A Critical Assessment of Combined Ligand- and Structure-Based Approaches to hERG Channel Blocker Modeling

Lei Du-Cuny,[†,§] Lu Chen,[†,§] and Shuxing Zhang*[,†,‡]

[†]Integrated Molecular Discovery Laboratory, Department of Experimental Therapeutics and [‡]Molecular Modeling and Structural Biology Core, The University of Texas, M.D. Anderson Cancer Center, 1901 East Rd., Houston, Texas

S *Supporting Information*

ABSTRACT: Blockade of human ether-à-go-go related gene (hERG) channel prolongs the duration of the cardiac action potential and is a common reason for drug failure in preclinical safety trials. Therefore, it is of great importance to develop robust in silico tools to predict potential hERG blockers in the early stages of drug discovery and development. Herein we described comprehensive approaches to assess the discrimination of hERG-active and -inactive compounds by combining quantitative structure—activity relationship (QSAR) modeling, pharmacophore analysis, and molecular docking. Our consensus models demonstrated high-predictive capacity and improved enrichment and could correctly classify 91.8% of 147 hERG blockers from 351 inactives. To further enhance our modeling effort, hERG homology models were constructed, and molecular docking studies were conducted, resulting in high correlations ($R^2 = 0.81$) between predicted and experimental pIC_{50}s. We expect our unique models can be applied to efficient screening for hERG blockades, and our extensive understanding of the hERG-inhibitor interactions will facilitate the rational design of drugs devoid of hERG channel activity and hence with reduced cardiac toxicities.
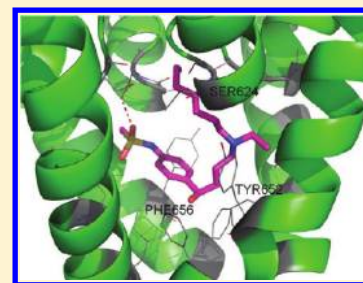
■ INTRODUCTION

Drug cardiac arrhythmia has become a major safety concern for both pharmaceutical development and health regulatory authorities.[1] In recent years, a number of clinically successful drugs have been withdrawn from the market due to their drug-induced sudden cardiac deaths.[2] This toxicity is caused by the blockade of the human ether-à-go-go related gene (hERG) channel which leads to the long QT interval, an abnormality in cardiac muscle repolarization. This syndrome has been implicated as a predisposing factor for torsades de points, a polymorphic ventricular tachycardia that can spontaneously degenerates to ventricular fibrillation and causes sudden death.[3] An experimental assessment of hERG inhibition is time-consuming and costly;[4] therefore, it is necessary to develop reliable in silico modeling methods to screen the hERG activity of drug candidates during the early stages of drug discovery and development. Furthermore, a detailed structural understanding of hERG channel would help design drugs with reduced hERG activity.

The Kv11.1 potassium (K^+) ion channel encoded by hERG is known for contributing to the electrical activity of the heart. However, the 3D structure is not yet available. Under this situation, various modeling studies have been performed to assess and predict potential hERG liability. For instance, ligand-based methods, such as principle component analysis (PCA), artificial neural network (ANN), and support vector machine (SVM), along with 2D molecular properties and structural fingerprints are frequently employed to predict hERG liability.[5−10] Descriptors with easily interpretable physicochemical meanings were applied by Yoshida et al.[5] to derive a quantitative structure—activity relationship (QSAR) model with

$r^2 = 0.70$ and $q^2 = 0.67$ for 104 hERG blockers. In a validation study, ANN models were used to classify 93% of 72 nonblocking agents and 71% of 23 hERG channel blockers.[6] Yufeng Tseng's group conducted a series of PLS and SVM classification studies of hERG blockers based on Pubchem bioassay data sets using 4D fingerprints and traditional 2D descriptors.[7,8] "Shape signature" molecular descriptors in conjunction with SVM were applied by Ekins et al., and the binary classification model achieved 69—73% cross-validation accuracy for the training set.[9] Other 2D fragment-based approaches, such as hologram QSAR[10] and substructure-based v-support vector regression,[11] were utilized without explicit consideration of the 3D binding modes and calculation of the molecular descriptors. Comparative molecular field analysis (CoMFA) studies of hERG channel blockers were also described, and the authors concluded that four features are critical to the hERG binding activity:[1] a flexible molecule, a central tertiary amine, at least two aromatic moieties, and a polar group on one side of the molecule. Ekins et al.[12] and Aronov et al.[13,14]generated 3D pharmacophore models for hERG blockers. Hydrophobic and positive ionizable features were found as important for hERG liability. To further understand the structure—function relationships of hERG channel, its 3D structure was constructed using homology modeling.[15,16] Li et al. combined GRIND pharmacophore descriptors with SVM method to predict hERG blockades.[17] However, this SVM model only achieved 72% accuracy on their test set. Combined receptor- and ligand-based approaches were also used to develop a universal pharmacophore
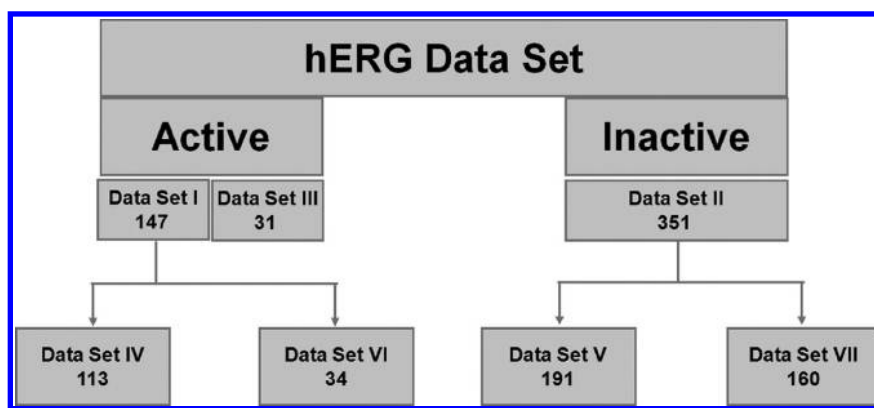
Figure 1. The hERG data sets used for the QSAR, pharmacophore, and docking analysis. Data Set I contains 147 hERG active compounds; Data Set II contains 351 inactive compounds; Data Set III contains the 31 actives compiled by Ekins et al.;[32] split from Data Set I, Data Set IV contains 113 actives with basic moieties; Data Set VI contains 34 actives without basic moieties; split from Data Set II, Data Set V contains 191 inactives with basic moieties; and Data Set VII contains 160 inactives without basic moieties.

models, providing rapid assessment of drug blocking ability to the hERG channel with high accuracy.[18] Unfortunately, many of these studies were conducted by commercial groups, and their models/data are not publicly available. These published works either employed relatively small collections of hERG blockades or used the data sets in which the IC_{50} or K_i values are divergent and lack data consistency. In addition, there has been limited work to comprehensively benchmark the performance of different computational methods for predictions of potential hERG inhibitors.

In the present study, we compiled a much larger data set containing 178 hERG blockers, and this enabled us to develop more robust hERG models with higher statistical significance and better predictive capability. Furthermore, we integrated both structure- and ligand-based methods and evaluated their performance. Specifically, we developed variable selection k nearest neighbor (kNN) QSAR and 3D pharmacophore models to predict the hERG inhibition. With QSAR methods, a collection of 2D descriptors was applied to a set of known hERG-active ("actives") and hERG-inactive ("inactives") drugs. For the 3D pharmacophore modeling, we employed an automated 3D pharmacophore ensemble approach. In both cases, classification models were developed to discriminate the actives from inactives in the data sets. This provides robust models for rapidly screening large virtual libraries to identify molecules with potential hERG liabilities. In addition, to probe hERG—inhibitor interactions, 3D homology models of hERG channel protein were constructed for molecular docking studies. A set of 12 hERG blockers were used to validate the structural models based on their known pIC_{50}s and available mutagenesis data.[19−26] Our results were consistent with experimental studies, and the pIC_{50} values could be reproduced for a wide range of ligands against the hERG channel.
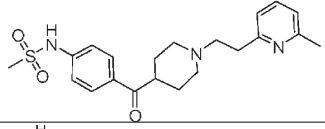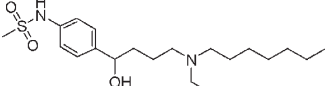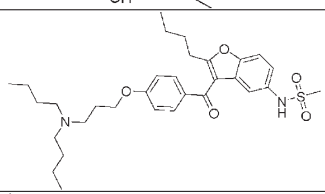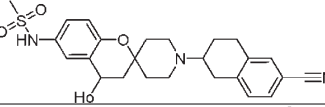
■ MATERIALS AND METHODS
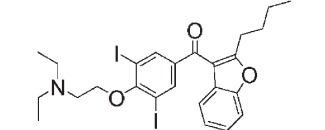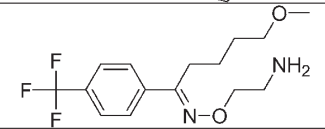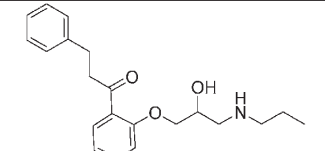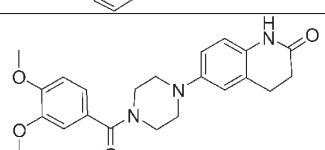
Data Sets. IC_{50} values for hERG blockers were collected from previous publications.[1,5,11,27−32] Some discrepancies were observed between the data set published by Ekins et al.[32] and others.[1,5,11,27−31] Based on thorough literature investigation, we found data from resources other than Ekins were more trustworthy, as they were reported by more than one research group and consistent to our own predictions. In total we collected 147 hERG active compounds as Data Set I, and although the 31 compounds

from Ekins et al[32] were excluded from model building, they were kept as Data Set III (Figure 1) for external testing of our models. In addition, a set of compounds was compiled by extracting orally deliverable drugs from the DrugBank,[33] with molecular weights less than 500 Da. Compounds in this compilation have very diverse structures, and they were for different disease treatments, including anti-HIV, anticancer, antiobesity, antidepressant, etc. We further investigated literature thoroughly for this chemical data set, and if any relation between any compound and the hERG protein is indicated, then the compound will be excluded. This resulted in 351 compoun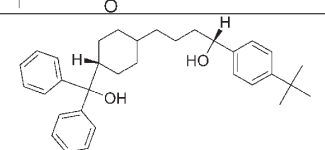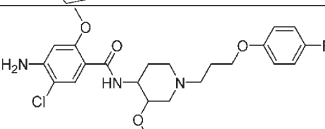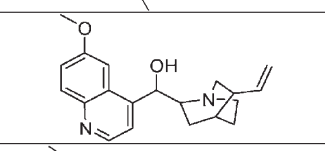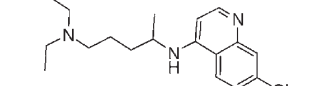ds forming the presumably hERG-inactive set (Data Set II). Furthermore, it is known that a basic nitrogen moiety is one of the major features[1,31,34] for hERG activities. Thereby, we divided Data Set I into two subsets: compounds with and without (Data Sets IV and VI, respectively) a basic moiety. Similarly, the inactives (Data Set II) were divided into Data Sets V and VII with and without the basic nitrogen moiety, respectively. The 3D structures of all compounds in our data sets were generated using the Molecular Operating Environment (MOE)[14] software. This includes the elimination of counterions, addition of hydrogen atoms, and calculation of standard protonation states for ionizable atoms at a physiological pH (pH = 7.4). The molecules were then subjected to energy minimization to generate the lowest energy conformation.

QSAR Modeling of hERG Blockers. The 147 hERG blockers (Data Set I) with known hERG inhibition activity (pIC_{50}) were used to build regression models. For the classification model construction, both Data Sets II (351 hERG-inactive compounds) and I (147 hERG-active compounds) were employed to form the training and test sets. MOE[14] was used to calculate 184 molecular descriptors, which were normalized to avoid disproportional weighting. Using the sphere exclusion algorithm, as we described previously,[35,36] the compounds were split into multiple training and test sets for both regression and classification modeling.

The kNN pattern recognition principle[37] and a variable selection procedure were applied to develop models of hERG inhibition. Briefly, a subset of n_{var} (number of selected variables) descriptors was selected randomly. The n_{var} parameter was tuned to obtain the best correlation for regression modeling (q^2 value) or best classification (accuracy value) for classification modeling. The models were optimized by leave-one-out cross validation, where each compound was eliminated from the training set and

Table 1A. Experimental Mutagenesis and pIC50 Data of 12 hERG Blockers

Compound	Structure	pIC$_{50 \text{ (Exp)}}$	Residues
e4031		7.74	T623, S624, V625, G648, Y652, F656, V659[20]
ibutilide		7.82	T623, S624, V625, G648, Y652, F656, V659[20]
dronedarone		4.23	F656[21]
mk499		7.47	T623, S624, V625, G648, Y652, F656, V659[20]
amiodarone		5.00	Y652, F656[21]
fluvoxamine		4.80	S631[22]
propafenone		6.36	F656[23]
vesnarinone		5.96	G628, S631[24]
terfenadine		6.70	Y652, F656[19]
cisapride		7.35	Y652, F656[19]
quinidine		6.49	F656 but not S631[25]
chloroquine		5.60	Y652, F656[26]

its hERG activity was predicted as the weighted average activity of the k-most similar molecules. The molecular similarity was characterized by the Euclidean distances between compounds in multidimensional descriptor space. A simulated annealing method with Metropolis-like acceptance criteria was used to optimize the selection of variables. A variety of models with $q^2 > 0.4$ (for regression) or accuracy > 40% (for classification) were obtained and validated on the test sets to select the predictive models with R^2 values > 0.5 or accuracy > 50%. To test the model reliability, Y-randomization[38] (randomization of response) was conducted and repeated three times for the training set using randomized activities. An applicability domain[39] with parameter $Z = 0.5$ was employed to ensure that only compounds similar to training set compounds were predicted and that the reliability was high.

3D Pharmacophore Modeling for the Prediction of hERG Blockers. Classic hERG compounds usually contain a positively charged nitrogen center;[40] however hERG activity has been observed in ligands lacking the moiety.[41−43] Therefore, two data sets were prepared for the hERG blockers with and without basic moieties to generate 3D pharmacophore models. For the data set with basic moieties, eight of the most active compounds studied by Cavalli et al.[1] with pIC_{50} values > 6.9 were used as actives and 191 drugs (Data Set V) as inactives. For the data set without basic moieties, 10 most active neutral compounds from Data Set VI with pIC_{50} values > 5.5, along with 160 drugs (Data Set VII) as hERG negatives, were used to generate pharmacophore models.

We used the pharmacophore elucidation method in MOE to generate the pharmacophore queries. The queries were selected according to the following steps: The input molecules were first prepared by assigning protonation states at pH of 7.0 and removing counterions. A collection of lowest energy conformers was generated and aligned in a flexible manner. Then the alignments were scored according to the internal strain and overlap of molecular features, e.g., aromaticity, hydrogen-bond donor/acceptor, hydrophobicity, etc.[44] Pharmacophore elucidation was subsequently performed to generate up to five-point pharmacophore queries using specified MOE pharmacophore annotation schemes, including unified,[45] PCHD,[45] and PPCH.[45] The "query spacing" was set to 0.8, and the "query cluster" was set to 1.25 during the query generation. Each query was then used to search the active molecules (8 for the charged and 10 for the uncharged data sets). Only those queries that lead to a good alignment of common pharmacophore features (overlap score[44] > 1.5 and coverage values > 50%) were retained. Finally, all queries were analyzed, and the best pharmacophore query was identified, according to its coverage of hERG actives and its ability to discriminate between the actives and inactives.

hERG Homology Modeling. Five different bacterial K[+] channels (KcsA, MthK, KvAP, KirBac1.1, and Kv1.2) have their crystal structures resolved. The crystal structure of KcsA is in the closed form, while MthK has a tetramer structure, but only part of KvAP, KirBac1.1, and Kv1.2 was resolved in open form. Hence, KcsA from *Streptomyces lividans*[46] (PDB ID: 1BL8) was used as the structural template (22% sequence identity) for homology modeling of hERG in the closed form, and MthK from *Methanothermobacter thermautotrophicus*[47] (PDB ID: 1LNQ) was used to model hERG in the open form (24% sequence identity). The sequence of hERG was aligned to KcsA and MthK using the software ALIGN[48] (Supplementary Figure 1, Supporting Information). Our alignments were identical to those published ones.[49] These sequence alignments were then used to generate 10 homology models with MODELLER 9v6.[50] The top model

with the lowest MODELLER energy scores was selected for further structural refinement as follows: The rotameric conformations of Tyr652 and Phe656 were manually adjusted so that they point toward the selectivity filter loops[15] of hERG; the steric conflicts were subsequently removed by refining the side chain conformations of the remaining residues based on the backbone-dependent rotamer library[51] in the Coot software;[52] and superimposition of the S5 helix of the selected lowest energy hERG models reveals that a bend in the S6 helix at Gly648 permits the structural differences that are observed between the closed and open states. The S6 helices in the closed and open state differentiate themselves about 40°. To derive the intermediate structures of hERG in different partially open states, the S6 helices in the closed form were bent toward (with Gly648 as the hinge point) the open form conformations using Coot[52] in a step of about 8° each time, as done by Rajamani.[16] The side chains of these derived partially open states were further refined using Coot.

Molecular Dynamics for Structural Optimization. Molecular dynamics (MD) simulations were carried out in AMBER 8[53] to further refine the hERG homology models using the generalized Born (GB) solvation model with the FF99 force field.[54] The side chains were minimized with 500 steps of steepest descent method followed by 500 conjugate gradient steps. Subsequently, the system was gradually heated to 300 K, and the dynamics time step is 2 fs. During the first 1.6 ps, the protein backbone coordinates were kept as rigid using harmonic constraints. With each monomer of only 56 residues, hERG pore domain is a relatively simple homotetramer, and its binding pocket consists of only four helices without any loops. Thus, we performed 500 ps MD structural refinement in GB solvation model with constraints removed, as suggested by some previous reports.[55]

Molecular Docking. The average structure of the ensembles of MD simulation was extracted for the molecular docking. Mutagenesis data provide important insights for binding pose selections. Herein, we collected 12 compounds, along with their binding data that change upon the mutation of several critical protein residues (Table 1A), to validate the 3D structure of the hERG models by analyzing the interactions between the hERG channel and its blockers. The program GOLD[56] was employed for the molecular docking of these 12 compounds with default parameters unless otherwise noted. We found Phe656 is critical for ligand binding[57] and sits in the middle of the channel. Therefore, the binding pocket was defined as 15 Å around the centroid of the four Phe656 residues of the tetramer that forms the pore domain (subunits S5−S6) of the hERG channel. For the selection of docking poses, residues important for ligand binding based on mutagenesis data (Table 1A) were employed to assess the hERG-inhibitor binding.

Evaluation of the hERG Homology Models Using Macro-Model and SIMCA. Six models were generated to capture the flexibility of the hERG channel of different open−close conformations. In order to discover which particular models are accurately constructed, each model was evaluated using 12 hERG blockers with known pIC_{50} values and mutagenesis data as follows: the selected docking pose of each ligand was minimized within the protein binding pocket using the Macro-Model module in Maestro.[58] The OPLS_2005 force field[59] and the generalized Born/surface area (GB/SA) continuum water model were applied. The minimization was continued until either the energy converged (0.05 convergence criterion) or the maximum 5000 iterations was reached. The protein residues within 8 Å of the inhibitors were optimized during the

Table 1B. Calculated Energitic Terms and Predicted pIC$_{50}$ for 12 Compounds Docked into the hERG Homology Model[a]

compound	Δstretch (kJ/mol)	Δbend (kJ/mol)	Δtorsion (kJ/mol)	Δimproper torsion (kJ/mol)	ΔvdW (kJ/mol)	Δelectrostatic (kJ/mol)	pIC$_{50(pred)}$	ΔpIC$_{50}$ (exp-pred)
e4031	1963.11	450.26	1217.21	48.38	−1285.49	−11115.75	6.78	0.96
ibutilide	1967.33	439.19	1253.39	49.69	−1360.04	−11500.08	7.02	0.80
dronedarone	1872.40	486.90	1160.45	44.68	−1430.63	−11979.89	4.14	0.09
mk499	1924.13	494.60	1194.96	43.61	−1418.60	−11282.03	7.04	0.43
amiodarone	1931.13	443.83	1178.21	43.28	−1313.19	−11378.87	5.20	−0.20
fluvoxamine	1978.35	473.81	1023.87	51.95	−1188.80	−11393.08	4.67	0.13
propafenone	2002.00	393.07	1044.21	39.84	−1224.08	−9887.91	6.65	−0.29
vesnarinone	1919.92	474.72	1301.54	47.78	−1401.05	−11608.96	6.91	−0.95
terfenadine	1939.98	489.88	1109.64	45.58	−1428.26	−11126.23	6.92	−0.22
cisapride	1915.36	521.30	1358.98	57.91	−1449.23	−12169.97	7.67	−0.32
quinidine	2058.25	388.36	968.05	36.35	−1072.18	−9693.92	6.57	−0.08
chloroquine	2011.19	400.62	1120.64	47.96	−1168.41	−10785.15	5.95	−0.35

[a] Note: Δstretch, Δbend, Δtorsion, Δimporper torsion, ΔvdW, and Δelectrostatic are estimated interaction energy differences for each ligand in complex and free states. Mutagenesis data were reported for an individual drug on their ability to block hERG functions.

minimization. All other residues were held as rigid. After the ligand−receptor complex was minimized, the ligand was extracted from the complex. The 3D structures of the protein and ligand were then minimized separately using a GB/SA continuum water model to obtain reference energies for their free states. The computed differences (ΔG_{bind}) in the stretch, bend, torsion, improper torsion, van der Waals, and electrostatic energies between the bound ($G_{complex}$) and free states ($G_{receptor}$, G_{ligand}) (eq 1) were used to correlate with the experimental pIC$_{50}$ values:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \qquad (1)$$

PLS (partial least squares) was used to predict the pIC$_{50}$ value of hERG blockers using the program SIMCA.[60] To this end, the above energetic terms (Table 1B) were used as descriptors (Δstretch, Δbend, Δtorsion, Δimproper torsion, ΔvdW, and Δelectrostatic) to establish their correlation with pIC$_{50}$. These descriptors were normalized to the unit variance. PCA was used to detect outliers as follows: New latent variables were calculated by summarizing the information contained in the original variables. The compounds which could not be well explained with the latent variables were classified as outliers. Outliers were identified using the 95% tolerance interval, which is signified as an ellipse in the PCA loading plot[60] (Supplementary Figure 2A, Supporting Information). The goodness of the fit (q^2) of a PLS model was evaluated with leave-one-out cross-validation. The q^2 value was used as the main criterion for assessing the quality of a given model. Once a model was chosen, it was validated by a Y-randomization test by scrambling the Y values to ensure that the model was not obtained by chance. For a generated model, the influence of every original variable on the matrix **Y** was estimated using the variable influence on projection (VIP) value[60] (Supplementary Figure 2C, Supporting Information). In a case where a strong correlation between pIC$_{50}$ and energy descriptors was found, the corresponding model structure can be employed to guide the hERG liability optimization.

■ RESULTS AND DISCUSSION

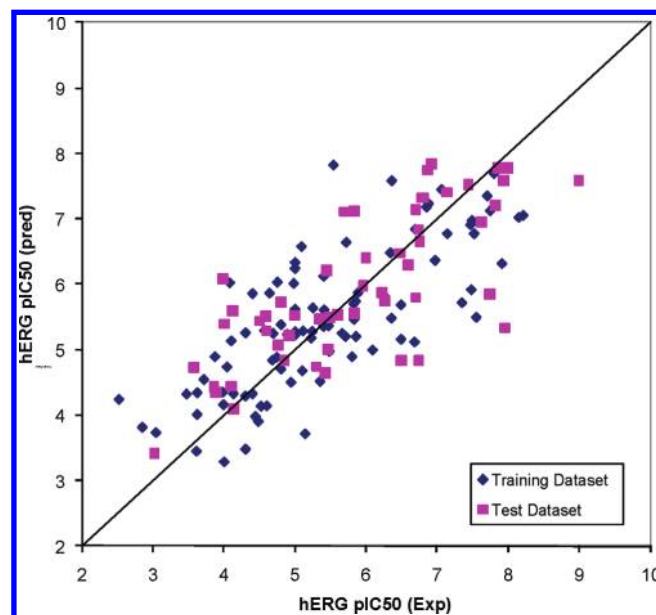**QSAR Regression Models for the Quantitative Prediction of hERG Inhibition.** Six QSAR models (Supplementary Table 1,

Figure 2. Experimental versus predicted hERG pIC$_{50}$ using one of our QSAR models. This model has 93 compounds in the training set and 54 in the test set.

Supporting Information) were obtained from kNN analyses with $q^2 > 0.55$ for the training set and $R^2 > 0.55$ for the test set compounds. The q^2, R^2, the number of compounds in the test set (n), and the number of descriptors (n_{var})[38] were used as criteria for selecting the top models. Generally, a predictive model was expected to have high q^2, R^2, and n values but with low n_{var} values and easily interpretable descriptors. The best model selected for the hERG prediction used eight descriptors to predict 93 compounds in the training set with a q^2 value of 0.56 and 54 compounds in the test set with a R^2 value of 0.59 (Figure 2). Y-randomization experiments did not generate any acceptable models, excluding the possibility of chance correlations (data not shown).

The analysis of our QSAR models provided insights into the relationship between the hERG channel and its blockers.

The descriptors selected by our best model were: diameter, a_nO, PEOE_VSA-6, KierFlex, a_acid, PC-, SlogP and SMR_VSA6 (Supplementary Table 2A, Supporting Information). These normalized descriptors were plotted against hERG activity. Supplementary Figure 3, Supporting Information, shows that the hERG pIC_{50} value increases with the diameter, KierFlex, and SlogP but decreases with the PEOE_VSA-6 and a_acid descriptors. This is in agreement with generally accepted observations that a hERG inhibitor needs to have one or two hydrophobic moieties to participate in π-stacking interactions with Phe656 in the hERG channel.[61] Hence, high lipophilicity (SlogP) increases the hERG liability. Moreover, basic compounds tend to be hERG active, while acidic groups mitigate hERG activity.[29] Hence, Supplementary Figure 3C, Supporting Information, shows that most hERG active compounds do not have acidic atoms (a_acid = 0), and moreover for those compounds that contain acidic atoms, a higher number of acidic atoms (a_acid) leads to lower hERG liability. In addition, a hERG inhibitor should be long and flexible so that it can orient itself along the pore axis of the hERG channel, with its lipophilic end facing the cytosol and its polar tail facing the selectivity filter area of the hERG pore domain. Therefore, compounds with high diameter and KierFlex value have a high hERG liability. In conclusion, according to the descriptors selected in our QSAR models, a potential hERG inhibitor needs to be long and flexible, possesses hydrophobic moieties interacting with Phe656, occupies a positive charge interacting with Tyr652, and has a polar tail interacting with Thr623 and Val625. All these structural features reflect the general architecture of the hERG channel.

Data Set III contains compounds studied by Ekins et al.,[32] and the generated kNN models were used to predict their hERG pIC_{50} values. We found that the correlation is poor with $R^2 = 0.21$, and overall the predicted values were much higher than the experimental values reported by Ekins et al. Some of the compounds in Data Set III were found to have very different pIC_{50} values than those published by other groups.[1,5,11,27−32] After careful evaluation of these data, we found the Ekins data has much lower pIC_{50} values,[1,5,11,27−32] in agreement with our predictions. Therefore, we think the data from other groups are more trustworthy in our case. For the eight compounds among them, if we use the data from other groups (Supplementary Table 3, Supporting Information), the R^2 for their predicted and corrected experimental values was increased from 0.21 to 0.66.

kNN Classification Models for Discrimination of hERG Blockers. Data Sets I and II were used to generate classification models for prediction of hERG liability. The best classification model used four descriptors. They were SMR_VSA4, radius, GCUT_SMR_0, and GCUT_SLOGP_3. The meaning of these descriptors is described in Supplementary Table 2B, Supporting Information. It demonstrated that polarizability (SMR_VSA4, GCUT_SMR_0), lipophilicity (GCUT_SLOGP_3), and size (radius) are properties highly correlated with the hERG liability. The domperidone (Data Set I) and seven compounds (bismuth, bismuth subsalicylate, flucytosine, hydroxyurea, methadyl acetate, modafinil, and valganciclvir) from Data Set II were characterized as strong outliers with significant dissimilarity to the training data set based on our applicability domain.[62,63] The model selected for the hERG blocker classification had 389 compounds in the training set, including 111 actives and 278 inactives. The recall rate (sensitivity or true positive: the proportion of hERG-active inhibitors that were correctly identified) was 83.8%, as were 226 inactives (specificity = 81.3%) correctly classified.

The confusion matrix of the model is shown in Supplementary Table 4A, Supporting Information. The ability of the model to correctly predict whether a compound is hERG active or not (accuracy) is 82.0%. The model was applied to the test data set containing 101 compounds, where 35 were actives and 66 were inactives. Among those, 82.7% of the actives and 83.3% of the inactives were correctly predicted. Moreover, all hERG inhibitors with pIC_{50} higher than 5.5 were correctly classified except loratadine (claritin). Different hERG pIC_{50} values were reported for loratadine, with 6.70 by Cavalli,[1] 5.64 by Song,[11] and 5.20 by Gavaghan.[27] According to the Canadian FDA report,[64] loratadine does not significantly block hERG potassium channels under the same in vitro conditions for terfenadine, which showed a strong hERG inhibition.[64] Furthermore, loratadine does not have a long shape as other highly active hERG inhibitors do. Maybe this is why it was not accurately predicted by our kNN classification model. In order to exclude the possibility of chance correlations, Y-randomization experiments were executed, and it did not yield any model that met our requirements (data not shown).

3D Pharmacophore Modeling. Different pharmacophore annotation schemes (Unified, PCHD, and PPCH) were applied in the pharmacophore elucidation module of MOE to search the best 3D pharmacophore models. The parameters used to select the model were: number of active molecules that match the query (coverage), the probability that the accuracy would be generated by chance alone (P), the accuracy of the query in matching the active compounds (acc1), and the accuracy of the query not matching the inactive compounds (acc0).

For the hERG blockers with basic charged moieties, the 8 most active compounds with pIC_{50} values > 6.9 and the 191 inactive compounds were employed to create the 3D pharmacophore. A model generated using the "unified" pharmacophore scheme was characterized as the most reasonable by its high coverage (8.00), low P (-2.75), and high acc1 (1) and acc0 (0.67) values. In order to refine the pharmacophore model, features obtained in this model were compared with those obtained using the pharmacophore consensus method, where the conformation alignment of the eight active compounds were employed to generate pharmacophore features. Subsequently, the radius of each pharmacophore feature was manually modified until it could cover as many hERG actives as possible and be able to distinguish the actives from inactives. The final query has six structural features (Figure 3A), consisting of four hydrophobic centroids, a cationic atom or hydrogen donor, and an aromatic or planar ring system. The mapping between the conformations of the eight active compounds in the training data set and the pharmacophore model is shown in Figure 3B. A cationic atom/hydrogen-bond donor (F4[Cat/Don]) and a hydrophobic atom (F5 [Hyd]) were features in our model that were found most important for the hERG-binding activity, in agreement with other reports.[1,13] Moreover, the hydrogen-bond acceptor, one of the pharmacophore features identified by Goldman,[13] has a position similar to our F3 hydrophobic centroid (F3 Hyd); this position is between the positively charged nitrogen and the aromatic ring. However, due to the differences between our active compounds in the training data set and those of Goldman's, the acceptor was not suggested in our case as a common feature by either the pharmacophore elucidation module or the consensus method. Instead, a F3 (Hyd) position was generated to improve the ability of the model to distinguish the actives from inactives. The quality of our pharmacophore models was further assessed using the
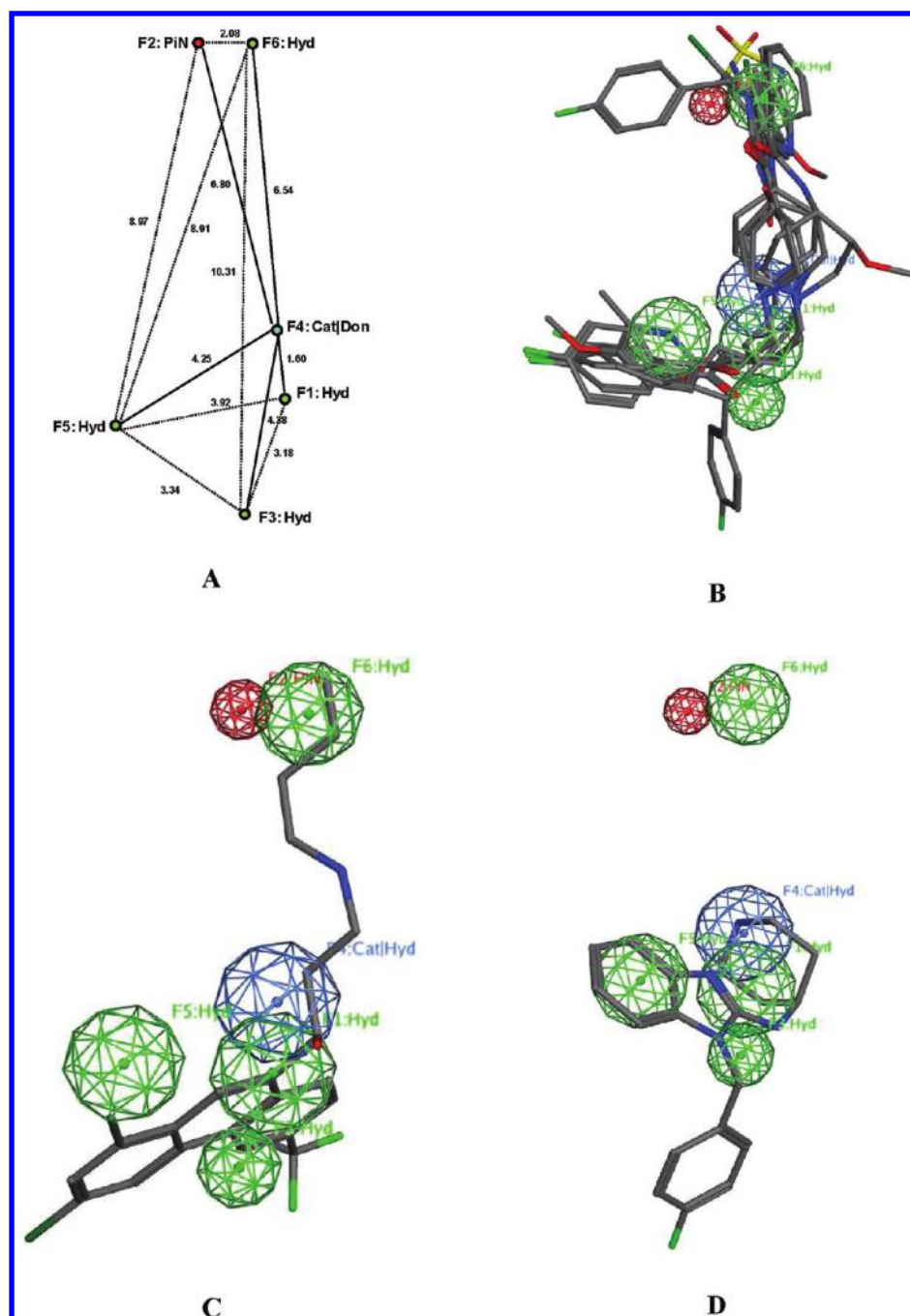
2953

dx.doi.org/10.1021/ci200271d |J. Chem. Inf. Model. 2011, 51, 2948–2960

Figure 3. (A) 3D pharmacophore model for compounds with basic moieties. Hyd: hydrophobic centroid; Cat: cationic atom; Don: hydrogen-bond donor; and PiN: aromatic or planar ring system. (B) Mapping of the eight active compounds in the training data set onto the 3D pharmacophore model. (C) Structure of *n*-desbutylhalofantrine fitted to the pharmacophore model. (D) Structure of tecastemizole fitted to the pharmacophore model.

pharmacophore search function for virtual screening. The con-fusion matrix of the model is shown in Supplementary Table 4B, Supporting Information. Additionally, of the highly active hERG blockers with a $pIC_{50} > 6.9$ ($n = 26$), the prediction accuracy (the proportion of the total number of predictions that were correct) is 96.2%, and the recall rate (sensitivity or true positive) is 92.3%. Among them, tecastemizole and *n*-desbutylhalofantrine ($pIC_{50} > 6.5$) were the only two compounds falsely predicted to be inactive (false negatives), probably because neither of them had an aromatic nor planar ring in the F2 position and a hydrophobic center

at either the F3 or F6 position (*n*-desbutylhalofantrine: Figure 3C; tecastemizole: Figure 3D). This low rate of false negative is what we need in hERG modeling (and other toxicity prediction) because we want to miss as few toxic compounds as possible in drug develop-ment to avoid potential toxic medications. When applied to the external Data Set III, 15 out of 24 hERG blockers with the amine moiety were correctly identified.

For the compounds without basic moieties, 10 active hERG blockers and 160 inactives (Data Set VII) were compiled as a training set for the generation of the pharmacophore model.
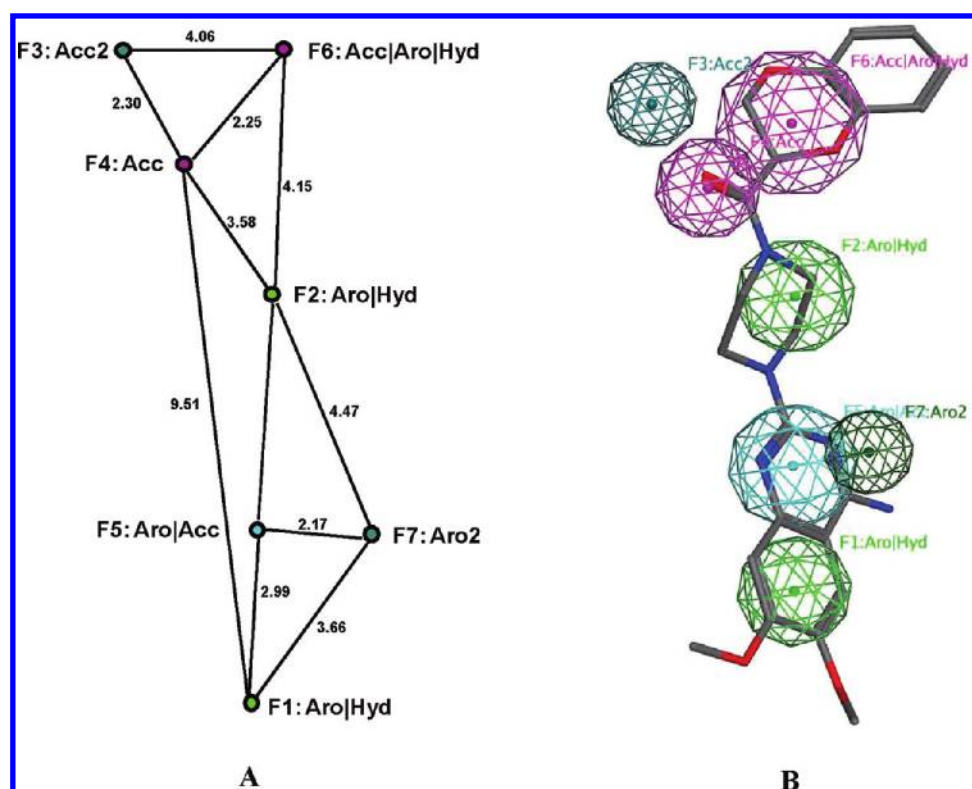
Figure 4. (A) 3D pharmacophore model for compounds without basic moieties. Acc: hydrogen-bond acceptor; Acc2: hydrogen-bond acceptor projected point; Aro: aromatic ring center; Aro2: aromatic ring center projected point; and Hyd: hydrophobic region. (B) Conformation of doxazosin mapped onto our seven-point pharmacophore model.

The PCHD was identified as the best pharmacophore scheme. The model selected for prediction had high coverage (9.00), low P (-2.26), high acc1 (0.90), and high acc0 (0.66) values. The final model has seven pharmacophore features (Figure 4). Among them, the hydrogen-bond acceptor and hydrophobic/aromatic features were also reported as being important for uncharged hERG blockers by Aronov et al.[14] Figure 4B showed the ketone oxygen of doxazosin matches the hydrogen-bond acceptor feature and its projected point, while its quinazoline, piperazine, and benzodioxine groups provide four hydrophobic features and one aromatic ring center projection point. Among the actives, 80% of the 10 compounds with pIC_{50} values > 5.5 matched the pharmacophore model. Data Sets VI and VII were searched using the pharmacophore query. The confusion matrix of the model is shown in Supplementary Table 4C, Supporting Information. Although the total accuracy and specificity are reasonable (75.8% and 80.0%, respectively), the recall rate (sensitivity or true positives) is relatively low (55.9% or 19 actives out of 34). This is probably due to the small number of active compounds used for pharmacophore generation in comparison to large number of inactive compounds (10 vs 160).

Consensus Results using kNN Classification and 3D Pharmacophore Models. Based on the developed kNN classification method and 3D pharmacophore models, a consensus method was suggested for predicting the activity of the hERG blockers. For toxicity prediction, we especially prefer the maximum coverage of the real hERG inhibitors (high sensitivity) in order to include as many hERG-actives as possible to avoid potential toxic agents. We also prefer the negative predictive value (NPV) over 95% so that only a very small portion of toxic agents (<5%) will

be forwarded to clinical use. According to this conception, the hERG active compounds were those identified as positives by either kNN or pharmacophore methods as shown in Table 2A. Table 2B demonstrates that the consensus prediction using these two methods led to a dramatically improved recall rate (sensitivity = 91.8%) of hERG blockers, with a slight increase of the prediction accuracy compared to pharmacophore modeling (70.3% vs 67.5%). NPV of consensus model was augmented to 94.8%. Therefore, this consensus model is able to cover 91% of hERG blockers, whereas it achieves 95% prediction accuracy for non-hERG blockers. Unfortunately, the specificity slightly decreased due to less coverage of the hERG-inactive agents. On the other hand, if we define the hERG active compounds as those identified as positives by both kNN and pharmacophore methods (data not shown), the accuracy, sensitivity, and specificity are obtained as 83.1%, 69.4% and 88.9%, respectively. Of note, with the new scheme we proposed (Supplementary Table 6, Supporting Information), we achieved 90.2% in accuracy, 89.5% in sensitivity, and 90.5% in specificity (Supplementary Table 4E, Supporting Information) for a majority of data sets (77.6%) consensually predicted by both kNN and pharmacophore methods. It implies that if a compound receives consensus prediction by both kNN and pharmacophore modeling, we are ~90% confident to believe this prediction is accurate. Finally, our kNN QSAR method could achieve the best balance in terms of the above three parameters (Table 2B). The confusion matrix of the consensus model is shown in Supplementary Table 4D, Supporting Information.

Homology Modeling of hERG Channel. Analogous to the Kv family of potassium channels, hERG contains a pore domain that

2955

dx.doi.org/10.1021/ci200271d |*J. Chem. Inf. Model.* 2011, 51, 2948–2960

is involved in drug binding. Electrophysiology experiments have shown that the pore domain must be in the open state for compounds to enter.[19,65] However, once a compound is in the cavity region, the four S6 helices can close to various degrees, depending on the bound ligand. To address the flexibility of the hERG channel, multiple homology models were generated to represent different open–close states. The binding energies for a diverse set of hERG blockers were calculated for each model and were then validated (via PLS regression) with the experimental pIC_{50} values of these blockers to determine which model best explains the conformation of the hERG channel in the bound form.

The hERG homology models of close and open states based on the KcsA and MthK structures, along with four additional models representing the different open–close intermediate states of hERG (Supplementary Figure 5, Supporting Information), were constructed and optimized. Compared with the close state of hERG, S6 helices bend approximately 40° at Gly648 toward solvents, thus opening up a channel for ligands entry. Intermediate states are predicted states in between fully opened and close states by bending S6 helices to different degrees. The question of whether these models can represent the hERG channel in the biological state was investigated by docking the 12 hERG blockers into the homology models and correlating their binding free energetic terms with pIC_{50} values. The docking poses were selected according to the pharmacophore match between the ligands and those relevant residues based on the mutagenesis data in Table 1A [19,66,67] of drugs blocking the hERG channel.

First, we investigate the energy values of the MacroModel calculation for the ligand and hERG complex based on the closed state (from KcsA) (Table 1B). We used SIMCA[68] to detect the outliers and established the linear correlation between the experimental and predicted pIC_{50} values of the hERG blockers. Supplementary Figure 2A and B, Supporting Information, demonstrated the contribution of all descriptors to IC_{50} using the

95% tolerance interval. PLS analysis was performed and yielded a model (Figure 5) with 3 principle components for the 12 compounds with $q^2 = 0.68$, $R^2 = 0.81$, and root mean squared error (RMSE) of 0.62.

$$pIC_{50} = 0.0282843*\Delta stretch + 0.0180008*\Delta bend + 0.00927848*\Delta torsion - 0.00771981*\Delta vdW + 0.00153115*\Delta electrostatic - 61.0522$$

(2)

Using eq 2, the pIC_{50} values of all compounds were predicted within one log unit of the experimental value (Table 1B). Descriptors used in eq 2 are shown in the VIP plot[60] (Supplementary Figure 2C, Supporting Information) according to their capacity in explaining the hERG activity. Both bonded (bend, torsion, and stretching) and nonbonded (electrostatic and van der Waals) interactions have VIP values > 0.8, which means that both types of interactions play an important role in the formation of the ligand–hERG complex. For the nonbonded term, the electrostatic interaction has a larger impact on binding than the van der Waals interaction. The larger the electrostatic interaction between the protein and ligand, the higher the hERG inhibition is. This is consistent with the observation that π-stacking between the aromatic groups of the ligand with Phe656 and Tyr652 of hERG is required for strong binding.[19,66]

We also investigated whether the pIC_{50} values of the hERG blockers can be correctly predicted using hERG open–close intermediate state homology models for docking and binding energy calculation. Unfortunately, no acceptable PLS regression models were found; therefore, only the homology model of hERG closed form which was based on KcsV structure was used for modeling the hERG pore domain. The stereochemical quality of the homology model was checked using PROCHECK.[69] No distorted dihedral angles were detected in the Ramachandran plot (Supplementary Figure 4, Supporting Information), and 94.4% of the residues in this homology model was in the most favored region, while 5.6% was in the allowed regions.

For the majority of these 12 compounds with known pIC_{50}, the docking poses with the highest GOLD scores matched the reported hERG mutagenesis data (Table 1A).[19–26] Figure 6 demonstrates that the binding mode of ibutilide which agrees with its experimental mutagenesis data (Table 1A). It has its phenyl ring as a hydrophobic group close to Tyr652 and Phe656. Tyr652 and Phe656 are two residues located in the S6 transmembrane domain of the hERG channel and have been shown in the site-directed mutagenesis[19,66,67] to be critical in ligand binding. The basic nitrogen atom of ibutilide makes a cation–π

Table 2A. Consensus Scheme of Classifying hERG-Actives and -Inactives Using kNN QSAR and 3D Pharmacophore Modeling Methods[a]

kNN	pharmacophore	consensus
1	1	1
0	1	1
1	0	1
0	0	0

[a] Note: "1" indicates the prediction as hERG-active, and "0" represents the prediction as hERG-inactive.

Table 2B. Comparison of kNN QSAR, 3D Pharmacophore, And Consensus Models in Classification of hERG-Actives and -Inactives[a]

method	hERG activity	N_{exp}	N_{pred}	accuracy	sensitivity	specificity	NPV
kNN classification	active	147	122				
	inactive	351	281	82.2%	83.0%	80.1%	91.8%
pharmacophore model	active	147	115				
	inactive	351	221	67.5%	78.2%	63.0%	87.4%
consensus method	active	147	135				
	inactive	351	220	70.3%	91.8%	62.7%	94.8%

[a] Note: NPV: negative predictive value; N_{exp} is the number of compounds experimentally observed; and N_{pred} is the number of active or inactive compounds correctly predicted.
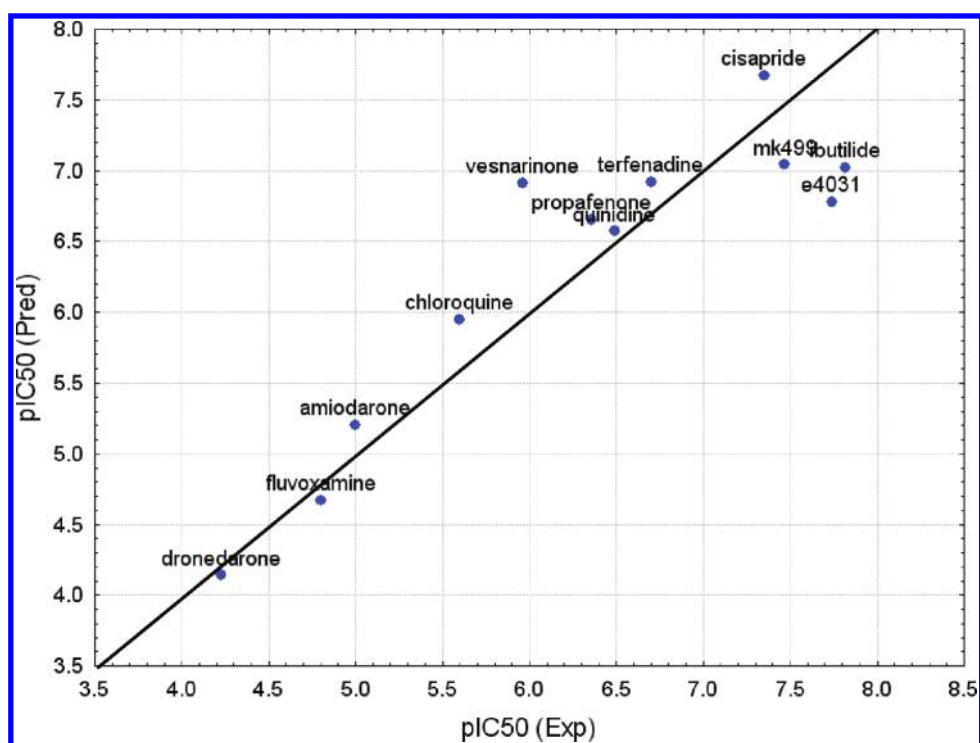
2956

dx.doi.org/10.1021/ci200271d |*J. Chem. Inf. Model.* 2011, 51, 2948–2960

Figure 5. Prediction of pIC_{50} for 12 compounds with docking poses selected based on the mutagenesis data.
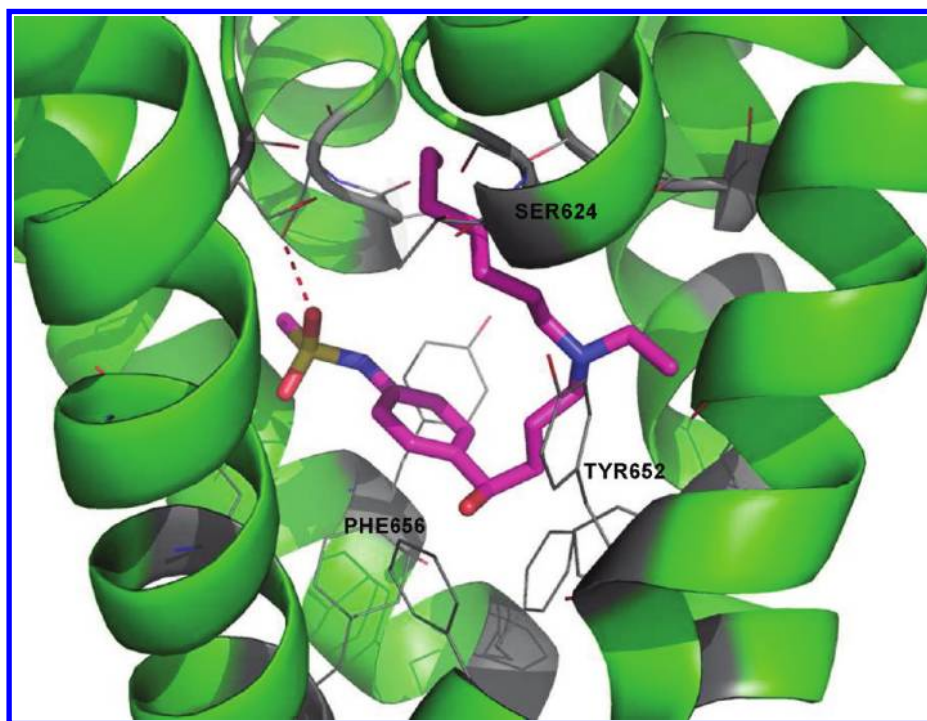

Figure 6. The docked pose of ibutilide (in magenta sticks) in the cavity of the hERG channel (green ribbons). It interacts strongly with residues Tyr652, Phe656, and Ser624 (displayed in line), which is in agreement with the mutagenesis data.

interaction with Tyr652. Its sulfonyl group acts as an acceptor and forms a hydrogen bond with the hydroxyl group of Ser624; interaction with this residue is reported to have pronounced effects on hERG blocking by ibutilide.[70,71] These consistencies validated our homology model, and our ability to predict the correct binding mode using the homology models is crucial for guiding the synthesis of inhibitors with less hERG liability.

Enrichment Analysis and Comparison of QSAR, Pharmacophore, and Molecular Docking Methods in Ranking the hERG Inhibitors. To rank the hERG channel blockers, three
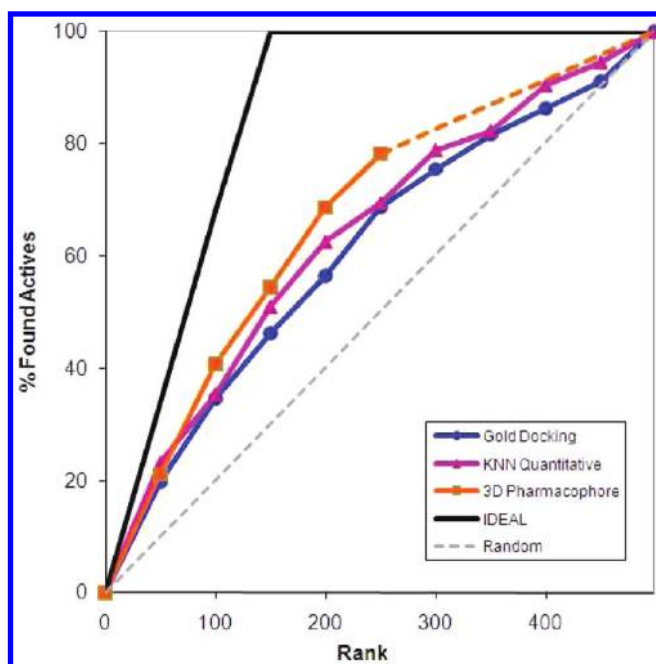
Figure 7. The enrichment plot of hERG inhibitors obtained from kNN QSAR, 3D pharmacophore, and GOLD docking methods. The gray dashed line represents random screening. The *X*-axis shows the compound ranking. The *Y*-axis shows the percentage of true actives predicted. For pharmacophore modeling, the program (MOE) only gave ranking for the top 248 hits which were identified as hERG-active. So the rest were represented with the orange dashed line for completeness.

methods were employed: kNN QSAR approach, 3D pharmacophore modeling, and homology-based GOLD docking. With the kNN method, the potency of hERG inhibitors was estimated using the predicted pIC_{50} values. For the 3D pharmacophore models, the root-mean-squared distance (rmsd) between the query features and their matching ligand annotation points was used as a criterion to assess the fit of a molecule to the pharmacophore features required for hERG toxicity. With the molecular docking method, we docked flexible ligand to rigid receptor based on the finding that only minor side-chain movements could occur on Phe656 and Tyr652 using flexible receptor docking by GOLD, and these local rearrangements could not affect the generation of docking poses (data not shown). The movements of these two aromatic residues indicate their critical role in ligand binding, which was consistent to the finding by Durdagi et al.[18] as well as our pharmacophore modeling. GOLD fitness was used as a parameter to describe the interaction between the ligand and the hERG channel.

Enrichment[36] was used to evaluate the ability of these three methods in ranking the 147 hERG inhibitors out of 498 compounds (Data Sets I and II). The percentage of true hERG inhibitors (*Y*-axis) identified using each method was plotted against their ranking. The ideal curve (black line in Figure 7) demonstrated the case where all actives can be recovered in the top 147 hits, and the gray dashed line represents a random screening. In comparison with kNN QSAR and GOLD docking, pharmacophore modeling performed slightly better for identification of top-ranked true hERG inhibitors (Figure 7). Unfortunately, this function on MOE can only be used to rank those compounds predicted as hERG active, as it does not give the score (rmsd) for compounds predicted as inactives (dashed orange line in

Figure 7). As demonstrated, with our pharmacophore model 115 out of 147 hERG blockers (78.2%) are true positives. Moreover, the enrichment plot shows all three methods (kNN, docking, and pharmacophore) have similar performance but significantly better than random screening. Since the mechanisms of each method of modeling are different, they have their respective advantages and disadvantages: kNN machine learning methodology uses 2D molecular descriptors of compounds for its predictions; pharmacophore modeling requires their 3D conformations; whereas molecular docking needs the accurate protein structure (homology model) for the analysis of binding affinities. More importantly, the current docking/scoring is derived using true receptor–ligand binding data, and they usually need more expert knowledge of the modeling system to obtain accurate predictions. Although they have achieved great successes in "specificity-focused" studies (e.g., in virtual screening to eliminate as many true negatives as possible), we suggest that its application to "sensitivity-focused" toxicity predictions (to cover as many true positives as possible) be used with cautions. A strong hERG blocker can be easily interpreted by its interaction with hERG in a docking study, but an active compound can also be perceived as a weak binder due to various reasons, such as the lack of correct receptor conformations and wrong parametrization. If this happens, then a toxic hERG blocker may go to a bedside, which is dangerous and is not what we want. To overcome this problem, we suggest ligand-based methods, such as the kNN and pharmacophore modeling, be used in this study. They require less expert knowledge (e.g., only 2D structures) and take into account both active and inactive hERG inhibitors. Therefore, due to the complementarities between ligand- and structure-based methods, the combined approach would give better results.

■ CONCLUSION

We describe a comprehensive assessment of different approaches to prediction hERG liability by combining, kNN regression and classification, 3D pharmacophore modeling, homology modeling of hERG 3D structures, and molecular docking. Ligand-based approaches, such as kNN and pharmacophore modeling, can be employed as a predictive filter for efficient screening of large chemical databases. Structure-based model of the hERG pore domain was constructed to investigate hERG–inhibitor interactions, and we could appropriately explain the binding of different inhibitors as well as the reported mutagenesis studies. Therefore we are convinced that reliable structural models have been obtained for hERG open and closed forms. In summary, these in silico models of hERG and its blockers are useful and can be employed to efficiently screen hERG liabilities in early lead discovery as well as to guide medicinal synthesis to reduce hERG binding during lead optimization stage.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Sequence alignments, hERG homology models, Ramachandran plot and details of the PLS, kNN, pharmacophore and consensus models. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: shuzhang@mdanderson.org. Telephone: (713) 745-2958.

2958

dx.doi.org/10.1021/ci200271d |*J. Chem. Inf. Model.* 2011, 51, 2948–2960

Author Contributions

§These authors contributed equally.

■ REFERENCES

(1) De Ponti, F.; Poluzzi, E.; Cavalli, A.; Recanatini, M.; Montanaro, N. Safety of non-antiarrhythmic drugs that prolong the QT interval or induce torsade de pointes: An overview. *Drug Safety* **2002**, *25* (4), 263–286.

(2) Pearlstein, R.; Vaz, R.; Rampe, D. Understanding the structure-activity relationship of the human ether-a-go-go-related gene cardiac K+ channel. A model for bad behavior. *J. Med. Chem.* **2003**, *46* (11), 2017–2022.

(3) Brown, A. M. Drugs, hERG, and sudden death. *Cell Calcium* **2004**, *35* (6), 543–547.

(4) Netzer, R.; Ebneth, A.; Bischoff, U.; Pongs, O. Screening lead compounds for QT interval prolongation. *Drug Discovery Today* **2001**, *6* (2), 78–84.

(5) Yoshida, K.; Niwa, T. Quantitative structure-activity relationship studies on inhibition of HERG potassium channels. *J. Chem. Inf. Model.* **2006**, *46* (3), 1371–8.

(6) 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* (5), 455–459.

(7) Shen, M. Y.; Su, B. H.; Esposito, E. X.; Hopfinger, A. J.; Tseng, Y. J. A comprehensive support vector machine binary hERG classification model based on extensive but biased end point hERG data sets. *Chem. Res. Toxicol.* **2011**, *24* (6), 934–49.

(8) Su, B. H.; Shen, M. Y.; Esposito, E. X.; Hopfinger, A. J.; Tseng, Y. J. In silico binary classification QSAR models based on 4D-fingerprints and MOE descriptors for prediction of hERG blockage. *J. Chem. Inf. Model.* **2010**, *50* (7), 1304–18.

(9) Ekins, S.; Chekmarev, D. S.; Kholodovych, V.; Balakin, K. V.; Ivanenkov, Y.; Welsh, W. J. Shape signatures: New descriptors for predicting cardiotoxicity in silico. *Chem. Res. Toxicol.* **2008**, *21* (6), 1304–1314.

(10) Keseru, G. M. Prediction of hERG potassium channel affinity by traditional and hologram qSAR methods. *Bioorg. Med. Chem. Lett.* **2003**, *13* (16), 2773–5.

(11) Song, M.; Clark, M. Development and Evaluation of an in Silico Model for hERG Binding. *J. Chem. Inf. Model.* **2006**, *46* (1), 392–400.

(12) Ekins, S.; Crumb, W. J.; Sarazan, R. D.; Wikel, J. H.; Wrighton, S. A. Three-dimensional quantitative structure-activity relationship for inhibition of human ether-a-go-go-related gene potassium channel. *J. Pharmacol. Exp. Ther.* **2002**, *301* (2), 427–434.

(13) Aronov, A. M.; Goldman, B. B. A model for identifying HERG K+ channel blockers. *Bioorg. Med. Chem.* **2004**, *12* (9), 2307–15.

(14) Aronov, A. M. Common pharmacophores for uncharged human ether-a-go-go-related gene (hERG) blockers. *J. Med. Chem.* **2006**, *49* (23), 6917–21.

(15) Stansfeld, P. J.; Gedeck, P.; Gosling, M.; Cox, B.; Mitcheson, J. S.; Sutcliffe, M. J. Drug block of the hERG potassium channel: insight from modeling. *Proteins: Struct., Funct., Bioinf.* **2007**, *68* (2), 568–580.

(16) Rajamani, R.; Tounge, B. A.; Li, J.; Reynolds, C. H. A two-state homology model of the hERG K+ channel: application to ligand binding. *Bioorg. Med. Chem. Lett.* **2005**, *15* (6), 1737–1741.

(17) Taboureau, O.; Li, Q. Y.; Jorgensen, F. S.; Oprea, T.; Brunak, S. hERG classification model based on a combination of support vector machine method and GRIND descriptors. *Mol. Pharmaceutics* **2008**, *5* (1), 117–127.

(18) Durdagi, S.; Duff, H. J.; Noskov, S. Y. Combined receptor and ligand-based approach to the universal pharmacophore model development for studies of drug blockade to the hERG1 pore domain. *J. Chem. Inf. Model.* **2011**, *51* (2), 463–74.

(19) Mitcheson, J. S.; Chen, J.; Lin, M.; Culberson, C.; Sanguinetti, M. C. A structural basis for drug-induced long QT syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (22), 12329–12333.

(20) Kamiya, K.; Niwa, R.; Mitcheson, J. S.; Sanguinetti, M. C. Molecular determinants of HERG channel block. *Mol. Pharmacol.* **2006**, *69* (5), 1709–16.

(21) Ridley, J. M.; Milnes, J. T.; Witchel, H. J.; Hancox, J. C. High affinity HERG K(+) channel blockade by the antiarrhythmic agent dronedarone: resistance to mutations of the S6 residues Y652 and F656. *Biochem. Biophys. Res. Commun.* **2004**, *325* (3), 883–91.

(22) Milnes, J. T.; Crociani, O.; Arcangeli, A.; Hancox, J. C.; Witchel, H. J. Blockade of HERG potassium currents by fluvoxamine: incomplete attenuation by S6 mutations at F656 or Y652. *Br. J. Pharmacol.* **2003**, *139* (5), 887–98.

(23) Witchel, H. J.; Dempsey, C. E.; Sessions, R. B.; Perry, M.; Milnes, J. T.; Hancox, J. C.; Mitcheson, J. S. The low-potency, voltage-dependent HERG blocker propafenone--molecular determinants and drug trapping. *Mol. Pharmacol.* **2004**, *66* (5), 1201–12.

(24) Kamiya, K.; Mitcheson, J. S.; Yasui, K.; Kodama, I.; Sanguinetti, M. C. Open channel block of HERG K(+) channels by vesnarinone. *Mol. Pharmacol.* **2001**, *60* (2), 244–53.

(25) Lees-Miller, J. P.; Duan, Y.; Teng, G. Q.; Duff, H. J. Molecular determinant of high-affinity dofetilide binding to HERG1 expressed in Xenopus oocytes: involvement of S6 sites. *Mol. Pharmacol.* **2000**, *57* (2), 367–74.

(26) Sanchez-Chapula, J. A.; Ferrer, T.; Navarro-Polanco, R. A.; Sanguinetti, M. C. Voltage-dependent profile of human ether-a-go-go-related gene channel block is influenced by a single residue in the S6 transmembrane domain. *Mol. Pharmacol.* **2003**, *63* (5), 1051–8.

(27) Gavaghan, C. L.; Arnby, C. H.; 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* (4), 189–206.

(28) Keseru, G. M. Prediction of hERG potassium channel affinity by traditional and hologram qSAR methods. *Bioorg. Med. Chem. Lett.* **2003**, *13* (16), 2773–2775.

(29) Sun, H. An accurate and interpretable Bayesian classification model for prediction of hERG liability. *ChemMedChem* **2006**, *1* (3), 315–322.

(30) Recanatini, M.; Poluzzi, E.; Masetti, M.; Cavalli, A.; De Ponti, F. QT prolongation through hERG K(+) channel blockade: current knowledge and strategies for the early prediction during drug development. *Med. Res. Rev.* **2005**, *25* (2), 133–66.

(31) Pearlstein, R. A.; Vaz, R. J.; Kang, J.; Chen, X.-L.; Preobrazhenskaya, M.; Shchekotikhin, A. E.; Korolev, A. M.; Lysenkova, L. N.; Miroshnikova, O. V.; Hendrix, J.; Rampe, D. Characterization of HERG potassium channel inhibition using CoMSiA 3D QSAR and homology modeling approaches. *Bioorg. Med. Chem. Lett.* **2003**, *13* (10), 1829–1835.

(32) Ekins, S.; Balakin, K. V.; 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* (17), 5059–5071.

(33) Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **2008**, *36* (Database Iss), D901–D906.

(34) Morgan, T. K., Jr.; Sullivan, M. E. An overview of class III electrophysiological agents: a new generation of antiarrhythmic therapy. *Prog. Med. Chem.* **1992**, *29*, 65–108.

(35) 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* (2−4), 241–253.

(36) Du-Cuny, L.; Song, Z.; Moses, S.; Powis, G.; Mash, E. A.; Meuillet, E. J.; Zhang, S. Computational modeling of novel inhibitors targeting the Akt pleckstrin homology domain. *Bioorg. Med. Chem.* **2009**, *17* (19), 6983–6992.

(37) Sharaf, M. A.; Illman, D. L.; Kowalski, B. R. *Chemical Analysis*, 1986; Vol. 82: Chemometrics, p 332.

(38) 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* (5), 1984–1995.

(39) Zheng, W.; Tropsha, A. Novel variable selection quantitative structure-property relationship approach based on the K-nearest-neighbor principle. *J. Chem. Inf. Comput. Sci.* **2000**, *40* (1), 185–194.

(40) Testai, L.; Bianucci, A. M.; Massarelli, I.; Breschi, M. C.; Martinotti, E.; Calderone, V. Torsadogenic cardiotoxicity of antipsychotic drugs: A structural feature, potentially involved in the interaction with cardiac HERG potassium channels. *Curr. Med. Chem.* **2004**, *11* (20), 2691–2706.

(41) Ok, D.; Li, C.; Abbadie, C.; Felix, J. P.; Fisher, M. H.; Garcia, M. L.; Kaczorowski, G. J.; Lyons, K. A.; Martin, W. J.; Priest, B. T.; Smith, M. M.; Williams, B. S.; Wyvratt, M. J.; Parsons, W. H. Synthesis and SAR of 1,2-trans-(1-hydroxy-3-phenylprop-1-yl)cyclopentane carboxamide derivatives, a new class of sodium channel blockers. *Bioorg. Med. Chem. Lett.* **2006**, *16* (5), 1358–1361.

(42) Sisko, J. T.; Tucker, T. J.; Bilodeau, M. T.; Buser, C. A.; Ciecko, P. A.; Coll, K. E.; Fernandes, C.; Gibbs, J. B.; Koester, T. J.; Kohl, N.; Lynch, J. J.; Mao, X.; McLoughlin, D.; Miller-Stein, C. M.; Rodman, L. D.; Rickert, K. W.; Sepp-Lorenzino, L.; Shipman, J. M.; Thomas, K. A.; Wong, B. K.; Hartman, G. D. Potent 2-[(pyrimidin-4-yl)amine]-1,3-thiazole-5-carbonitrile-based inhibitors of VEGFR-2 (KDR) kinase. *Bioorg. Med. Chem. Lett.* **2006**, *16* (5), 1146–1150.

(43) Friesen, R. W.; Ducharme, Y.; Ball, R. G.; Blouin, M.; Boulet, L.; Cote, B.; Frenette, R.; Girard, M.; Guay, D.; Huang, Z.; Jones, T. R.; Laliberte, F.; Lynch, J. J.; Mancini, J.; Martins, E.; Masson, P.; Muise, E.; Pon, D. J.; Siegl, P. K. S.; Styhler, A.; Tsou, N. N.; Turner, M. J.; Young, R. N.; Girard, Y. Optimization of a Tertiary Alcohol Series of Phospho-diesterase-4 (PDE4) Inhibitors: Structure-Activity Relationship Related to PDE4 Inhibition and Human Ether-a-go-go Related Gene Potassium Channel Binding Affinity. *J. Med. Chem.* **2003**, *46* (12), 2413–2426.

(44) Labute, P.; Williams, C.; Feher, M.; Sourial, E.; Schmidt, J. M. Flexible Alignment of Small Molecules. *J. Med. Chem.* **2001**, *44* (10), 1483–1490.

(45) Lauri, G.; Bartlett, P. A. CAVEAT: a program to facilitate the design of organic molecules. *J. Comput.-Aided Mol. Des.* **1994**, *8* (1), 51–66.

(46) Doyle, D. A.; Morais Cabral, J.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* **1998**, *280* (5360), 69–77.

(47) Jiang, Y.; Lee, A.; Chen, J.; Cadene, M.; Chait Brian, T.; MacKinnon, R. Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **2002**, *417* (6888), 515–22.

(48) Pearson, W. R. *ALIGN*, 2.0u66; The University of Virginia: Charlottesville, VA, 1997.

(49) Jiang, Y.; Lee, A; Chen, J.; Cadene, M.; Chait, B. T.; MacKinnon, R. The open pore conformation of potassium channels. *Nature (London, U. K.)* **2002**, *417* (6888), 523–526.

(50) *MODELLER* [9v6], UCSF: San Francisco, CA, 2009.

(51) Bower, M. J.; Cohen, F. E.; Dunbrack, R. L., Jr. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *J. Mol. Biol.* **1997**, *267* (5), 1268–1282.

(52) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60* (12, Pt. 1), 2126–2132.

(53) Case David, A.; Cheatham Thomas, E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz Kenneth, M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods Robert, J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–88.

(54) Wang, J. M.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21* (12), 1049–1074.

(55) Jin, L.; Wu, Y. Molecular mechanism of the sea anemone toxin ShK recognizing the Kv1.3 channel explored by docking and molecular dynamic simulations. *J. Chem. Inf. Model.* **2007**, *47* (5), 1967–72.

(56) *GOLD* [3.2], CCDC: Cambridge, U.K., 2007.

(57) Aronov, A. M. Predictive in silico modeling for hERG channel blockers. *Drug Discovery Today* **2005**, *10* (2), 149–55.

(58) Yang, H.; Shen, Y.; Chen, J.; Jiang, Q.; Leng, Y.; Shen, J. Structure-based virtual screening for identification of novel 11beta-HSD1 inhibitors. *Eur. J. Med. Chem.* **2009**, *44* (3), 1167–71.

(59) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26* (16), 1752–80.

(60) Eriksson, L.; Johansson, N.; Kettaneh-Wold, N.; Wold, S. *Multi-and Megavariate Data Analysis: Principles and Applications*. Umetrics Academy: Umea, Sweden, 1999.

(61) Aronov, A. M. Predictive in silico modeling for hERG channel blockers. *Drug Discovery Today* **2005**, *10* (2), 149–155.

(62) Kozikowski, A. P.; Sun, H.; Brognard, J.; Dennis, P. A. Novel PI Analogues Selectively Block Activation of the Pro-survival Serine/ Threonine Kinase Akt. *J. Am. Chem. Soc.* **2003**, *125* (5), 1144–1145.

(63) Friesner Richard, A.; Banks Jay, L.; Murphy Robert, B.; Halgren Thomas, A.; Klicic Jasna, J.; Mainz Daniel, T.; Repasky Matthew, P.; Knoll Eric, H.; Shelley, M.; Perry Jason, K.; Shaw David, E.; Francis, P.; Shenkin Peter, S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–49.

(64) Lapenta, J. R. The loratadine and their 110 adverse effects; Dermagic: Venezuela ; http://www.dermagicjournal.20m.com/year2001/ 11B.html.

(65) Spector, P. S.; Curran, M. E.; Keating, M. T.; Sanguinetti, M. C. Class III antiarrhythmic drugs block HERG, a human cardiac delayed rectifier K+ channel. Open-channel block by methanesulfonanilides. *Circulation research* **1996**, *78* (3), 499–503.

(66) Lees-Miller, J. P.; Duan, Y.; Teng, G. Q.; Thorstad, K.; Duff, H. J. Novel gain-of-function mechanism in K(+) channel-related long-QT syndrome: altered gating and selectivity in the HERG1 N629D mutant. *Circ. Res.* **2000**, *86* (5), 507–13.

(67) Chen, J.; Seebohm, G.; Sanguinetti Michael, C. Position of aromatic residues in the S6 domain, not inactivation, dictates cisapride sensitivity of HERG and eag potassium channels. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (19), 12461–6.

(68) *SIMCA*, Umetrics AB: Umea, Sweden, 2011.

(69) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26* (2), 283–91.

(70) Perry, M.; De Groot, M. J.; Helliwell, R.; Leishman, D.; Tristani-Firouzi, M.; Sanguinetti, M. C.; Mitcheson, J. Structural determinants of HERG channel block by clofilium and ibutilide. *Mol. Pharmacol.* **2004**, *66* (2), 240–249.

(71) Perry, M.; Stansfeld, P. J.; Leaney, J.; Wood, C.; de Groot, M. J.; Leishman, D.; Sutcliffe, M. J.; Mitcheson, J. S. Drug binding interactions in the inner cavity of hERG channels: molecular insights from structure-activity relationships of clofilium and ibutilide analogs. *Mol. Pharmacol.* **2006**, *69* (2), 509–519.