibition of hERG K+ channels on HEK and CHO cells evaluated by patch clamp.	157	Various	Molecular, electronic	NN	$r^2=0.73$ $q^2=0.64$	Yes	Yes	[60]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	529	Various	MOE	KNN	$r^2=0.59$ $q^2=0.56$	Yes	Yes	[61]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	293	In house and ref.[62]	FCFP_4	NB RF	Acc=0.82-0.96 Se=0.78-0.93 Sp=0.88-1.00	No	No	[63]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch clamp and HTS.	2214	PubChem BioAssay AID 376	MOE, fingerprints	SVM	Acc=0.73-0.90	No	No	[64]

Type of assay/Endpoint	Number of compounds	Dataset reference	Descriptors	Machine Learning	Performance	AD	Y-rand	Ref.
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	803	Various	Dragon, MOE, Molconnz, MACCS, CDK, fingerprints, VolSurf	GA-kNN, LDA, RF, PLSD	Acc=0.76–0.97 F <sub>score</sub> =0.77–0.96	No	No	[65]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	242	[55]	Cerius, Dragon, PaDEL	LDA, PLS, GFA, G/PLS	r <sup>2</sup> <sub>pred</sub> =0.53–0.60	Yes	Yes	[66]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	113	[67]	CODESSA	Heuristic	r <sup>2</sup> =0.91 r <sup>2</sup> <sub>cv</sub> =0.87	No	No	[68]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	806	Various	Molecular, ECFP_8 fingerprints	NB, RP	Acc=0.85–0.89	No	No	[69]
Inhibition of hERG K+ channels on mammalian cells evaluated by patch-clamp.	1686	Various	MACCS	NB, KNB, PWM	SE=0.55–0.66 3-CCC=0.48–0.58	No	No	[70]
Multiple assays	3721	ChEMBL	RDKit	RF	AUC=-0.66	No	Yes	[71]
	694	PubChem			AUC=0.56–0.66			
IonWorks™ HT high throughput electrophysiology assay.	45	[72, 73]	Dragon	SW-MLR GA-MLR GA-SVM	q <sup>2</sup> =0.569 q <sup>2</sup> =0.797 q <sup>2</sup> =0.894	Yes	Yes	[74]
	59	Various	Docking-based, PaDEL	GA-MLR	q <sup>2</sup> <sub>Looc</sub> =0.57 r <sup>2</sup> <sub>pred</sub> =0.67	No	No	[75]
	25	[76, 77]	MOE	MLR	q <sup>2</sup> <sub>LMO</sub> =0.71–0.87 q <sup>2</sup> <sub>ext</sub> =0.82–0.88	Yes	Yes	[78]
Inhibition of hERG K+ channels on various cells evaluated by patch clamp.	98	Various	Physico-chemical, <i>in vitro</i> parameters	Various	RMSE=0.86–1.17	No	No	[79]

Acc: accuracy; AD: applicability domain; BM: baseline model; CDK: chemistry development kit; CHO: Chinese hamster ovary; CoMFA: comparative molecular field analysis; CoMSIA: comparative molecular similarity index analysis; CPG-NN: counter-propagation neural networks; CT: classification tree; ECFP: extended connectivity fingerprints; FCFP: functional class fingerprints; G/PLS: genetic partial least squares; GA-kNN: k-nearest neighbor with genetic algorithm variable selection; GFA: genetic function approximation; GFA-MLR: genetic function approximation with multiple linear regression; GP: genetic programming; GP-FVS: Gaussian process with forward variable selection; GP-Nest: Gaussian process with nested sampling; GP-Opt: Gaussian process with conjugate gradient optimization; HEK: human embryonic kidney; HPLS: hierarchical PLS; HTS: high-throughput screening; KNB: kernelized naïve Bayes; kNN: k-nearest neighbors; LDA: linear discriminant analysis; LLR: local lazy regression; MNA: multilevel neighborhoods of atoms; MOE: molecular operating environment; NB: naïve Bayes; NN: neural networks; PLS: partial least squares; PLSD: partial least squares discriminant; parzen-window based model; rho: Spearman's rank correlation coefficient; RMSE: root-mean-square error; ROC: receiver operating characteristic; RP: recursive partitioning; RR: ridge regression; Se: sensitivity; SOM: self-organizing map; Sp: specificity; SR: stepwise regression; SVM: support vector machines; SVR: support vector regression; Y-rand: Y-randomization.

**Table 2**

Number of compounds of hERG modeling set after curation.

	1 $\mu$ M Threshold		10 $\mu$ M Threshold		20 $\mu$ M Threshold	
	Blocker	Non-blocker	Blocker	Non-blocker	Blocker	Non-blocker
Modeling set	795	3,156	2,277	1,590	2,666	970
Evaluation Set	199	788	611	355	713	195

Summarized statistical characteristics of QSAR models for hERG liability assessed by 5-fold external cross-validation for the modeling set at the threshold level of 10  $\mu\text{M}$ .

Table 3

Model name	Modeling Set (threshold level 10 $\mu\text{M}$ ; $n=3,867$ )						
	Accuracy	BAC	MCC	Sensitivity	Specificity	AUC	Coverage
MACCS-SVM	0.84	0.75	0.52	0.86	0.64	0.75	0.70
featMorgan-SVM	0.86	0.78	0.57	0.86	0.69	0.78	0.74
Pharm. FP-SVM	0.84	0.74	0.50	0.85	0.63	0.74	0.78
PubChem-SVM	0.85	0.77	0.54	0.85	0.68	0.77	0.70
MACCS-RF	0.84	0.75	0.52	0.85	0.66	0.75	0.70
featMorgan-RF	0.85	0.77	0.55	0.86	0.68	0.77	0.74
Pharm. FP-RF	0.86	0.77	0.55	0.84	0.71	0.77	0.78
PubChem-RF	0.84	0.77	0.55	0.85	0.69	0.77	0.70
MACCS-TreeBag	0.84	0.76	0.53	0.85	0.67	0.76	0.70
featMorgan-TreeBag	0.85	0.76	0.52	0.84	0.67	0.76	0.74
Pharm. FP-TreeBag	0.85	0.74	0.50	0.83	0.66	0.74	0.78
PubChem-TreeBag	0.84	0.76	0.53	0.84	0.69	0.76	0.70
MACCS-GBM	0.83	0.74	0.49	0.85	0.62	0.74	0.70
featMorgan-GBM	0.85	0.74	0.50	0.84	0.64	0.74	0.74
Pharm. FP-GBM	0.86	0.75	0.51	0.84	0.65	0.75	0.78
PubChem-GBM	0.85	0.76	0.52	0.85	0.66	0.76	0.70
Consensus	0.89	0.83	0.65	0.85	0.81	0.83	0.74
Consensus Rigor	0.91	0.87	0.71	0.89	0.84	0.87	0.34

SVM: Support Vector Machine; RF: Random Forest; TreeBag: Tree Bagging method; MACCS: MACCS keys; PubChem: PubChem Fingerprints; Pharm. FP: ChemAxon Pharmacophore Fingerprint; BAC: Balanced Accuracy; MCC: Matthews correlation coefficient; AUC: area under the receiver operating characteristic curve; Consensus and Consensus Rigor models were built by averaging the predicted values from the individual model for each machine learning technique that yielded the best performance with higher coverage (featMorgan-GBM, PubChem-TreeBag, Pharm. FP-RF, MACCS-SVM).

Statistical results of some QSAR models for hERG liability for the external test set at the threshold level of 10  $\mu\text{M}$ .

**Table 4**

Model name	Test Set (threshold 10 $\mu\text{M}$ , $n=966$ )						
	Accuracy	BAC	MCC	Sensitivity	Specificity	AUC	Coverage
MACCS-SVM	0.83	0.79	0.60	0.91	0.66	0.79	0.85
Pharm. FP-RF	0.84	0.81	0.64	0.91	0.70	0.81	0.88
PubChem-TreeBag	0.83	0.80	0.61	0.88	0.73	0.80	0.71
featMorgan-GBM	0.83	0.80	0.61	0.90	0.69	0.80	0.80
Consensus	0.90	0.91	0.76	0.89	0.93	0.91	0.78

SVM: Support Vector Machine; RF: Random Forest; TreeBag: Tree Bagging method; GBM: Gradient boosting method; MACCS: MACCS keys; PubChem: PubChem Fingerprints; Pharm. FP: ChemAxon Pharmacophore Fingerprint; BAC: Balanced Accuracy; MCC: Matthews correlation coefficient; AUC: the area under receiver operating characteristic curve; Consensus was built by averaging the predicted value from each individual model (featMorgan-GBM, PubChem-TreeBag, Pharm. FP-RF, MACCS-SVM).

Comparison of the statistical values between the best models generated in this research and the work of Li et al. [48].

**Table 5**

Model name	Accuracy	BAC	MCC	Sensitivity	Specificity	AUC	Coverage
	WOMBAT-PK dataset (threshold 1 $\mu$ M, n=66)						
SVM linear <sup>d</sup>	-	0.72	-	0.68	0.74	-	-
SVM non linear <sup>d</sup>	-	0.72	-	0.83	0.47	-	-
Consensus	0.96	0.95	0.89	0.92	0.98	0.95	0.85
Model name	WOMBAT-PK dataset (threshold 10 $\mu$ M, n=66)						
SVM linear <sup>d</sup>	-	0.57	-	0.56	0.59	-	-
SVM non linear <sup>d</sup>	-	0.71	-	0.77	0.59	-	-
Consensus	0.98	0.98	0.96	0.98	0.99	0.99	0.89
Model name	WOMBAT-PK dataset (threshold 20 $\mu$ M, n=66)						
SVM linear <sup>d</sup>	-	0.60	-	0.60	0.60	-	-
SVM non linear <sup>d</sup>	-	0.74	-	0.91	0.35	-	-
Consensus	0.96	0.97	0.91	0.95	0.99	0.97	0.91

<sup>d</sup>The models of Li et al. [48].