Insights for Human Ether-a-Go-Go-Related Gene Potassium Channel Inhibition Using Recursive Partitioning and Kohonen and Sammon Mapping Techniques

Sean Ekins,*[,†,#,‡] Konstantin V. Balakin,[§,‡] Nikolay Savchuk,[§] and Yan Ivanenkov[§]

ACT LLC, 601 Runnymede Avenue, Jenkintown, Pennsylvania 19046, Department of Pharmaceutical Sciences, University of Maryland, 20 Penn Street, Baltimore, Maryland 21201, and Chemical Diversity, Inc., 11558 Sorrento Valley Road, San Diego, California 92121

Received January 23, 2006

The human ether-a-go-go related gene (hERG) can be inhibited by marketed drugs, and this inhibition may lead to QT prolongation and possibly fatal cardiac arrhythmia. We have collated literature data for 99 diverse hERG inhibitors to generate Kohonen maps, Sammon maps, and recursive partitioning models. Our aim was to investigate whether these computational models could be used either individually or together in a consensus approach to predict the binding of a prospectively selected test set of 35 diverse molecules and at the same time to offer further insights into hERG inhibition. The recursive partitioning model provided a quantitative prediction, which was markedly improved when Tanimoto similarity was included as a filter to remove molecules from the test set that were too dissimilar to the training set ($r^2 = 0.83$, Spearman rho $= 0.75$, $p = 0.0003$ for the 18 remaining molecules, >0.77 similarity). This model was also used to screen and prioritize a database of drugs, recovering several hERG inhibitors not used in model building. The mapping approaches used molecular descriptors required for hERG inhibition that were not reported previously and in particular highlighted the importance of molecular shape. The Sammon map model provided the best qualitative classification of the test set (95% correct) compared with the Kohonen map model (81% correct), and this result was also superior to the consensus approach. This study illustrates that patch clamping data from various literature sources can be combined to generate valid models of hERG inhibition for prospective predictions.

Introduction

Numerous classes of drugs have been shown to prolong the QT interval, which reflects a slowing of repolarization of the ventricular myocardium.[1,2] Excessive QT interval prolongation can lead to the potentially life-threatening ventricular tachyarrhythmia, torsade de pointes. In cardiac tissue, inhibition of potassium channels is associated with QT interval prolongation.[3,4] The most common potassium channel linked to drug-induced QT interval prolongation is also responsible for the rapid component of the delayed rectifier potassium current (I_{Kr}). The human ether-a-go-go-related gene (hERG) is believed to encode the protein that underlies the delayed rectifier potassium current I_{Kr},[5,6] and many drugs associated with QT interval prolongation have been found to block hERG.[7−9] Drugs such as cisapride, terfenadine, astemizole, sertindole, and grepafloxacin have been withdrawn from the market in recent years in some degree because of cardiovascular toxicity associated with undesirable blockade of this channel. It is therefore important for drug discovery scientists to understand the structural requirements of molecules binding to this potassium channel to avoid potential toxicity.

To date, the in vitro assessment of the drug-mediated interaction with these channels uses various cell systems expressing the hERG channel as well as methods such as patch clamping, radioligand binding, fluorescent probes, and rubidium flux studies. These methods produce data of varying quality and reliability that can be potentially modeled computationally. Understanding the molecular features that confer hERG inhibition activity has therefore become a recent focus of considerable computational and statistical modeling efforts. Previously published studies have generated either numerous quantitative structure−activity relationship (QSAR) or pharmacophore models for hERG. These have aided in a ligand-based understanding of the molecular descriptors or three-dimensional disposition of molecular features that are important for hERG inhibition.[10−23]

A preliminary pharmacophore was derived from a structurally diverse literature set of 15 molecules known to inhibit hERG in patch clamping studies with predominantly HEK 293 cells.[16] This model was also assessed using a test set of 22 molecules including known inhibitors such as antipsychotics and their metabolites that had been tested with hERG-expressing HEK 293 cells. Classes of hERG inhibitors using data from Chinese hamster ovary (CHO) cells expressing the channel were predicted in a computational neural network analysis trained with 244 molecules.[11] A further pharmacophore was derived using the comparative molecular field analysis (CoMFA) method for 31 inhibitors.[20] Recently a comparative molecular similarity analysis (CoMSiA) used 28 molecules with patch clamping data dominated by sertindole analogues tested in CHO cells expressing hERG.[13] The model was also placed in the context of the protein using a homology model based on the MthK channel[13] to illustrate key interactions with Phe 656 and Tyr 652 residues,[113] the same residues as identified by Mitcheson et al. earlier. Attempts at building a QSAR model with a recursive partitioning algorithm and based on patch clamping data for hERG with 66 molecules have also been described along with a diverse test set of 25 molecules.[18] An updated version of this model was also built with 99 literature molecules[18] that has been used recently to rank the 23 sertindole analogues generated by Pearlstein et al.[114] (Spearman's rho $= 0.74$, $p < 0.0001$, $r^2 = 0.53$).[17] This type of model is perhaps less visually interpretable

* To whom correspondence should be addressed. Address: ACT LLC, 601 Runnymede Ave, Jenkintown, PA 19046. Phone: 269-930-0974. Fax: 215-481-0159. E-mail: ekinssean@yahoo.com.
† ACT LLC.
# University of Maryland.
‡ Denotes authors who contributed equally.
§ Chemical Diversity, Inc.

10.1021/jm060076r CCC: $33.50
Published on Web 08/02/2006

Table 1. hERG Training Set Molecules Obtained from Various Literature Sources and Cell Types Used as the Training Set for Computational Model Building[a]

compd	$\log_{10}$ IC_{50} (μM)	cell type	ref	compd	$\log_{10}$ IC_{50} (μM)	cell type	ref
2-hydroxymethylolanzapine	1.06	HEK	16	ibutilide	−1.82	AT-1	45
9-hydroxyrisperidone	0.11	HEK	16	imipramine	0.53	CHO	57
alosetron	0.505	HEK	16	isradipine	1	CHO	11
amiodarone	0	AT-1	32	josamycin	2.01	HEK	43
amitriptyline	1	CHO	33	ketoconazole	0.278	HEK	16
astemizole	−3.04	HEK	34	laam	0.34	HEK	38
azimilide	−0.22	CHO	35	levofloxacin	2.96	CHO	42
bepridil	−0.26	COS	36	lidocaine	2.42	HEK	54
BRL 32872	−1.7	HEK	37	loratadine	−0.76	HEK	58
buprenorphine	0.87	HEK	38	LY97241	−0.82	XO	59
carvedilol	1.02	XO	39	MDL-74156	0.77	HEK	49
cetirizine	1.48*	XO	40	mefloquine	0.75	CHO	60
chlorpheniramine	1.32	XO	41	meperidine	1.87	HEK	38
chlorpromazine	0.167	CHO	33	mesoridazine	−0.495	HEK	16
ciprofloxacin	2.98	CHO	42	methadone	0.99	HEK	38
cisapride	−2.174	HEK	16	methylecgonidine	2.23	HEK	44
clarithromycin	1.52	HEK	43	mexiletine	1	CHO	11
clozapine	−0.49	HEK	16	mibefradil	0.16	COS	36
clozapine *N*-desmethyl	0.65	HEK	16	mizolastine	−0.455	HEK	45
clozapine *N*-oxide	2.12	HEK	16	MK499	−1.49	XO	61
cocaethylene	0.079	HEK	44	morphine	3	HEK	38
cocaine	0.64	HEK	44	moxifloxacin	2.11	CHO	42
codeine	2.48	HEK	38	*N*-desbutylhalofantrine	−1.14	HEK	62
descarboxyl-loratadine	0.26	HEK	45	nicotine	2.39	HEK	16
desipramine	0.14	HEK	16	nifedipine	1.7	HEK	46
desmethyl-astemizole	−3	HEK	34	nitrendipine	1*	COS	63
desmethyl-erythromycin	2.17	HEK	43	norastemizole	−1.56	HEK	34
diltiazem	1.23	HEK	46	olanzapine	−0.63	HEK	16
diphenhydramine	1.43	XO	47	olanzapine desmethyl	1.15	HEK	16
disopyramide	1.96	AT-1	32	oleandomycin	2.53	HEK	43
dofetilide	−1.92	HEK	48	ondansetron	−0.09	HEK	49
dolasetron	1.083	HEK	49	perhexiline	0.89	CHO	64
domperidone	−0.79	CHO	50	pimozide	−1.26	HEK	16
droperidol	−1.49	HEK	51	propafenone	−0.36	HEK	54
E-4031	−1.74	HEK	16	quetiapine	0.76	CHO	65
ebastine	−0.85	XO	52	quinidine	−0.495	AT-1	32, 66
EDDP	1.7	HEK	38	risperidone	−0.829	HEK	16
epinastine	2	XO	53	roxithromycin	1.56	HEK	43
erythromycin	1	HEK	43	sertindole	−1.83	HEK	16
erythromyclamine	2.44	HEK	43	sildenafil	0.518	HEK	16
fentanyl	0.25	HEK	38	sparfloxacin	1.25	CHO	42
fexofenadine	1.33	HEK	45	sulfamethoxazole	4	CHO	11
flecainide	0.59	HEK	54	terfenadine	−0.67	HEK	16
gatifloxacin	2.11	CHO	42	terikalant	−0.6	XO	67
glimepiride	2.7	neuroblastoma cells	55	thioridazine	−1.479	HEK	16
glyburide	1.87	neuroblastoma cells	55	trimethoprim	2.38	CHO	11
granisetron	0.57	HEK	49	verapamil	−0.84	HEK	46
grepafloxacin	1.64	CHO	42	vesnarinone	0.041	HEK	68
halofantrine	−0.708	CHO	56	ziprasidone	−0.9	HEK	16
haloperidol	−1.57	HEK	16				

[a] HEK = HEK-293. XO = Xenopus oocytes. CHO = Chinese hamster ovary cells. ∗ = maximum concentration studied.

than a pharmacophore because it is based on 2D descriptors, but it enables a very high throughput for scoring virtual libraries of molecules. More recently another molecular descriptor-based approach using a partial least squares algorithm was applied to a 55-molecule training set and found to be predictive for a 13-compound test set ($r^2 = 0.81$).[15] More recent studies include a combined classification and pharmacophore approach,[22] a genetic programming classification model,[21] stepwise regression,[23] consensus of neural network and Bayesian classification models,[14] a support vector machine model,[10] and finally a method using PCA and PLS.[19]

The focus to date has been primarily on individual models of a "global" nature consisting of structurally divergent molecules across several therapeutic targets (antipsychotics, antihistamines, antibiotics, etc.). "Local" models have also been generated around narrow structural series such as the sertindole analogues.[13] For example, publications containing [^{3}H]dofetilide binding data for the $5HT_{2A}$ class of molecules,[24,25] 3-aminopyrrolidinone farnesyltransferase inhibitors,[26] have been used to produce individual pharmacophores that were ultimately combined to suggest common areas of positive ionizable features and hydrophobicity from aromatic rings.[17,27] These ligand-based computational models along with a growing number of homology models[28,29] have provided insights that complement the experimental studies such as site directed mutagenesis.[30,31]

The present study illustrates for the first time the comparison of multiple different modeling approaches including Kohonen map, Sammon map, and recursive partitioning with the same training set to assess whether one or a combination of approaches is preferable. All hERG models were assessed with a large external data set of published molecules and demonstrate good predictivity. In addition, we demonstrate that a similarity measure is of considerable value to limit extrapolations far beyond the training set of the quantitative recursive partitioning model when performing prospective testing. The descriptors selected for the qualitative mapping methods provide further insight into the structural features of hERG inhibitors when compared with those generated by other methods previously. Finally, we have also shown how the recursive partitioning model can be used to rapidly screen a database of known drugs in an attempt to prospectively identify molecules that may bind

to hERG. The selected molecules may be tested in vitro in future to evaluate these predictions.

Methods

Training Set. The literature was used to derive a structurally diverse training set for model building following from previous publications that had collated patch clamping hERG data ($\log_{10}$(IC_{50} in μM) = −3.04 to 4, Table 1).[16,18] Most of these data are from HEK-293 and CHO cells with a small number of data points from Xenopus oocytes (XO) and other cell types. Attempts to restrict the data to predominantly HEK-293 cell data were made to limit IC_{50} variability (seen when compared with XO, as well as when derived from many different laboratories using subtle experimental differences in the patch clamping procedure), which has been described previously by us.[16,18] Although a small number of molecules have been well studied by several laboratories in the same cell type, in the vast majority of cases a single study is taken as representative of the patch clamping IC_{50} because for most molecules this was the only example available. The reader is referred to the original reference for each molecule for details on the standard errors obtained in these studies. No other attempts at data standardization were performed. Initial efforts at model construction without the XO data did not appear to be significantly different. Therefore, the total training set included 99 structures, which were assigned to 3 principal classes: class 0, 34 compounds with $\log_{10} IC_{50} < 0$; class 1, 31 compounds with $\log_{10} IC_{50} = 0-1$; class 2, 34 compounds with $\log_{10} IC_{50} > 1$.

Computational Methods

Recursive Partitioning. The ChemTree recursive partitioning software (GoldenHelix, Bozeman, MT) was run on a Pentium 4 processor. The molecules and $\log_{10}$-transformed IC_{50} data (μM) were imported as an sdf file. ChemTree was used to generate 564 path length molecular descriptors,[69,70] which were used with the $\log_{10} IC_{50}$ data to generate 100 random tree models using some of the descriptors available with the following options: Bonferroni adjusted p value threshold for splits, 0.99; maximal segments, 10; parallel threads, 1; Segmenting algorithm, Approximate, 0 ($n^{1.5}$) and resampling iterations, 10 000. A second tree model was later generated to incorporate the test set molecules and consisted of 134 molecules. ChemTree was used to generate 694 path length molecular descriptors for these molecules, which were used with the $\log_{10} IC_{50}$ data to generate 100 random tree models (using some of the descriptors available) with the same options described above.

Molecular Descriptors. A wide range of different molecular descriptors were calculated for all compounds with the Smart Mining v1.01 (ChemDiv, Inc., San Diego, CA) software tool. These included electronic, topological, spatial, and structural descriptors. A total of more than 150 initial descriptors were calculated for each compound. To reduce the number of descriptors that could contain redundant information, principal component analysis (PCA) was performed. About 90% of the variance could be explained by the first 8−10 PCs. Descriptors maximally contributing to the first significant PCs were selected on the basis of these results as the most relevant, and these descriptors were used as input parameters in all further computational experiments.

Training Set for Mapping Methods. To clean the input data for their subsequent use in mapping experiments, the training set molecules were filtered on the basis of molecular weight (<700) to limit the data to small molecules. With this threshold, the following six molecules were removed: erythromycin, clarithromycin, roxithromycin, josamycin, desmethyl-erythromycin, erythromyclamine. The filtered training set for mapping methods included a total of 93 structures: 32 compounds with $\log_{10} IC_{50} < 0$ (class 0); 29 compounds with $\log_{10} IC_{50} = 0-1$ (class 1); 32 compounds with $\log_{10} IC_{50} > 1$ (class 2).

Kohonen Self-Organizing Maps. The generation of the Kohonen self-organizing maps (SOMs)[71] was also conducted using the Smart Mining software. The training parameters for the SOM were as follows: the number of interactions for the training runs was 2000, the starting adjustment radius for the training runs was 5, and the learning rate factor was 0.01. After the SOM was generated, we studied the distribution of opposite compound groups (potent or nonpotent hERG inhibitors) as separate maps.

Sammon Nonlinear Mapping. The Sammon map[72] generation was conducted using a program developed internally at ChemDiv as part of the Smart Mining, v1.01 software suite. The nonlinear map (NLM) was built on the basis of the following parameters: maximal number of iterations 500; optimization step 0.1. Euclidean distance was used as a similarity measure. The Sammon NLM procedure allows the creation of a 2-D image of the studied multidimensional property space. For visual discrimination of the studied drug categories on the nonlinear map, we used the separation lines. The positioning of the separation line was determined using the nonlinear Support Vector Machine (SVM) algorithm[73] as implemented in the LibSVM-2.4 program (URL: http://www.csie.ntu.edu.tw/~cjlin/libsvm/). The separation line represents the largest margin separating the studied compounds classes, which is defined as the sum of the shortest distances from the decision line to the closest points for both classes and thus can serve as an optimal discriminator between the two studied compound categories.

External Validation and Statistical Analysis. The 35 molecules used to test all the models were obtained from the literature ($\log_{10}$(IC_{50} in μM) = −1.79 to 3.64) following the initial QSAR model construction with the training set. These 35 molecules represent those that were not found initially when searching for the training set or were published later. These test set molecules were input into ChemTree as an sdf file, and predictions were made with the 99-molecule, 100-tree models generated previously. The in vitro observed data and the computationally predicted data for the average of the models were compared and assessed using the r^2 and the Spearman rho correlation available in JMP 5.1 (SAS Institute Inc., Cary, NC). The similarity of the molecules in the test set to those in the recursive partitioning model were assessed using the Accord software molecule descriptors (Accelrys, San Diego, CA) and the Tanimoto similarity coefficient.[74] The Tanimoto coefficient is $a/(a + b + c)$ where a is the number of bits common to both the query and target structures, b is the number of bits exclusively in the query structure, and c is the number of bits exclusively in the target structure. In this case a value of 1 indicates that the molecule is identical to one in the training set. As this value decreases, the less similar the molecule is to molecules in the training set. Similarly, the molecular descriptors selected for the mapping methods were calculated for these test set molecules followed by Kohonen and Sammon mapping in order to obtain qualitative predicted classifications of hERG activity with these models. These predictions were compared with the observed classifications of hERG activity.

Consensus Analysis of hERG Models. We also performed a consensus analysis of the test set molecules using predicted data from all three computational methods for modeling hERG inhibition. The consensus score $S_{c,i}$ for each of the two

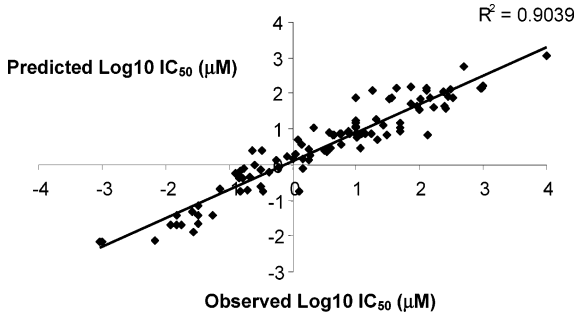

Figure 1. Training set correlation for 99 molecules using 100 random recursive partitioning models.

considered classes i ($i = 0$ and 2) was calculated as $S_{c,i} = (S_{rp,i} + S_{km,i} + S_{sm,i})/3$, where $S_{rp,i}$, $S_{km,i}$, and $S_{sm,I}$ are the predicted classes for the recursive partitioning, Kohonen mapping, and Sammon mapping, respectively. We compared the consensus scores for classes 0 and 2 and assigned a compound to a particular class with the higher consensus score. It was demonstrated that the mapping techniques do not provide a useful classification for the intermediate class 1. Therefore, in this analysis, we calculated the consensus scores for only 21 molecules in the test set belonging to experimental classes 0 and 2.

Database Searching. An in-house database was generated as described previously[75] consisting of 578 known drugs from the Clinician's Pocket Drug Reference.[76] These represent drugs that are in clinical use in the U.S.. This database was created using structures in the sdf format. The hERG recursive partitioning model ($n = 99$ molecules) developed with ChemTree was used to search this database. The similarity of the input database molecules to those in the individual QSAR model training set was calculated using the Tanimoto coefficient[74] as described above. An sdf file of predictions for hERG and Tanimoto similarity values were then output for all database molecules. This file was sorted in ChemDraw for Excel (CambridgeSoft, Cambridge, MA) and ranked by predicted $\log_{10}$ IC_{50}. The 25 most potent predicted inhibitors were evaluated, and any molecules in the model training set or test set were removed from this list. The remaining molecules were then used as queries for searches in PubMed to find relevant literature relating to hERG inhibition.

Results

Recursive Partitioning. The training set for 99 molecules with patch clamping hERG data, consisting primarily of data from HEK-293 cells, was used to generate 100 random models with an observed versus predicted $\log_{10}$ IC_{50} correlation $r^2 = 0.90$ (Figure 1).

Analysis of the 35-molecule test set (Table 2) using quantitative predictions derived from this recursive partitioning model resulted in a relatively low though statistically significant correlation ($r^2 = 0.33$, Spearman rho $= 0.55$, $p = 0.0006$) (Figure 2a). When the Tanimoto similarity analysis of the test set molecules was used (Figure 2b), it was observed that the log of the difference between observed and predicted $\log_{10}$ IC_{50} increased as the Tanimoto similarity declined. Therefore, a Tanimoto similarity index greater than 0.77 was found to contain

Table 2. Observed and Predicted hERG IC_{50} Results for the Test Set Compounds Derived with the Different QSAR Models[a]

			recursive partitioning			predicted class[b]				
no.	name	exptl $\log_{10}(IC_{50}$ in $\mu M)$ (class[b])	prediction $\log_{10}(IC_{50}$ in $\mu M)$ (class[b])	prediction std dev	similarity (arbitrary units)	on Sammon map	on Kohonen map	predicted consensus class[b]	cell type	ref
1	pilsicainide	1.3 (2)	0.6 (1)	0.65	0.82	2	0	unclass	HEK	77
2	ajmaline	0 (1)	0.78 (1)	0.54	0.74				HEK	78
3	budipine	1.0 (1)	0.73 (1)	0.89	0.77				XO	79
4	tamoxifen	1.65 (2)	0.84 (1)	0.66	0.76	2	2	2	XO	80
5	prulifloxacin	2.52 (2)	2.12 (2)	0.28	0.91	2	2	2	HEK	81
6	saxitoxin	−0.32 (0)	1.38 (2)	0.66	0.73	2	2	2	HEK	82
7	phenytoin	2.38 (2)	0.96 (1)	0.33	0.70	2	2	2	HEK	83
8	4AP	3.64 (2)	0.79 (1)	1.09	0.34	2	2	2	HEK	84
9	phenobarbital	3.48 (2)	1.12 (2)	0.55	0.72	2	2	2	HEK	83
10	miconazole	0.32 (1)	0.49 (1)	0.46	0.80				HEK	85
11	trazodone	0.46 (1)	0.04 (1)	0.66	0.8				HEK	86
12	trifluoperazone	0.96 (1)	0.019 (1)	0.85	0.74				HEK	87
13	lamotrigine	2.36 (2)	1.33 (2)	0.62	0.74	2	2	2	HEK	88
14	naringenin	1.56 (2)	1.08 (2)	0.48	0.82	2	2	2	HEK	89
15	pentamidine	0.7 (1)	0.98 (1)	0.89	0.68				HEK	90
16	clenbuterol	1.9 (2)	1.36 (2)	0.95	0.63	2	2	2	HEK	91
17	metoclopramide	0.73 (1)	0.93 (1)	1.03	0.81				HEK	92
18	dronedarone	−1.23 (0)	−1.35 (0)	1.17	0.86	0	0	0	HEK	93
19	GF109203X	0 (1)	0.27 (1)	0.47	0.85				HEK	94
20	vardenafil	1.1 (2)	0.43 (1)	0.53	0.97	2	0	unclass	HEK	95
21	tadalafil	2 (2)	1.19 (2)	0.89	0.88	2	0	2	HEK	95
22	tolteridone	−1.77 (0)	1.24 (2)	0.56	0.76	0	0	0	CHO	96
23	chloroquine	0.4 (1)	0.15 (1)	0.57	0.74				HEK	97
24	lumefantrine	0.9 (1)	−0.32 (0)	0.61	0.74	0	0		HEK	97
25	desbutyllumefantrine	0.74 (1)	−0.33 (0)	0.62	0.74				HEK	97
26	prazosin	0.19 (1)	0.59 (1)	0.75	0.79				HEK	98
27	doxazosin	−0.23 (0)	0.61 (1)	0.76	0.81	0	0	0	HEK	98
28	terazosin	1.24 (2)	0.59 (1)	0.75	0.79	2	2	2	HEK	98
29	clomiphene	−0.74 (0)	0.016 (1)	0.44	0.72	0	0	0	HEK	99
30	lidoflazine	−1.79 (0)	−1.18 (0)	0.68	0.79	0	0	0	HEK	100
31	amsacrine	−0.67 (0)	0.15 (1)	2.06	0.68	0	0	0	HEK	101
32	DW286a	1.94 (2)	1.57 (2)	1.28	0.88	2	2	2	CHO	102
33	AVE0118	1 (1)	0.62 (1)	1.02	0.80				CHO	103
34	AMP579	1 (1)	0.50 (1)	1.19	0.85				HEK	104
35	propranolol	1.9 (2)	1.14 (2)	0.91	0.73	2	2	2	HEK	91

[a] The observed data was retrieved from different literature sources and cell types. [b] $\log_{10}$ $IC_{50} < 0$ (class 0); $\log_{10}$ $IC_{50} = 0-1$ (class 1); $\log_{10}$ $IC_{50} > 1$ (class 2); HEK = HEK-293; XO = Xenopus oocytes; CHO = Chinese hamster ovary cells; unclass = unclassified.

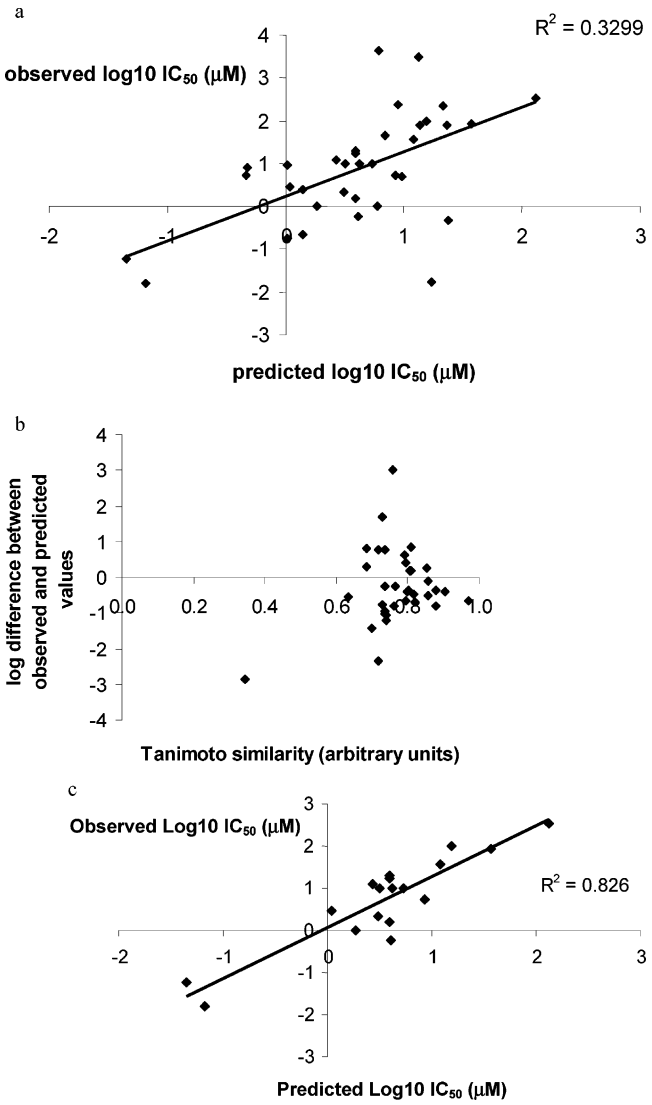

Figure 2. (a) Test set correlation using the 99 molecule, 100 random recursive partitioning models with the 35 molecule test set. (b) Plot of the log of the difference between observed and predicted IC_{50} values for the training set versus the Tanimoto similarity distance. (c) Test set correlation using the 99 molecule, 100 random recursive partitioning model with the filtered test set of 18 molecules (Tanimoto similarity index greater than 0.77). Note the removal of extreme outliers when compared with part a.

molecules with $\log_{10}$ IC_{50} that were less than 1 log unit different from the observed data. These 18 remaining molecules alone produced dramatically improved testing correlation statistics ($r^2 = 0.83$, Spearman rho $= 0.75$, $p = 0.0003$, Figure 2c) compared with using the whole test set. There did not appear to be any relationship between the log of the difference between the observed and predicted $\log_{10}$ IC_{50} and the standard deviation for the predictions for the recursive partitioning models.

The original training set and the test set were combined to produce an updated model for future testing. A 100-tree recursive partitioning model with these 134 molecules produced a training correlation of $r^2 = 0.85$ for observed and predicted $\log_{10}$ IC_{50} (data not shown). The mining of the literature in the future will enable us to derive additional molecules that can be used to test this model in turn.

The quantitative test set predictions were also placed into the qualitative three classes based on the activity cut-offs described earlier. This produced 22 of 35 (63%) correct class predictions. When the Tanimoto similarity cut-off described above was used to limit the test set to those molecules that possessed a Tanimoto similarity index greater than 0.77, 14 of 18 molecules (77.8%) were correctly classified.

Table 3. Selected Molecular Descriptors Used in Kohonen and Sammon Map Models[a]

descriptor abbreviation	descriptor name
W_{rms}	topological mean-square Wiener index
J_t	topological Balaban index of modified connection distance
$S(-CH_2-)$	electrotopological index of CH_2 group
S(=CH−)	electrotopological index of CH group
S(>N−)	electrotopological index of >N group
$HS(CH_2)$	electrotopological index of CH_2 group
Don	number of H-bond donors
Hy	hydrophilicity index

[a] These represent the remaining descriptors from an initial set of over 150 that were generated with the Smart Mining software tool followed by principal component analysis.

Descriptor Selection for Mapping Studies. PCA was used to select eight descriptors maximally contributing to the first significant principal components (Table 3). These descriptors were calculated as defined by Todeschini and Consonni.[105] The topological mean square Wiener index represents a measure of molecular branching and is related to molecular surface area, thus reflecting molecular compactness. The topological Balaban index of modified connection distance provides information on the intramolecular relationships between atoms and accounts for both bond multiplicity and heteroatoms. The number of H-bond donors represents a measure of the hydrogen-bonding ability of a molecule expressed as a count of hydrogen atoms bonded to oxygen, nitrogen, or sulfur. The hydrophilicity index is a simple empirical index based on count descriptors.[105] The electrotopological state (E-state) indices encode information about both the topological environment of an atom and the electronic interactions resulting from all other atoms in the molecule. E-state indices encode some essential molecular features characterizing the topology, polarity, and hydrogen-bonding capabilities of a molecule.

Sammon Nonlinear Mapping. The filtered training set of 93 molecules (Table 1) with 8 molecular descriptors (Table 3) calculated for each compound was used for the generation of a Sammon nonlinear map. After the nonlinear map was generated, we identified the positions of compounds belonging to classes 0−2 (Figure 3). Obviously, compounds belonging to the opposite classes 0 and 2 (white and gray circles, respectively) occupied distinctly different areas on the map. Compounds belonging to the intermediate class 1 with $IC_{50} = 0-1$ μM (shown as diamonds) occupied a wide area on the map overlapping with the sites of location of classes 0 and 2 compounds (Figure 3).

The observed differences in mapping to the Sammon map were used to assess the hERG inhibition profile for novel compounds. The positions of the 35 test set compounds (Table 2) on the same map are shown in Figure 4. The test set predictions based on the cut-offs described earlier are shown in Table 2. Thus, 6 of 7 (86%) molecules in class 0 and 14 of 14 (100%) in class 2 were correctly classified, resulting in a total correct classification of 95% (Table 4). As observed with the training set, compounds belonging to the intermediate class 1 occupied a wide area on the map overlapping with the sites of location of classes 0 and 2 compounds.

Kohonen Self-Organizing Maps. The same training set of 93 molecules (Table 1) with 8 molecular descriptors (Table 3) calculated for each compound was used to generate a Kohonen self-organizing map. After the map was generated, we identified

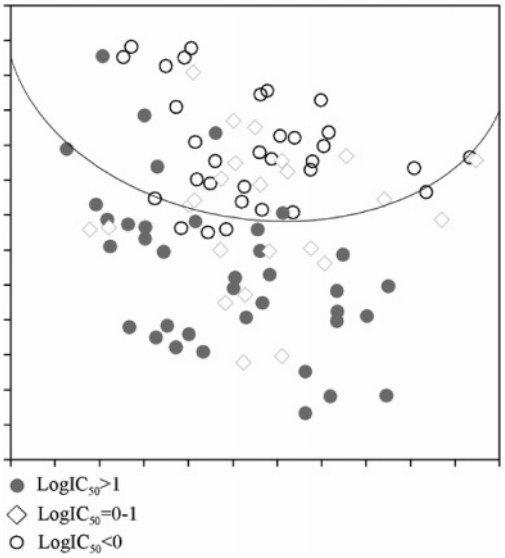

Figure 3. Sammon Map for the 93 hERG training set molecules generated for 8 molecular descriptors illustrating the positions of compounds in classes 0 (white circles), 1 (diamonds), and 2 (gray circles).

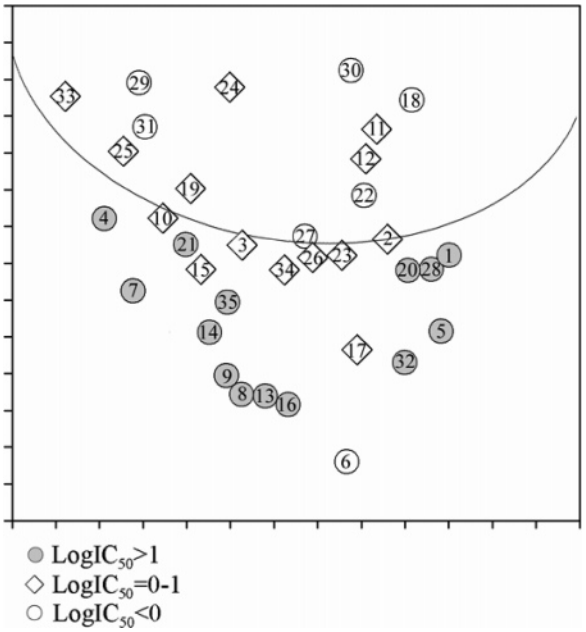

Figure 4. Distribution of the test set molecules on the Sammon map. The numbers shown correspond to the record numbers in Table 2.

the positions of compounds belonging to molecules belonging to classes 0−2 (Figure 5a−c, respectively). According to the figures, compounds belonging to the opposite classes 0 and 2 (parts a and c of Figure 5, respectively) occupied distinctly different areas on the map. Compounds belonging to class 1 with $IC_{50} = 0-1$ μM (Figure 5b) occupy an intermediate area on the map partially overlapping with the sites of classes 0 and 2 compounds.

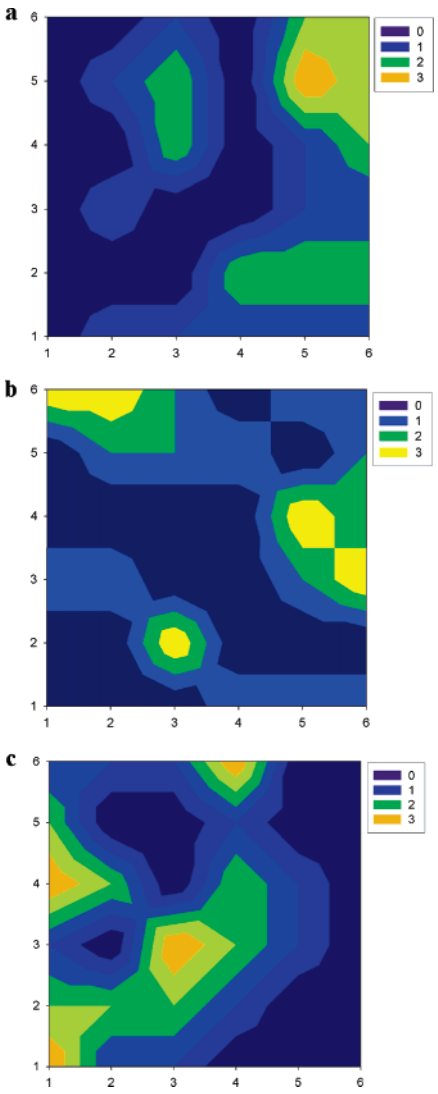

Figure 5. Kohonen maps for hERG training set data individually showing the distribution of molecules in each of the activity classes using the same molecular descriptors: (a) the area of compounds with $\log_{10} IC_{50} < 0$ (32 compounds); (b) with $\log_{10} IC_{50} = 0-1$ (29 compounds); (c) with $\log_{10} IC_{50} > 1$ (32 compounds). Colors indicate the intensity of the molecule distribution and have been smoothed for presentation purposes. Lighter colors represent higher densities of distribution of molecules.

The distance between nodes on the Kohonen map is a dimensionless parameter; it represents an abstract, discrete distance between the points in a multidimensional property space. For the training set, the areas of strong and poor binders can be identified as the nodes on the map, where the percentage of strong or poor binders (total number equal to 100%) is higher than the percentage of compounds belonging to the opposite category. The positions of 35 test set compounds on this map (Figure 6) and the experimental and predicted data are individually summarized (Table 2). The map correctly classified 6 of 7

Table 4. Percentage of Correctly Classified Compounds from the External Test Set for hERG Determined Using the Kohonen and Sammon Mapping Techniques

	training set			test set		
class	no. compds	Kohonen, no. compds, %	Sammon, no. compds, %	no. compds	Kohonen, no. compds, %	Sammon, no. compds, %
$\log_{10} IC_{50} < 0$ (class 0)	32	26, 81	28, 88	7	6, 86	6, 86
$\log_{10} IC_{50} > 1$ (class 2)	32	27, 84	27, 84	14	11, 79	14, 100
total	93	53, 83	55, 86	35	17, 81	20, 95

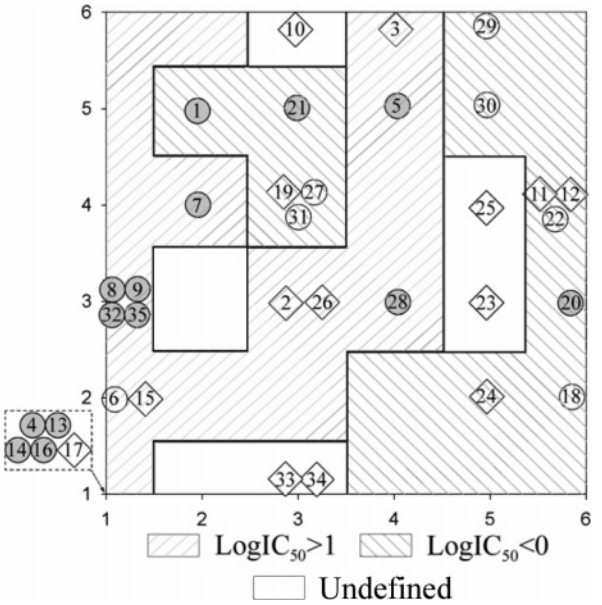

Figure 6. Distribution of classification areas on the Kohonen SOM (shaded background) and positions of the test set compounds: compounds with $\log_{10}$ IC_{50} > 1 (gray circles); compounds with $\log_{10}$ IC_{50} < 0 (white circles); compounds with $\log_{10}$ IC_{50} = 0−1 (diamonds). The numbers shown correspond to the record numbers in Table 2.

(86%) of class 0 compounds and 11 of 14 (79%) of class 2 compounds with a total of 81% correct predictions (Table 4). The quality of this discrimination is statistically significant only in the case of classes 0 and 2 because molecules belonging to class 1 occupied a wide area on the map. Although the distribution of compounds on the map suggests the method may be able to discriminate between all the studied classes 0−2, more data are required for a statistically valid result.

Consensus Analysis of Models. A consensus analysis of the 21 test set molecules available across all three computational methods for the two classes resulted in 6 of 7 (86%) compounds for class 0 and in 12 of 14 (86%) compounds for class 2 that were correctly predicted. These results are not representative of a marked improvement over the individual methods alone because the Sammon mapping predictions alone are preferable. One compound, saxitoxin (no. 6 in Table 2), was classified incorrectly by all three computational methods. Two other compounds, pilsicainide and vardenafil (nos. 1 and 20, Table 2), gave different results with all three methods, and these compounds therefore remained unclassified using the approach for consensus score calculation.

Database Searching. The recursive partitioning hERG model (n = 99 molecules) was used to search a database of 578 known drugs.[76] Out of the top 25 molecules ranked by the predicted hERG $\log_{10}$ IC_{50} value, there were 12 molecules (2.1% of the whole database, Table 5) that were not in the model training set. The remaining 13 molecules retrieved in this set (present in the training set) had predicted $\log_{10}$ IC_{50} values between −0.36 and −1.92, strongly indicative that the other molecules were also likely to be hERG inhibitors. Eleven of these 12 nontraining set molecules remaining had Tanimoto similarity values greater than or equal to 0.77 (Table 5), indicating that they were on or above the threshold described earlier and likely to provide good predictions of binding less than 1 log unit different from the observed data. A search of the literature using PubMed to find published data on these molecules indicated that citalopram, pergolide, and sotalol had confirmed interactions with hERG at nanomolar to micromolar concentrations. In the case of pergolide the prediction is within 1 log of the observed $\log_{10}$ IC_{50}. In addition, QTc prolongation in humans for citalopram,[106] paroxetine,[107] and the voltage- and use-dependent block of sotalol[108] have been recognized in the literature. The remaining molecules have not been tested in vitro to our knowledge. A similar approach to database searching with the Sammon and Kohonen mapping approaches for hERG has not been performed.

Table 5. Molecules Retrieved and Scored with the hERG Recursive Partitioning Model[a]

molecule	Tanimoto similarity	predicted hERG $\log_{10}$(IC_{50}, μM)	published hERG IC_{50}, μM ($\log_{10}$ IC_{50})	ref
delavirdine	0.83	−1.37	na	
levocabastine	0.77	−1.36	na	
meperidide	0.81	−1.12	na	
citalopram	0.80	−0.84	3.97 (0.59)	107, 109
flurazepam	0.78	−0.75	na	
pergolide	0.78	−0.62	0.12 (−0.92)	110
fluvastatin	0.85	−0.29	na	
sotalol	0.85	−0.28	∼30−269 (1.48−2.42)	111, 112
paroxetine	0.77	−0.23	na	
pramoxine	0.67	−0.17	na	
mezlocillin	0.82	−0.16	na	
loperamide	0.85	−0.13	na	

[a] The 99-molecule recursive partioning model was used to search a database of 578 known drugs that are used in the U.S.[76] The molecules were sorted by the predicted hERG $\log_{10}$(IC_{50} in μM), and the 25 highest affinity predictions were assessed. Molecules present in the training set were removed to leave unique predictions. The literature was then searched to obtain verification and examples of IC_{50} values. na = not available.

Discussion

There have been considerable efforts to address the attrition of drug candidates in drug discovery using earlier and more extensive in vitro testing as well as computational approaches for absorption, distribution, metabolism, excretion, and toxicity (ADME/Tox).[27] Cardiotoxicity was identified as one area for particular attention because several drugs were removed from the market as a result of undesirable QT prolongation associated with binding to the hERG potassium channel.[5−9] Since 2002, there have been at least 14 studies that have described individual QSAR models, statistical models, or pharmacophores for hERG.[10−23] These studies have encompassed a wide set of data generation and modeling techniques as well as an array of molecules for model building and testing. The results of these methods have produced some pharmacophore features and molecular descriptors for identifying hERG inhibitors that can be briefly summarized as follows: (A) four hydrophobes between 5.2 and 7.5 Å from a positive ionizable feature;[16] (B) three aromatic (hydrophobic) features 4.6−9.1Å from a central nitrogen were key;[20] (C) at least two hydrophobic features 6−8 Å from a basic nitrogen;[114] (D) a hydrophobe (or hydrogen bond acceptor), a ring aromatic, and a positive ionizable feature are key;[22] (E) two hydrophobes with an asymmetrically placed protonated nitrogen between them;[21] (F) a basic protonated nitrogen or hydrogen bond donor, hydrogen bond acceptor, and hydrophobes;[19] (G) two ring aromatic features, a hydrophobe, and a positive ionizable feature;[17,27] (H) calculated log P (ClogP, a measure of hydrophobicity) is higher in potent hERG inhibitors compared with nonpotent inhibitors;[11] (I) ClogP, molar refractivity, partial negative surface area, polarizability, and hydrophobicity are differentiating descriptors.[15]

It is evident that from all of these preceding studies that there are some gross similarities in the suggested requirements for hERG inhibitors requiring hydrophobic features surrounding a

positive ionizable/basic nitrogen feature. However, depending on the molecules and techniques used for model building, the pharmacophore or descriptors suggested may vary. In the current study we have used a consistent set of literature derived molecules with hERG patch clamping data for model building and testing with different descriptors and three model building techniques, namely, recursive partitioning, Kohonen mapping, and Sammon mapping. The aim was to assess whether these methods could be used alone or together to accurately predict hERG inhibition as well as to gain some further insights into the structural requirements for binding. The descriptors selected by PCA analysis and used for the Sammon and Kohonen mapping techniques (Table 3) in general agree with the prior observations above in the need for hydrophobic features. However, the inclusion of the Wiener index represents a measure of molecular branching and molecular compactness, and the Balaban index connection distance provides information on the intramolecular relationships between atoms. These descriptors suggest that molecular shape or topological characteristics are also very important for hERG inhibitors and were not explicitly described previously in the literature.

Considerable research efforts have focused on machine learning algorithms, which can be used for drug discovery for predicting molecular properties for very large numbers of molecules, to ultimately act as a filter. Recursive partitioning methods have been used widely with large sets of molecules and either continuous[115,116] or binary data, for therapeutic target end points as well as cytochrome P450 inhibition[117] and AMES mutagenicity.[70] In the current study the recursive partitioning model generated a statistically significant rank ordering of the 35 molecule test set (Figure 2a) that was improved dramatically when only molecules with a Tanimoto similarity greater than 0.77 were considered (Table 2, parts b and c of Figure 2). The role of molecular similarity as a descriminator for prediction accuracy has been described;[118] however, this represents the first time that it has been applied to an hERG model, to our knowledge. This recursive partitioning model combined with the similarity calculation represents a fast method to score for potential interactions with hERG. We have also used this recursive partitioning model to quantitatively assess a library of known drugs and predict those that may have interactions with hERG. Three known hERG inhibitors that were not in the training set were identified in the top 25 scored molecules alongside 13 molecules from the training set that were potent inhibitors. Using both the similarity score and the predicted IC_{50} indicates that most of the remaining molecules may be considered for future testing in vitro to further verify our predictions. This example also highlights the limited amount of data in the public domain relating to known drugs and hERG interaction. Such database searching efforts could enable prioritization of larger compound databases for in vitro testing such as the NIH molecular libraries initiative (http://nihroadmap.nih.gov/molecularlibraries/).

The Sammon maps describe all relative distances between all pairs of compounds, and the distance of two points on the map directly reflects the similarity of the compounds. Among the other dimensionality reduction techniques that have appeared in the statistical literature, Sammon nonlinear maps are unique because of their conceptual simplicity and ability to reproduce the topology and structure of the data space in a faithful and unbiased manner. This method has a practical value and can be recommended for analysis of small-sized combinatorial libraries (up to several thousand compounds) aimed at the selection of subsets with enhanced knowledge-based information content.

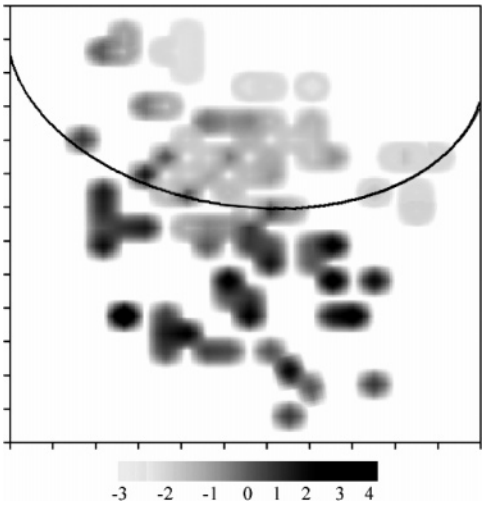

Figure 7. Heat map distribution of hERG IC_{50} values on the Sammon nonlinear map. Darker shading represents higher IC_{50} values.

We have described the use of Sammon maps with various ADME data sets previously.[119] The Sammon map that was generated in this study (Figure 3) is a useful tool for visual analysis of various quantitative and qualitative dependencies in the studied set of hERG inhibitors (Figure 4) with high classification rates for the test set (Tables 2 and 4). The distribution of IC_{50} values throughout the nonlinear map demonstrates an obvious gradient (Figure 7), which suggests that the map is able to reflect even more subtle structure−activity relationships for the studied hERG inhibitors. It should also be stressed there are disadvantages of this technique particularly for large data sets were the Sammon nonlinear map computation is increasingly intractable. The approach may also generate a 2D mapping that poorly approximates the original distances when the number of compounds is large.[120] Also, if additional data points or data sets are to be included in the projection, a complete recomputation of the map is required for all data points, hence further complicating testing.

To overcome these drawbacks we performed an additional series of experiments using self-organizing Kohonen maps for hERG inhibitors. Kohonen maps[71] were originally designed in an attempt to model intelligent information processing. The general idea behind this method is to map a set of vectorial samples onto a two-dimensional lattice in a way that preserves the topology of the original space. Kohonen maps belong to a class of neural networks known as competitive learning or self-organizing networks. The Kohonen map consists of artificial neurons that are characterized by weight vectors with the same dimensionality as the descriptor set. The neurons are connected by a distance-dependent function. In an unsupervised training algorithm the neurons self-organize until their pairwise neighborhoods represent the correct topology of the original data set. Kohonen maps have recently been applied to successfully model cytochrome P450 mediated drug metabolism.[121,122] In the present study, the nonlinear mapping algorithm provides clear visual discrimination between the principal classes of hERG inhibitors (Figure 5). The average classification quality is high for both training and test selections; up to 86% and 95% of compounds were classified correctly in the corresponding data sets (Table 4). At the same time, insufficient statistics prevent correct assignment of compounds belonging to the intermediate class 1. Figure 5 shows the distribution of the training set compounds on the Kohonen map that was generated. Obviously, compounds belonging to classes 0 (Figure 5a) and 2 (Figure 5c) occupy

distinctly different sites on the map. Nevertheless, while the potent and nonpotent inhibitors are located in distinctly different sites on the map, the class 1 compounds typically occupy the intermediate positions near the line of separation. The same conclusion can be made from analysis of location of class 1 compounds from the set on the Sammon map (Figure 4). At present after these two methods were used, compounds predicted in this region as class 1 would need to be tested in vitro to verify the correct classification. Our results also demonstrate a good generalization ability of the model with 6 of 7 (86%) of class 0 compounds and 11 of 14 (79%) of class 2 compounds being classified correctly in the external test set (Figure 6, Table 4). In the present case, the Sammon mapping technique outperformed the Kohonen maps in classification of compounds from the external test set (Table 4), but both are acceptable. This suggests that there is a high probability that a molecule predicted as class 0 or 2 actually belongs to these classes using these methods. In general, the Kohonen maps demonstrate a significant speed advantage compared to Sammon maps and they allow the instant inclusion of new individual or multiple data points on the map without the need for recomputing the entire data set. This enables visualization and analysis of larger databases or virtual libraries compared with Sammon maps. On the other hand, Sammon maps provide better distance and topology preservation compared with Kohonen maps, and the latter maps often contain gaps or undefined regions of chemical space. Therefore, there is a tradeoff to be made for these methods that will be the subject of further study.

There have been several published cases were a consensus scoring scheme can increase the predictive power of a computational method.[14,123,124] We calculated the consensus scores for the test set compounds belonging to the opposite experimental classes 0 and 2 for all three modeling approaches. The classification quality in this case is slightly better than the results provided by the Kohonen mapping and recursive partitioning alone: 6 of 7 (86%) compounds in class 0 and 12 of 14 (86%) compounds in class 2 were correctly predicted (compare with Table 4). However, the prediction accuracy is still lower than in the case of Sammon mapping alone with up to 95% compounds predicted correctly overall. This is primarily due to the consensus scoring failing to classify pilsicainide and vardenafil (nos. 1 and 20, Table 2); all three methods assigned these compounds to different hERG inhibitory activity classes. Both compounds have IC_{50} values close to those of the intermediate class 1 (1.1 and 1.3 μM, respectively) and were incorrectly classified using the Kohonen map (Figure 6), while the other algorithms made correct predictions. It should therefore be noted that consensus scoring in this case was not as successful as in previously published studies.[14,123,124] There was only one distinct outlier in this test set, saxitoxin (no. 6 in Table 2), which is incorrectly classified by all the methods (Table 2, Figures 4 and 6). Saxitoxin is a tricyclic molecule with two guanidino groups, which can carry a positive charge or can be partially deprotonated under physiological conditions. We suggest that the polar nature of saxitoxin and the presence of several tautomeric forms seriously complicate the calculation of relevant molecular descriptors and the in silico classification of this molecule. The calculation of molecular properties for minor tautomeric forms under physiological conditions has already been shown to lead to incorrect predictions.[125]

Although the described map-based models predict the extremes of activity correctly for the class 2 and class 0 compounds, insufficient statistics prevent correct assignment of compounds belonging to the intermediate class 1. However, in any sort of prospective analysis, the user of the model would not know a priori the potency of a molecule. To give an idea of the predictability of the best classification approach, Sammon mapping, on the full activity range, we reduced the classification problem into a two-class model. Arbitrarily, we considered compounds with $\log_{10} IC_{50} < 0.5$ as inhibitors and compounds with $\log_{10} IC_{50} > 0.5$ as noninhibitors. Using the same Sammon map model depicted in Figure 4, we calculated the prediction accuracy for the entire test set of 35 compounds (Table 6). The two-class model correctly predicted 64% of inhibitors and 86% of noninhibitors, with an average classification accuracy of 77%. Because of the aforementioned poor predictability of the mapping models for the intermediate potency agents, the last exercise has an illustrative character. Nevertheless, the two-class model still demonstrates good predictability and can be used for rough estimation of the hERG inhibitory potential.

Table 6. Percentage of Correctly Classified Compounds from the External Test Set in a Two-Class Classification Task Using the Sammon Mapping Technique

class	total no. compds	correctly classified, no. compds, %
$\log_{10} IC_{50} < 0.5$ (inhibitors)	14	9, 64
$\log_{10} IC_{50} > 0.5$ (noninhibitors)	21	18, 86
total	35	27, 77

It is certainly worth noting that the statistics-oriented QSAR approaches are sensitive to the properties of the molecules to be tested. In cases when the molecular parameters are too far from the parameters of the training set (for example, large molecular weight, high lipophilicity, increased number of polar groups, etc.), the predictions can be incorrect. Therefore, some preprocessing of the tested structural data sets is usually required to ensure their compatibility with the training set. Simple threshold filtering criteria can be used for such preprocessing, as well as the more sophisticated methods of removal of outliers using special statistical techniques.[121,122,126] In addition, in this study we have used a postfilter analysis in the form of a Tanimoto similarity assessment with the recursive partitioning model to remove molecules from the test set that are too dissimilar to the training set as described with the recursive partitioning approach. At present the different hERG modeling methods described herein represent the state of the art and could be combined and used in parallel as a consensus-modeling approach,[14] which might offer some improvement for the prediction of external molecules across different compound classes.

In this study, we have studied consistent sets of compounds with literature hERG $\log_{10} IC_{50}$ data using recursive partitioning, Sammon nonlinear maps, and Kohonen self-organizing maps. Our results with an external test set, also from the literature, indicate that these computational methods are all valuable as clustering, classification, and visualization tools either used individually or combined in a consensus model. We have also provided further structural insights for molecular requirements for hERG inhibition beyond our previous pharmacophore studies.[16,17,27] This study shows how the careful analysis of literature data from many different laboratories can be a valuable asset for computational model building[127−129] along with the appropriate analysis of similarity to the training set to avoid extrapolation.[118] As we have demonstrated with the recursive partitioning model, molecules used initially as a test set can be used to further expand the training set to update the model for future testing. These computational models described therefore represent the latest high-throughput filters for screening of diverse virtual libraries of molecules to classify the hERG

inhibitory potency of compounds prior to in vitro screening, allowing the potential avoidance of potent inhibitors and more efficient utilization of experimental resources.

Acknowledgment. S.E. gratefully acknowledges Dr. Maggie A. Z. Hupcey for editorial assistance.

Supporting Information Available: Molecules used in this study for HERG models. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Tan, H. L.; Hou, C. J.; Lauer, M. R.; Sung, R. J. Electrophysiologic mechanisms of the long QT interval syndromes and torsade de pointes. *Ann. Intern. Med.* **1995**, *122* (9), 701−714.
(2) Thomas, S. H. Drugs, QT interval abnormalities and ventricular arrhythmias. *Adverse Drug React. Toxicol. Rev.* **1994**, *13* (2), 77−102.
(3) Barry, D. M.; Xu, H.; Schuessler, R. B.; Nerbonne, J. M., Functional knockout of the transient outward current, long-QT syndrome, and cardiac remodeling in mice expressing a dominant-negative Kv4 alpha subunit. *Circ. Res.* **1998**, *83* (5), 560−567.
(4) Jeron, A.; Mitchell, G. F.; Zhou, J.; Murata, M.; London, B.; Buckett, P.; Wiviott, S. D.; Koren, G. Inducible polymorphic ventricular tachyarrhythmias in a transgenic mouse model with a long Q−T phenotype. *Am. J. Physiol.: Heart Circ. Physiol.* **2000**, *278* (6), H1891−H1898.
(5) Trudeau, M. C.; Warmke, J. W.; Ganetzky, B.; Robertson, G. A. HERG, a human inward rectifier in the voltage-gated potassium channel family. *Science* **1995**, *269* (5220), 92−95.
(6) Warmke, J. W.; Ganetzky, B. A family of potassium channel genes related to eag in Drosophila and mammals. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91* (8), 3438−3442.
(7) Curran, M. E.; Splawski, I.; Timothy, K. W.; Vincent, G. M.; Green, E. D.; Keating, M. T. A molecular basis for the cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* **1995**, *80*, 795−803.
(8) Rampe, D.; Murawsky, M. K.; Grau, J.; Lewis, E. W. The antipsychotic agent sertindole is a high affinity antagonist of the human cardiac potassium channel HERG. *J. Pharmacol. Exp. Ther.* **1998**, *286* (2), 788−793.
(9) Suessbrich, H.; Schonherr, R.; Heinemann, S. H.; Attali, B.; Lang, F.; Busch, A. E. The inhibitory effect of the antipsychotic drug haloperidol on HERG potassium channels expressed in Xenopus oocytes. *Br. J. Pharmacol.* **1997**, *120* (5), 968−974.
(10) 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* (11), 2886−2890.
(11) Roche, O.; Trube, G.; Zuegge, J.; Pflimlin, P.; Alanine, A.; Schneider, G. A virtual screening method for the prediction of the hERG potassium channel liability of compound libraries. *ChemBioChem* **2002**, *3*, 455−459.
(12) Pearlstein, R. A.; Vaz, R. J.; 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*, 2017−2022.
(13) 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.* **2003**, *13*, 1829−1835.
(14) O'Brien, S. E.; de Groot, M. J. Greater than the sum of its parts: combining models for useful ADMET prediction. *J. Med. Chem.* **2005**, *48* (4), 1287−1291.
(15) Kesuru, G. M. Prediction of hERG potassium channel affinity by traditional and hologram QSAR methods. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2773−2775.
(16) Ekins, S.; Crumb, W. J.; Sarazan, R. D.; Wikel, J. H.; Wrighton, S. A. Three dimensional quantitative structure activity relationship for the inhibition of the hERG (human ether-a-gogo related gene) potassium channel. *J. Pharmacol. Exp. Ther.* **2002**, *301*, 427−434.
(17) Ekins, S. Predicting undesirable drug interactions with promiscuous proteins in silico. *Drug Discovery Today* **2004**, *9*, 276−285.
(18) Ekins, S. In silico approaches to predicting metabolism, toxicology and beyond. *Biochem. Soc. Trans.* **2003**, *31*, 611−614.
(19) Cianchetta, G.; Li, Y.; Kang, J.; Rampe, D.; Fravolini, A.; Cruciani, G.; Vaz, R. J. Predictive models for hERG potassium channel blockers. *Bioorg. Med. Chem. Lett.* **2005**, *15* (15), 3637−3642.
(20) Cavalli, A.; Poluzzi, E.; De Ponti, F.; Recanatini, M. 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.
(21) 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* (2), 205−233.
(22) Aronov, A. M.; Goldman, B. B. A model for identifying HERG K+ channel blockers. *Bioorg. Med. Chem.* **2004**, *12* (9), 2307−2315.
(23) Aptula, A. O.; Cronin, M. T. Prediction of hERG K+ blocking potency: application of structural knowledge. *SAR QSAR Environ Res.* **2004**, *15* (5−6), 399−411.
(24) Rowley, M.; Hallett, D. J.; Goodacre, S.; Moyes, C.; Crawforth, J.; Sparey, T. J.; Patel, S.; Marwood, R.; Patel, S.; Thomas, S.; Hitzel, L.; O'Connor, D.; Szeto, N.; Castro, J. L.; Hutson, P. H.; MacLeod, A. M. 3-(4-Fluoropiperidin-3-yl)-2-phenylindoles as high affinity, selective, and orally bioavailable h5-HT2A receptor antagonists. *J. Med. Chem.* **2002**, *44*, 1603−1614.
(25) Fletcher, S. R.; Burkhamp, F.; Blurton, P.; Cheng, S. K. F.; Clarkson, R.; O'Connor, D.; Spinks, D.; Tudge, M.; van Niel, M. B.; Patel, S.; Chapman, K.; Marwood, R.; Shepheard, S.; Bentley, G.; Cook, G. P.; Bristow, L. J.; Castro, J. L.; Hutson, P. H.; MacLeod A. M. 4-(Phenylsulfonyl)piperidines: novel, selective, and bioavailable 5-HT2A receptor antagonists. *J. Med. Chem.* **2002**, *45*, 492−503.
(26) Bell, I. M.; Gallicchio, S. N.; Abrams, M.; Beese, L. S.; Beshore, D. C.; Bhimnathwala, H.; Bogusky, M. J.; Buser, C. A.; Culberson, J. C.; Davide, J.; Ellis-Hutchings, M.; Fernphenylsulfonylandes, C.; Gibbs, J. B.; Graham, S. L.; Hamilton, K. A.; Hartman, G. D.; Heimbrook, D. C.; Homnick, C. F.; Huber, H. E.; Huff, J. R.; Kassahun, K.; Koblan, K. S.; Kohl, N. E.; Lobell, R. B.; Lynch, J. J., Jr.; Robinson, R.; Rodrigues, A. D.; Taylor, J. S.; Walsh, E. S.; Williams, T. M.; Zartman, C. B. 3-Aminopyrrolidinone farnesyltransferase inhibitors: design of macrocyclic compounds with improved pharmacokinetics and excellent cell potency. *J. Med. Chem.* **2002**, *45*, 2388−2409.
(27) Ekins, S.; Swaan, P. W. Development of computational models for enzymes, transporters, channels and receptors relevant to ADME/TOX. *Rev. Comput. Chem.* **2004**, *20*, 333−415.
(28) Osterberg, F.; Aqvist, J. Exploring blocker binding to a homology model of the open hERG K+ channel using docking and molecular dynamics methods. *FEBS Lett.* **2005**, *579* (13), 2939−2944.
(29) 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.
(30) Fernandez, D.; Ghanta, A.; Kauffman, G. W.; Sanguinetti, M. C. Physicochemical features of the HERG channel drug binding site. *J. Biol. Chem.* **2004**, *279* (11), 10120−10127.
(31) Sanguinetti, M. C.; Mitcheson, J. S. Predicting drug−hERG channel interactions that cause acquired long QT syndrome. *Trends Pharmacol. Sci.* **2005**, *26* (3), 119−124.
(32) Yang, T.; Snyders, D.; Roden, D. M. Drug block of I(kr): model systems and relevance to human arrhythmias. *J. Cardiovasc. Pharmacol.* **2001**, *38* (5), 737−744.
(33) Tie, H.; Walker, B. D.; Valenzuela, S. M.; Breit, S. N.; Campbell, T. The heart of psychotropic drug therapy. *Lancet* **2000**, *355*, 1825.
(34) Zhou, Z.; Vorperian, V. R.; Gong, Q.; Zhang, S.; January, C. T. Block of HERG potassium channels by the antihistamine astemizole and its metabolites desmethylastemizole and norastemizole. *J. Cardiovasc. Electrophysiol.* **1999**, *10* (6), 836−843.
(35) Walker, B. D.; Singleton, C. B.; Tie, H.; Bursill, J. A.; Wyse, K. R.; Valenzuela, S. M.; Breit, S. N.; Campbell, T. J. Comparative effects of azimilide and ambasilide on the human ether-a-go-go-related gene (HERG) potassium channel. *Cardiovasc. Res.* **2000**, *48* (1), 44−58.
(36) Chouabe, C.; Drici, M. D.; Romey, G.; Barhanin, J.; Lazdunski, M. HERG and KvLQT1/IsK, the cardiac K+ channels involved in long QT syndromes, are targets for calcium channel blockers. *Mol. Pharmacol.* **1998**, *54* (4), 695−703.
(37) Thomas, D.; Wendt-Nordahl, G.; Rockl, K.; Ficker, E.; Brown, A. M.; Kiehn, J. High-affinity blockade of human ether-a-go-go-related gene human cardiac potassium channels by the novel antiarrhythmic drug BRL-32872. *J. Pharmacol. Exp. Ther.* **2001**, *297* (2), 753−761.
(38) Katchman, A. N.; McGroary, K. A.; Kilborn, M. J.; Kornick, C. A.; Manfredi, P. L.; Woosley, R. L.; Ebert, S. N. Influence of opioid agonists on cardiac human ether-a-go-go-related gene K(+) currents. *J. Pharmacol. Exp. Ther.* **2002**, *303* (2), 688−694.
(39) Karle, C. A.; Kreye, V. A.; Thomas, D.; Rockl, K.; Kathofer, S.; Zhang, W.; Kiehn, J. Antiarrhythmic drug carvedilol inhibits HERG potassium channels. *Cardiovasc. Res.* **2001**, *49* (2), 361−370.

(40) Taglialatela, M.; Pannaccione, A.; Castaldo, P.; Giorgio, G.; Zhou, Z.; January, C. T.; Genovese, A.; Marone, G.; Annunziato, L. Molecular basis for the lack of HERG K+ channel block-related cardiotoxicity by the H1 receptor blocker cetirizine compared with other second-generation antihistamines. *Mol Pharmacol.* **1998**, *54* (1), 113−121.
(41) Suessbrich, H.; Waldegger, S.; Lang, F.; Busch, A. E. Blockade of HERG channels expressed in Xenopus oocytes by the histamine receptor antagonists terfenadine and astemizole. *FEBS Lett.* **1996**, *385* (1−2), 77−80.
(42) Kang, J.; Wang, L.; Chen, X.-L.; Triggle, D. J.; Rampe, D. Interactions of a series of fluoroquinolone antibacterial drugs with the human cardiac K^+ channel HERG. *Mol. Pharmacol.* **2001**, *59*, 122−126.
(43) Volberg, W. A.; Koci, B. J.; Su, W.; Lin, J.; Zhou, J. Blockade of human cardiac potassium channel human ether-a-go-go-related gene (HERG) by macrolide antibiotics. *J. Pharmacol. Exp. Ther.* **2002**, *302* (1), 320−327.
(44) Ferreira, S.; Crumb, W. J., Jr.; Carlton, C. G.; Clarkson, C. W. Effects of cocaine and its major metabolites on the HERG-encoded potassium channel. *J. Pharmacol. Exp. Ther.* **2001**, *299* (1), 220−226.
(45) Fenichel, R. R. Receptor Binding Database. http://www.fenichel.net/pages/Professional/subpages/QT/Tables/pbydrug.htm (accessed 2005).
(46) Zhang, S.; Zhou, Z.; Gong, Q.; Makielski, J. C.; January, C. T. Mechanism of block and identification of the verapamil binding domain to HERG potassium channels. *Circ. Res.* **1999**, *84* (9), 989−998.
(47) Taglialatela, M.; Timmerman, H.; Annunziato, L. Cardiotoxic potential and CNS effects of first-generation antihistamines. *Trends Pharmacol. Sci.* **2000**, *21*, 52−56.
(48) Snyders, D. J.; Chaudhary, A. High affinity open channel block by dofetilide of HERG expressed in a human cell line. *Mol. Pharmacol.* **1996**, *49* (6), 949−955.
(49) Kuryshev, Y. A.; Brown, A. M.; Wang, L.; Benedict, C. R.; Rampe, D. Interactions of the 5-hydroxytryptamine 3 antagonist class of antiemetic drugs with human cardiac ion channels. *J. Pharmacol. Exp. Ther.* **2000**, *295* (2), 614−620.
(50) Drolet, B.; Rousseau, G.; Daleau, P.; Cardinal, R.; Turgeon, J. Domperidone should not be considered a no-risk alternative to cisapride in the treatment of gastrointestinal motility disorders. *Circulation* **2000**, *102* (16), 1883−1885.
(51) Drolet, B.; Zhang, S.; Deschenes, D.; Rail, J.; Nadeau, S.; Zhou, Z.; January, C. T.; Turgeon, J. Droperidol lengthens cardiac repolarization due to block of the rapid component of the delayed rectifier potassium current. *J. Cardiovasc. Electrophysiol.* **1999**, *10* (12), 1597−1604.
(52) Ko, C. M.; Ducic, I.; Fan, J.; Shuba, Y. M.; Morad, M. Suppression of mammalian K+ channel family by ebastine. *J. Pharmacol. Exp. Ther.* **1997**, *281* (1), 233−244.
(53) Chachin, M.; Katayama, Y.; Yamada, M.; Horio, Y.; Ohmura, T.; Kitagawa, H.; Uchida, S.; Kurachi, Y. Epinastine, a nonsedating histamine H1 receptor antagonist, has a negligible effect on HERG channel. *Eur J. Pharmacol.* **1999**, *374* (3), 457−460.
(54) Paul, A. A.; Witchel, H. J.; Hancox, J. C. Inhibition of the current of heterologously expressed HERG potassium channels by flecainide and comparison with quinidine, propafenone and lignocaine. *Br. J. Pharmacol.* **2002**, *136* (5), 717−729.
(55) Rosati, B.; Rocchetti, M.; Zaza, A.; Wanke, E. Sulfonylureas blockade of neural and cardiac HERG channels. *FEBS Lett.* **1998**, *440* (1−2), 125−130.
(56) Tie, H.; Walker, B. D.; Singleton, C. B.; Valenzuela, S. M.; Bursill, J. A.; Wyse, K. R.; Breit, S. N.; Campbell, T. J. Inhibition of HERG potassium channels by the antimalarial agent halofantrine. *Br. J. Pharmacol.* **2000**, *130* (8), 1967−1975.
(57) Teschemacher, A. G.; Seward, E. P.; Hancox, J. C.; Witchel, H. J. Inhibition of the current of heterologously expressed HERG potassium channels by imipramine and amitriptyline. *Br. J. Pharmacol.* **1999**, *128* (2), 479−485.
(58) Crumb, W. J., Jr. Loratadine blockade of K(+) channels in human heart: comparison with terfenadine under physiological conditions. *J. Pharmacol. Exp. Ther.* **2000**, *292* (1), 261−264.
(59) Suessbrich, H.; Schonherr, R.; Heinemann, S. H.; Lang, F.; Busch, A. E. Specific block of cloned Herg channels by clofilium and its tertiary analog LY97241. *FEBS Lett.* **1997**, *414* (2), 435−438.
(60) Kang, J.; Chen, X. L.; Wang, L.; Rampe, D. Interactions of the antimalarial drug mefloquine with the human cardiac potassium channels KvLQT1/minK and HERG. *J. Pharmacol. Exp. Ther.* **2001**, *299* (1), 290−296.
(61) Mitcheson, J. S.; Chen, J.; Sanguinetti, M. C. Trapping of a methanesulfonanilide by closure of the HERG potassium channel activation gate. *J. Gen Physiol.* **2000**, *115* (3), 229−240.
(62) Mbai, M.; Rajamani, S.; January, C. T. The anti-malarial drug halofantrine and its metabolite *N*-desbutylhalofantrine block HERG potassium channels. *Cardiovasc. Res.* **2002**, *55* (4), 799−805.
(63) Chouabe, C.; Drici, M. D.; Romey, G.; Barhanin, J. Effects of calcium channel blockers on cloned cardiac K+ channels IKr and IKs. *Therapie* **2000**, *55* (1), 195−202.
(64) Walker, B. D.; Valenzuela, S. M.; Singleton, C. B.; Tie, H.; Bursill, J. A.; Wyse, K. R.; Qiu, M. R.; Breit, S. N.; Campbell, T. J. Inhibition of HERG channels stably expressed in a mammalian cell line by the antianginal agent perhexiline maleate. *Br. J. Pharmacol.* **1999**, *127* (1), 243−251.
(65) Kongsamut, S.; Kang, J.; Chen, X. L.; Roehr, J.; Rampe, D. A comparison of the receptor binding and HERG channel affinities for a series of antipsychotic drugs. *Eur J. Pharmacol.* **2002**, *450* (1), 37−41.
(66) Po, S. S.; Wang, D. W.; Yang, I. C.; Johnson, J. P., Jr.; Nie, L.; Bennett, P. B. Modulation of HERG potassium channels by extracellular magnesium and quinidine. *J. Cardiovasc. Pharmacol.* **1999**, *33* (2), 181−185.
(67) Jurkiewicz, N. K.; Wang, J.; Fermini, B.; Sanguinetti, M. C.; Salata, J. J. Mechanism of action potential prolongation by RP 58866 and its active enantiomer, terikalant. Block of the rapidly activating delayed rectifier K+ current, IKr. *Circulation* **1996**, *94* (11), 2938−2946.
(68) Katayama, Y.; Fujita, A.; Ohe, T.; Findlay, I.; Kurachi, Y. Inhibitory effects of vesnarinone on cloned cardiac delayed rectifier K(+) channels expressed in a mammalian cell line. *J. Pharmacol. Exp. Ther.* **2000**, *294* (1), 339−346.
(69) Young, S. S.; Ekins, S.; Lambert, C. So many targets, so many compounds, but so few resources. *Curr. Drug Discovery* **2002**, December, 17−22.
(70) Young, S. S.; Gombar, V. K.; Emptage, M. R.; Cariello, N. F.; Lambert, C. Mixture deconvolution and analysis of Ames mutagenicity data. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 5−11.
(71) Kohonen, T. *Self-Organizing Maps*, 3rd ed.; Springer-Verlag: New York, 2000.
(72) Sammon, J. W. A non-linear mapping for data structure analysis. *IEEE Trans. Comput.* **1969**, *C18*, 401−409.
(73) Vapnik, V. *Statistical Learning Theory*; Wiley: New York, 1998.
(74) Willet, P. Similarity-based approaches to virtual screening. *Biochem. Soc. Trans.* **2003**, *31*, 603−606.
(75) Ekins, S.; Johnston, J. S.; Bahadduri, P.; D'Souzza, V. M.; Ray, A.; Chang, C.; Swaan, P. W. In vitro and pharmacophore based discovery of novel hPEPT1 inhibitors. *Pharm. Res.* **2005**, *22*, 512−517.
(76) Gomella, L.; Haist, S. *Clinician's Pocket Drug Reference*, *2004*; McGraw-Hill: New York, 2004.
(77) Wu, L. M.; Orikabe, M.; Hirano, Y.; Kawano, S.; Hiraoka, M. Effects of Na+ channel blocker, pilsicainide, on HERG current expressed in HEK-293 cells. *J. Cardiovasc. Pharmacol.* **2003**, *42* (3), 410−418.
(78) Kiesecker, C.; Zitron, E.; Luck, S.; Bloehs, R.; Scholz, E. P.; Kathofer, S.; Thomas, D.; Kreye, V. A.; Katus, H. A.; Schoels, W.; Karle, C. A.; Kiehn, J. Class Ia anti-arrhythmic drug ajmaline blocks HERG potassium channels: mode of action. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2004**, *370* (6), 423−435.
(79) Scholz, E. P.; Zitron, E.; Kiesecker, C.; Lueck, S.; Kathofer, S.; Thomas, D.; Weretka, S.; Peth, S.; Kreye, V. A.; Schoels, W.; Katus, H. A.; Kiehn, J.; Karle, C. A. Drug binding to aromatic residues in the HERG channel pore cavity as possible explanation for acquired long QT syndrome by antiparkinsonian drug budipine. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2003**, *368* (5), 404−414.
(80) Thomas, D.; Gut, B.; Karsai, S.; Wimmer, A. B.; Wu, K.; Wendt-Nordahl, G.; Zhang, W.; Kathofer, S.; Schoels, W.; Katus, H. A.; Kiehn, J.; Karle, C. A. Inhibition of cloned HERG potassium channels by the antiestrogen tamoxifen. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2003**, *368* (1), 41−48.
(81) Lacroix, P.; Crumb, W. J.; Durando, L.; Ciottoli, G. B. Prulifloxacin: in vitro (HERG current) and in vivo (conscious dog) assessment of cardiac risk. *Eur. J. Pharmacol.* **2003**, *477* (1), 69−72.
(82) Wang, J.; Salata, J. J.; Bennett, P. B. Saxitoxin is a gating modifier of HERG K+ channels. *J. Gen. Physiol.* **2003**, *121* (6), 583−598.
(83) Danielsson, B. R.; Lansdell, K.; Patmore, L.; Tomson, T. Phenytoin and phenobarbital inhibit human HERG potassium channels. *Epilepsy Res.* **2003**, *55* (1−2), 147−57.
(84) Ridley, J. M.; Milnes, J. T.; Zhang, Y. H.; Witchel, H. J.; Hancox, J. C. Inhibition of HERG K+ current and prolongation of the guinea-pig ventricular action potential by 4-aminopyridine. *J. Physiol* **2003**, *549* (Part 3), 667−672.
(85) Kikuchi, K.; Nagatomo, T.; Abe, H.; Kawakami, K.; Duff, H. J.; Makielski, J. C.; January, C. T.; Nakashima, Y. Blockade of HERG cardiac K+ current by antifungal drug miconazole. *Br. J. Pharmacol.* **2005**, *144* (6), 840−848.

(86) Zitron, E.; Kiesecker, C.; Scholz, E.; Luck, S.; Bloehs, R.; Kathofer, S.; Thomas, D.; Kiehn, J.; Kreye, V. A.; Katus, H. A.; Schoels, W.; Karle, C. A. Inhibition of cardiac HERG potassium channels by the atypical antidepressant trazodone. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2004**, *370* (2), 146−156.
(87) Choi, S. Y.; Koh, Y. S.; Jo, S. H. Inhibition of human ether-a-go-go-related gene K+ channel and IKr of guinea pig cardiomyocytes by antipsychotic drug trifluoperazine. *J. Pharmacol. Exp. Ther.* **2005**, *313* (2), 888−895.
(88) Danielsson, B. R.; Lansdell, K.; Patmore, L.; Tomson, T. Effects of the antiepileptic drugs lamotrigine, topiramate and gabapentin on hERG potassium currents. *Epilepsy Res.* **2005**, *63* (1), 17−25.
(89) Zitron, E.; Scholz, E.; Owen, R. W.; Luck, S.; Kiesecker, C.; Thomas, D.; Kathofer, S.; Niroomand, F.; Kiehn, J.; Kreye, V. A.; Katus, H. A.; Schoels, W.; Karle, C. A. QTc prolongation by grapefruit juice and its potential pharmacological basis: HERG channel blockade by flavonoids. *Circulation* **2005**, *111* (7), 835−838.
(90) Kuryshev, Y. A.; Ficker, E.; Wang, L.; Hawryluk, P.; Dennis, A. T.; Wible, B. A.; Brown, A. M.; Kang, J.; Chen, X. L.; Sawamura, K.; Reynolds, W.; Rampe, D. Pentamidine-induced long QT syndrome and block of hERG trafficking. *J. Pharmacol. Exp. Ther.* **2005**, *312* (1), 316−323.
(91) Dupuis, D. S.; Klaerke, D. A.; Olesen, S. P. Effect of beta-adrenoceptor blockers on human ether-a-go-go-related gene (HERG) potassium channels. *Basic Clin. Pharmacol. Toxicol* **2005**, *96* (2), 123−130.
(92) Claassen, S.; Zunkler, B. J. Comparison of the effects of metoclopramide and domperidone on HERG channels. *Pharmacology* **2005**, *74* (1), 31−36.
(93) 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), 883−891.
(94) Thomas, D.; Hammerling, B. C.; Wimmer, A. B.; Wu, K.; Ficker, E.; Kuryshev, Y. A.; Scherer, D.; Kiehn, J.; Katus, H. A.; Schoels, W.; Karle, C. A. Direct block of hERG potassium channels by the protein kinase C inhibitor bisindolylmaleimide I (GF109203X). *Cardiovasc. Res.* **2004**, *64* (3), 467−476.
(95) Sarazan, D. R.; Crumb, W. J., Jr.; Beasley, C. M., Jr.; Emmick, J. T.; Ferguson, K. M.; Strnat, C. A.; Sausen, P. J. Absence of clinically important HERG channel blockade by three compounds that inhibit phosphodiesterase 5--sildenafil, tadalafil, and vardenafil. *Eur J. Pharmacol.* **2004**, *502* (3), 163−167.
(96) Kang, J.; Chen, X. L.; Wang, H.; Ji, J.; Reynolds, W.; Lim, S.; Hendrix, J.; Rampe, D. Cardiac ion channel effects of tolterodine. *J. Pharmacol. Exp. Ther.* **2004**, *308* (3), 935−940.
(97) Traebert, M.; Dumotier, B.; Meister, L.; Hoffmann, P.; Dominguez-Estevez, M.; Suter, W. Inhibition of hERG K+ currents by antimalarial drugs in stably transfected HEK293 cells. *Eur. J. Pharmacol.* **2004**, *484* (1), 41−48.
(98) Thomas, D.; Wimmer, A. B.; Wu, K.; Hammerling, B. C.; Ficker, E. K.; Kuryshev, Y. A.; Kiehn, J.; Katus, H. A.; Schoels, W.; Karle, C. A. Inhibition of human ether-a-go-go-related gene potassium channels by alpha 1-adrenoceptor antagonists prazosin, doxazosin, and terazosin. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2004**, *369* (5), 462−472.
(99) Yuill, K. H.; Borg, J. J.; Ridley, J. M.; Milnes, J. T.; Witchel, H. J.; Paul, A. A.; Kozlowski, R. Z.; Hancox, J. C. Potent inhibition of human cardiac potassium (HERG) channels by the anti-estrogen agent clomiphene without QT interval prolongation. *Biochem. Biophys. Res. Commun.* **2004**, *318* (2), 556−561.
(100) Ridley, J. M.; Dooley, P. C.; Milnes, J. T.; Witchel, H. J.; Hancox, J. C. Lidoflazine is a high affinity blocker of the HERG K(+)channel. *J. Mol. Cell. Cardiol.* **2004**, *36* (5), 701−705.
(101) Thomas, D.; Hammerling, B. C.; Wu, K.; Wimmer, A. B.; Ficker, E. K.; Kirsch, G. E.; Kochan, M. C.; Wible, B. A.; Scholz, E. P.; Zitron, E.; Kathofer, S.; Kreye, V. A.; Katus, H. A.; Schoels, W.; Karle, C. A.; Kiehn, J. Inhibition of cardiac HERG currents by the DNA topoisomerase II inhibitor amsacrine: mode of action. *Br. J. Pharmacol.* **2004**, *142* (3), 485−494.
(102) Kim, E. J.; Kim, K. S.; Shin, W. H. Electrophysiological safety of DW-286a, a novel fluoroquinolone antibiotic agent. *Hum. Exp. Toxicol.* **2005**, *24* (1), 19−25.
(103) Gogelein, H.; Brendel, J.; Steinmeyer, K.; Strubing, C.; Picard, N.; Rampe, D.; Kopp, K.; Busch, A. E.; Bleich, M. Effects of the atrial antiarrhythmic drug AVE0118 on cardiac ion channels. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2004**, *370* (3), 183−192.
(104) Saegusa, N.; Sato, T.; Ogura, T.; Komuro, I.; Nakaya, H. Inhibitory effects of AMP 579, a novel cardioprotective adenosine A1/A2A receptor agonist, on native IKr and cloned HERG current. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2004**, *370* (6), 492−499.
(105) Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, Germany, 2000; Vol. 11.
(106) Kelly, C. A.; Dhaun, N.; Laing, W. J.; Strachan, F. E.; Good, A. M.; Bateman, D. N. Comparative toxicity of citalopram and the newer antidepressants after overdose. *J. Toxicol. Clin. Toxicol.* **2004**, *42* (1), 67−71.
(107) Isbister, G. K.; Bowe, S. J.; Dawson, A.; Whyte, I. M. Relative toxicity of selective serotonin reuptake inhibitors (SSRIs) in overdose. *J. Toxicol. Clin. Toxicol.* **2004**, *42* (3), 277−285.
(108) Ducroq, J.; Printemps, R.; Le Grand, M. Additive effects of ziprasidone and D,L-sotalol on the action potential in rabbit Purkinje fibres and on the hERG potassium current. *J. Pharmacol. Toxicol. Methods* **2005**, *52* (1), 115−122.
(109) Witchel, H. J.; Pabbathi, V. K.; Hofmann, G.; Paul, A. A.; Hancox, J. C. Inhibitory actions of the selective serotonin re-uptake inhibitor citalopram on HERG and ventricular L-type calcium currents. *FEBS Lett.* **2002**, *512* (1−3), 59−66.
(110) Hurst, R. S.; Higdon, N. R.; Lawson, J. A.; Clark, M. A.; Rutherford-Root, K. L.; McDonald, W. G.; Haas, J. V.; McGrath, J. P.; Meglasson, M. D. Dopamine receptor agonists differ in their actions on cardiac ion channels. *Eur. J. Pharmacol.* **2003**, *482* (1−3), 31−7.
(111) Kirsch, G. E.; Trepakova, E. S.; Brimecombe, J. C.; Sidach, S. S.; Erickson, H. D.; Kochan, M. C.; Shyjka, L. M.; Lacerda, A. E.; Brown, A. M. Variability in the measurement of hERG potassium channel inhibition: effects of temperature and stimulus pattern. *J. Pharmacol. Toxicol. Methods* **2004**, *50* (2), 93−101.
(112) Diaz, G. J.; Daniell, K.; Leitza, S. T.; Martin, R. L.; Su, Z.; McDermott, J. S.; Cox, B. F.; Gintant, G. A. The [3H]dofetilide binding assay is a predictive screening tool for hERG blockade and proarrhythmia: Comparison of intact cell and membrane preparations and effects of altering [K^+]o. *J. Pharmacol. Toxicol. Methods* **2004**, *50* (3), 187−199.
(113) Mitcheson, J. S.; Chen, J.; Lin, M.; Culberson, C.; Sanguinetti, M. C. A structural basis for the drug-induced log QT syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12329−12333.
(114) 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.
(115) Chen, X.; Rusinko, A., III; Young, S. S. Recursive partitioning analysis of a large structure−activity data set using three-dimensional descriptors. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1054−1062.
(116) Chen, X.; Rusinko, I. A.; Tropsha, A.; Young, S. S. Automated pharmacophore identification for large chemical data sets. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 887−896.
(117) Ekins, S.; Berbaum, J.; Harrison, R. K. Generation and validation of rapid computational filters for CYP2D6 and CYP3A4. *Drug Metab. Dispos.* **2003**, *31*, 1077−1080.
(118) Sheridan, R. P.; Feuston, B. P.; Maiorov, V. N.; Kearsley, S. K. Similarity to molecules in the training set is a good discriminator for prediction accuracy in QSAR. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (6), 1912−1928.
(119) Balakin, K. V.; Ivanenkov, Y. A.; Savchuk, N. P.; Ivaschenko, A. A.; Ekins, S. Comprehensive computational assessment of ADME properties using mapping techniques. *Curr. Drug Discovery Technol.* **2005**, *2*, 99−113.
(120) Clark, R. D.; Patterson, D. E.; Soltanshahi, F.; Blake, J. F.; Matthew, J. B. Visualizing substructural fingerprints. *J. Mol. Graphics Modell.* **2000**, *18* (4−5), 404−411, 527−532.
(121) Balakin, K. V.; Ekins, S.; Bugrim, A.; Ivanenkov, Y. A.; Korolev, D.; Nikolsky, Y.; Skorenko, S. A.; Ivashchenko, A. A.; Savchuk, N. P.; Nikolskaya, T. Kohonen maps for prediction of binding to human cytochrome P450 3A4. *Drug Metab. Dispos.* **2004**, *32*, 1183−1189.
(122) Korolev, D.; Balakin, K. V.; Nikolsky, Y.; Kirillov, E.; Ivanenkov, Y. A.; Savchuk, N. P.; Ivashchenko, A. A.; Nikolskaya, T. Modeling of human cytochrome P450-mediated drug metabolism using unsupervised machine learning approach. *J. Med. Chem.* **2003**, *46*, 3631−3643.
(123) So, S.-S.; Karplus, M. A comparitive study of ligand−receptor complex binding affinity prediction methods based on glycogen phosphorylase inhibitors. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 243−258.
(124) Ekins, S.; Durst, G. L.; Stratford, R. E.; Thorner, D. A.; Lewis, R.; Loncharich, R. J.; Wikel, J. H. Three-dimensional quantitative structure permeability relationship analysis for a series of inhibitors of rhinovirus replication. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1578−1586.
(125) Pospisil, P.; Ballmer, P.; Folkers, G.; Scapozza, L. Tautomerism of Nucleobase Derivatives and Their Score in Virtual Screening to Thymidine Kinase. Presented at the 224th National Meeting of the American Chemical Society, Boston, MA, August 18−22, 2002.

(126) Balakin, K. V.; Ekins, S.; Bugrim, A.; Ivanenkov, Y. A.; Korolev, D.; Nikolsky, Y.; Ivashchenko, A. A.; Savchuk, N. P.; Nikolskaya, T. Quantitative structure−metabolism relationship modeling of the metabolic N-dealkylation rates. *Drug Metab. Dispos.* **2004**, *32*, 1111−1120.
(127) Chang, C.; Pang, K. S.; Swaan, P. W.; Ekins, S. Comparative pharmacophore modeling of organic anion transporting polypeptides: A meta-analysis of rat Oatp1a1 and OATP1B1. *J. Pharmacol. Exp. Ther.* **2005**, *314*, 533−541.
(128) Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Three dimensional quantitative structure−activity relationships of inhibitors of P-glycoprotein. *Mol. Pharmacol.* **2002**, *61*, 964−973.
(129) Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Application of three-dimensional quantitative structure−activity relationships of P-glycoprotein inhibitors and substrates. *Mol. Pharmacol.* **2002**, *61*, 974−981.

JM060076R