

Estimation of hERG inhibition of drug candidates using multivariate property and pharmacophore SAR

Stephen R. Johnson,^{a,*} Hongwen Yue,^b Mary Lee Conder,^c Hong Shi,^b
Arthur M. Doweiko,^a John Lloyd^d and Paul Levesque^b

^aComputer-Assisted Drug Design, Bristol-Myers Squibb, PO Box 4000, Princeton, NJ 08543, USA

^bDiscovery Toxicology, Bristol-Myers Squibb, PO Box 4000, Princeton, NJ 08543, USA

^cDiscovery Biology, Bristol-Myers Squibb, PO Box 4000, Princeton, NJ 08543, USA

^dMedicinal Chemistry, Bristol-Myers Squibb, PO Box 4000, Princeton, NJ 08543, USA

Received 14 March 2007; revised 21 May 2007; accepted 12 June 2007

Available online 14 June 2007

Abstract—We describe the development of a computational model for the prediction of the inhibition of K⁺ flow through the hERG ion channel. Using a collection of 1075 discovery compounds with hERG inhibition measured in our standard patch-clamp electrophysiology assay, molecular features important for drug-induced inhibition were identified using a combination of statistical inference algorithms and manual hypothesis generation and testing. While many of the features used in the model reflect those referenced in the literature, several aspects of the model provide new insight into the role of physicochemical properties, electrostatics, and novel pharmacophores in hERG inhibition. Coefficients for these 10 features were then determined by least median squares regression, resulting in a model with an $R^2 \sim 0.66$ and RMS error (RMSe) of 0.47 log units for an external test set. Significant additional validation performed using a large collection of subsequent discovery data has been very encouraging with an $R^2 = 0.54$ and an RMSe of 0.63 log units. The performance of the model across several different chemotypes is demonstrated and discussed.
© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

Pharmacological blockade of I_{Kr}, the cardiac potassium ion current encoded by the human *ether-a-go-go* related gene (hERG), has been linked to a delayed membrane potential repolarization, prolonged action potential duration, and increased QT interval on the ECG.^{1–3} This increase in the QT interval is a major risk factor for *Torsades de pointes*,⁴ a serious cardiac arrhythmia occasionally resulting in death. Potent blockade of the hERG channel is now a primary concern in the drug discovery process, resulting in a correspondingly significant chemistry effort directed toward minimizing this undesirable activity.

The number of reviews^{5–11} and primary reports^{5,6,9,12–30} of computational models of hERG inhibition continues to grow. A particularly notable review was made by Jamieson et al.,³¹ who discuss practical avenues (includ-

ing modeling) for mitigating hERG inhibition. For the most part, there appears to be a general consensus on the primary drivers of hERG binding, namely Pi-cation interactions and usually some combinations of Pi-Pi and/or hydrophobic interactions with F656. Still, the diversity of potent inhibitors and the broad range of activity for compounds with typical hERG pharmacophores continue to be a vexing problem without a firmly established path for optimization.

The growth in the literature regarding how potential drugs block the hERG channel is testimony to the importance of this liability in the drug discovery process. The primary efforts are the hERG alanine-scanning experiments conducted by Sanguinetti and coworkers³² that identified the aromatic residues F656 and Y652 as playing a significant role in drug-induced blockade. Of the compounds studied thus far, only some lower potency hERG inhibitors have not been highly affected by mutation of Y652.^{33,34} Still more recently, additional mutagenesis studies have shown the importance of Pi-Cation interactions between ligands and the Y652 of hERG,²⁴ as well as the role of Y652 in voltage-dependent inhibition.^{34–36}

Keywords: hERG; Novelty detection; Pharmacophore model; Prolonged QTc; QSAR; In silico.

* Corresponding author. Tel.: +1 609 252 3003; fax: +1 609 252 6030; e-mail: stephen.johnson@bms.com

The desire for a computational model for predicting hERG inhibition is twofold. First, the broad SAR demonstrated by the array of potent hERG blockers has often frustrated medicinal chemistry efforts to work around this critical liability. A tool that can provide ideas and guidance to balance against target potency models and SAR represents a critical need in speeding discovery programs around this liability hurdle. Second, the most trusted assay technology for assessing hERG blockade, patch-clamp electrophysiology, is very labor intensive and struggles to keep pace with the number of compounds being considered in an increasingly high-throughput environment. An *in silico* tool is just one approach to solving this throughput dilemma along with FLIPR assays,³⁷ radioligand binding assays,³⁸ and more recently, automated patch-clamp instruments.³⁹ The availability of a tool to prioritize compounds for assay would be expected to greatly impact the costs of toxicity screening in discovery research.

The entire model development cycle used in this report was comprised of the iterative development of a series of candidate models (hypotheses). Existing compounds in our inventory were then selected for testing in the manual patch-clamp assay to challenge particular features. This usually was accomplished by selecting compounds with a narrow range of values for one feature, but with a broad diversity of values for the other features in the model. In some instances, compounds with combinations of descriptor values that were unusual in

the pre-existing data were sought for testing. Several features were eliminated in this process when the coefficient for the feature derived for the new data was not significant or if the coefficient was substantially different from the original hypothesis. At this point, a new hypothesis was generated, and the process repeated. This report discusses the outcome of this model development process.

2. Results and discussion

The 10 descriptors and coefficients for the model derived using the approach explained in the Methods Section are shown in Tables 1 and 2. The calculated versus observed plot for the training and test data is shown in Figures 3a and b. The model is a combination of both pharmacophoric and physicochemical descriptors, several of which are novel compared to the literature. The prediction results for the 1679 compounds that have been assayed since this model was developed are shown in Figure 4.

One obvious trend from Figure 3 is that the largest deviations are observed for compounds at the extremes of the activity range. This is typical of most incompletely solved regression models, and indicates that the model is an incomplete representation of the forces involved in ligand affinity for the hERG channel. It is also consistent with trends seen in Figure 1 for the IC₅₀ estimation from the single point inhibition data. We expect that both scenarios are probably contributing factors to the

Table 1. Descriptors and coefficients used in the prediction of hERG IC₅₀s

Description	Label	Coefficient	Geometry Dep. ^a
LogD at pH = 6.5	LogD _{6.5}	−1.576	N
Dry probe interaction volume at −1.4 kcal/mol	Vol_D7	−1.236	Y
LogP (Octanol-methylformamide)	LogP _{Oct/NMF}	−2.699	Y
Aromatic ESP Interaction volume at −5 kcal/mol	aESP, −5	1.084	Y
Binary interaction pair: two aromatic atoms 14 bonds apart	BIP2214	−0.298	N
Binary interaction pair: H-bond acceptor 11 bonds from an aromatic atom	BIP2611	−0.317	N
DDRR311342 ^b	DDRR	−0.293	Y
BR4 ^b	BR4	−0.171	Y
AAL334 ^b	AAL	0.192	Y
Secondary and tertiary amine indicator	Am	−0.955	N
Intercept		4.744	—
Training results		$R^2 \sim 0.65$, RMSe ~ 0.58	
Validation results		$R^2 \sim 0.66$, RMSe ~ 0.47	

^a Notes that this descriptor is sensitive to the conformation used in the calculation.

^b Table 2 further elucidates the pharmacophore descriptors.

Table 2. Definitions of the pharmacophoric descriptors used in the model

BR4	Basic center (B)	Lipophilic (L)		
Basic Center		6.0–9.0		
AAL334	H-Acceptor (A)	H-Acceptor (A)	Lipophilic (L)	
H-Acceptor	—	4.0–6.0	4.0–6.0	
H-Acceptor	—	—	6.0–9.0	
DDRR311342	H-Donor (D)	H-Donor (D)	Aromatic (R)	Aromatic (R)
H-Donor	—	6.0–9.0	2.5–4.0	2.5–4.0
H-Donor	—	—	6.0–9.0	9.0–13.0
Aromatic	—	—	—	4.0–6.0

Distances between pharmacophoric points shown in Angstroms.

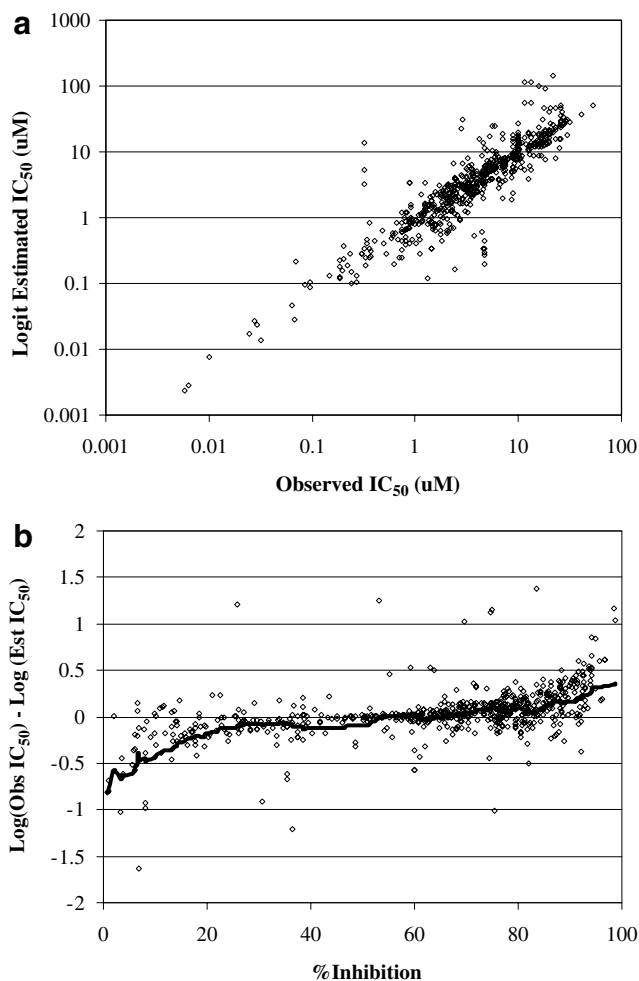


Figure 1. (a) The agreement between fully determined IC_{50} s from the hERG patch-clamp assay, and IC_{50} s estimated from single-point %-inhibition data. The $r^2 \sim 0.83$ with an RMS error of 0.27 log units. (b) Residuals of estimation versus %inhibition of the single point data. The line is a 20 compound moving average of the residuals.

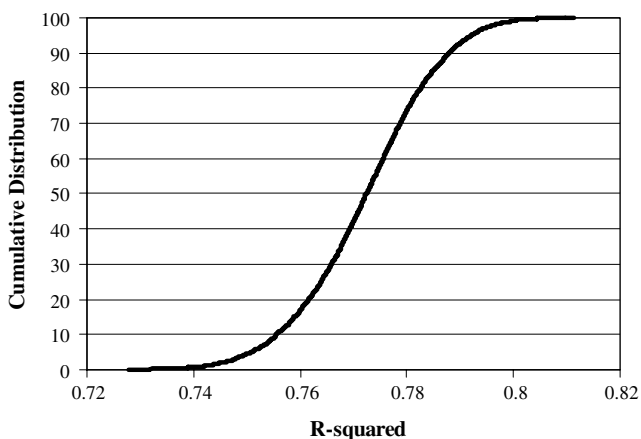


Figure 2. Distribution of the correlation coefficients from a Monte Carlo simulation of the effect of the error introduced by estimating IC_{50} s from the percent inhibition data. The simulation suggests that the upper bound on the in silico model performance is $r^2 \sim 0.8$. More likely, the actual performance must be lower as this simulation ignored other sources of experimental error.

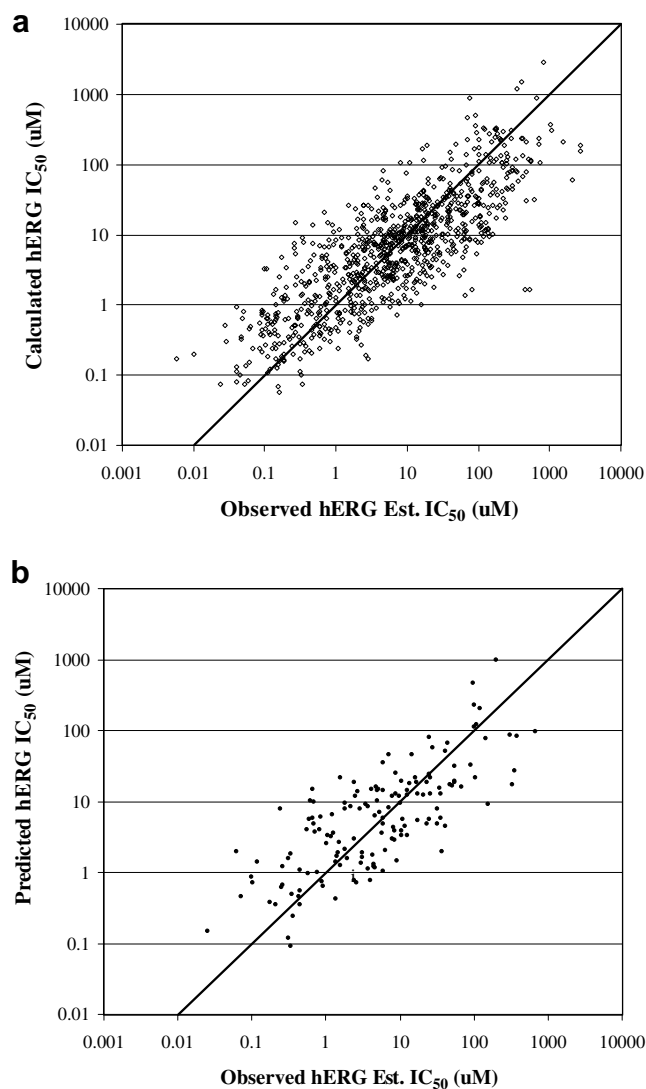


Figure 3. Calculated versus observed plots for the model defined in Table 1. (a) Results for the training data ($R^2 \sim 0.65$, RMSe ~ 0.58). (b) Results for the validation data ($R^2 \sim 0.66$, RMSe ~ 0.47).

quality of the fit here. Nonetheless, our experience has been that medicinal chemistry teams generally have less trouble decreasing hERG potency from the low nM to the low μ M range than from the low μ M range to double-digit μ M. For this reason, we believe the behavior at the extremes is somewhat less important than the behavior in the mid-potency range.

The validation data shown in Figure 4 are the more important tools for analyzing the performance of the model. These compounds were primarily tested in the natural course of discovery programs, although a subset of compounds was tested specifically to challenge the model in weakly populated descriptor space. The model predicts this large data set with an $R^2 = 0.54$ and an RMSe of prediction of 0.63. Compared to the training and original test data in Figure 3, there is wide scatter in the prediction plot. In addition, the skew at the extrema of the observed values are more pronounced than seen when the model was developed. Still, the overall

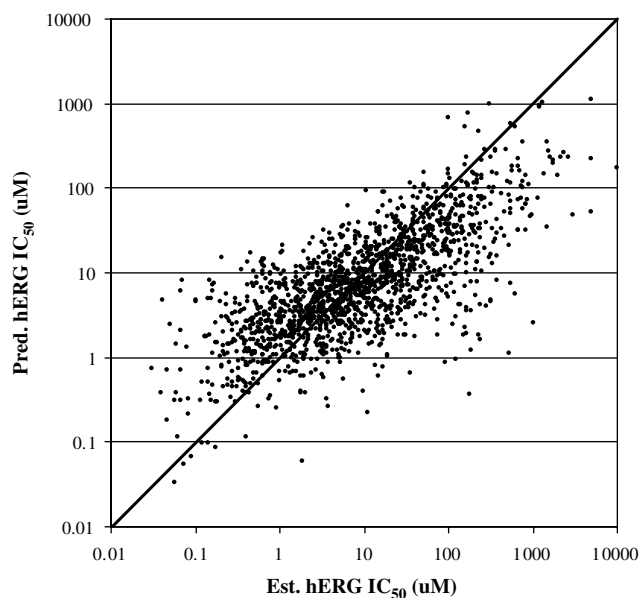


Figure 4. Prediction results on 1679 discovery compounds with hERG data measured post-model development. $R^2 = 0.54$, RMSE of Prediction = 0.63.

prediction is quite satisfactory as a true forward test of the model. More in-depth discussion of this validation data is given below.

Interpreting descriptors in QSAR models is always difficult for two basic reasons. First, the descriptors employed are often cryptic, and their interpretation is subject to the data used to derive the relationship. Second, QSAR models are inherently correlative and by assigning causality to the descriptors without significant additional information the user falls victim to a well-known logical fallacy. While empirical correlation is a necessary prerequisite for causation, it is not sufficient by itself to determine causation. Ideally, descriptors should serve as the basis for a hypothesis that can be tested to bolster or refute claims of causality. Still, the generation of such hypotheses is an important component of computational modeling.

The presence of a basic amine has been a feature in many computational models of hERG inhibition. Experimental evidence supports the hypothesis that a basic ionizable center can participate in a π -cation interaction with Y652.²⁴ The current model has two such features. The first is a simple indicator variable for a secondary or tertiary basic amine. The second is a pharmacophore, labeled as BR4, consisting of a basic center 6–9 Å from the centroid of an aromatic ring. Earlier versions of our model, developed when only a few hundred compounds were available, utilized a pharmacophore very similar to those in the literature.^{17,21,27} Each of these reports utilized basically the same pharmacophores with variations between the use of lipophilic points and aromatic points. In our hands, the statistical power of these pharmacophores has faded as additional compounds have been generated in the various discovery projects. More recent compounds that

contain the prototypical pharmacophore of a base with aromatic and lipophilic groups have had a wider distribution of hERG activity as medicinal chemistry optimization has focused on manipulating the properties of the aromatic systems or basic center that make up the pharmacophore. Additionally newer compounds that lack the distal hydrophobic points have maintained hERG potency. These two changes in trends over time highlight how the reductionist nature of pharmacophores can lead to dramatically different results depending on the particular collection of compounds used to derive them.

While searching for an explanation for the loss of power of the established hERG binding pharmacophore, a second pharmacophore was identified that seemed to directly mitigate the impact of the BR4 feature. Comprised of two H-bond acceptors and a lipophilic group, the AAL334 pharmacophore chiefly negates the impact of the BR4 pharmacophore when both are present simultaneously. This may prove a viable avenue for the optimization that allows medicinal chemists to retain a basic ionizable center. Additionally, compounds containing AAL334 appear to have slightly lower hERG potency in compounds not containing the BR4 pharmacophore. Figure 5 shows the cumulative distribution plots of hERG IC_{50} for compounds with respect to the presence or absence of BR4 and/or AAL334. The activity distribution for compounds with the BR4 pharmacophore but not the AAL334 pharmacophore is shifted toward more potent IC_{50} s. In contrast, compounds with both pharmacophores have an activity distribution that is nearly identical to those compounds that lack the BR4 pharmacophore. To our knowledge, no similar

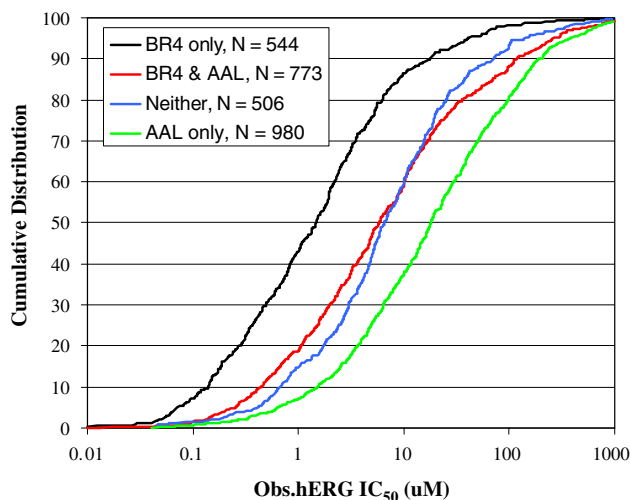


Figure 5. Cumulative distribution plots for the presence of the BR4 and AAL334 pharmacophores. Compounds containing the BR4 pharmacophore in the absence of the AAL334 pharmacophore are substantially more potent in the hERG assay than compounds that lack the BR4 pharmacophore. Conversely, compounds with the AAL334 pharmacophore appear substantially less potent on average. Compounds containing both pharmacophores have approximately the same activity distribution as those compounds containing neither pharmacophore.

pharmacophore has appeared in the literature as diminishing hERG activity.

Several additional features in the model highlight the importance of aromatic, lipophilic, and H-bond acceptor groups in hERG potency. While the precision of a 14-bond distance in BIP-6614 between aromatic atoms is difficult to justify, the parameter is identifying multiple aromatic rings that are at opposing ends of a molecule. This is an easily recognizable trait among many potent inhibitors of hERG, with Pi-stacking postulated as an important interaction with both F656 and Y652 in the pore of the channel. A combination of Pi-stacking to these residues coupled with the possibility of H-bond formation is captured in the BIP-2611 feature. Based on crude homology models derived from the KcsA and Kv1.2 structures (not shown), this may be consistent with H-bonds between the ligand and polar residues at the base of the selectivity filter (e.g., Thr623). Such interactions have been hypothesized previously.

Another pharmacophore present in the model is more novel compared to the literature. The feature labeled as DDRR contains two H-bond donors as well as two aromatic rings. Compounds possessing this feature are, on average, 3-fold more potent than compounds that do not contain this pharmacophore. It is actually not a trivial exercise to rationalize this pharmacophore with a homology model of hERG. Few opportunities are apparent for an H-bond donor making interactions with sidechains. It is possible for such interactions to form with Thr623, although the equivalent feature using an H-bond acceptor does not show a propensity toward increased hERG potency.

One feature of the model that arose specifically as a result of a series of compounds from a discovery project was the electrostatic potential (ESP) around aromatic rings contoured at -5 kcal/mol. A number of papers suggesting the importance of electrostatic potential in aromatic Pi-stacking have appeared in the literature.^{59–62} We observed a trend among a series of compounds in which compounds with electron withdrawing groups around a phenyl ring had a 5- to 30-fold increase in potency over that of the unsubstituted ring. Compounds with electron donating groups attached to the ring had a 2- to 5-fold decrease in potency compared to the unsubstituted ring. Based on this and similar trends in our corporate database, several series of compounds were selected for testing, and the trend appeared to be generally consistent. A feature (aESP, -5) was designed specifically to capture aromatic rings that were relatively electron deficient. This feature is now among the most important features in the model based on coefficient, and has already found significant use in internal discovery projects as an avenue for optimization.

Jamieson et al.³¹ propose the manipulation of the electrostatics around aryl rings as an avenue toward remediation of hERG liabilities. In contrast to our results, they propose the addition of electron withdrawing groups to decrease hERG potency. Perry et al.⁶³ discuss the effect of changing the *para* substituent on a series of clofilium

analogs. They observe that the potency is increased with a more polarizable group at the *para* position. This is more consistent with our findings. A number of the theoretical studies of aromatic Pi-stacking do show that the addition of electron donating groups can make Pi-stacking more favorable. However, these studies are performed relative to benzene, whereas phenol would be a better model system for interactions with Y652.

It has been appreciated for some time that the formal charge on a compound can have great impact on hERG inhibition. In general compounds possessing an acidic group have greatly diminished potency. This comes as little surprise as potassium ion channels are designed to stabilize a monovalent cation in the water filled cavity.⁶⁴ Nonetheless, not all acidic compounds are completely inactive against hERG. In our database acidic compounds are as potent as 60% inhibition at $1\text{ }\mu\text{M}$ (est. $\text{IC}_{50} \sim 0.6\text{ }\mu\text{M}$) with the median potency being 24% inhibition at $30\text{ }\mu\text{M}$ (est. $\text{IC}_{50} \sim 95\text{ }\mu\text{M}$). In contrast, the median potency of basic compounds is $4.7\text{ }\mu\text{M}$ and for neutral compounds the median potency is 50% inhibition at $10\text{ }\mu\text{M}$ (est. $\text{IC}_{50} \sim 10\text{ }\mu\text{M}$). The $\text{LogP}_{\text{Oct/NMF}}$ feature (octanol-*N*-methylformamide partition coefficient) captures this trend quite well, as well as identifying compounds with low $\text{LogD}_{6.5}$ that still show significant potency. This feature is calculated using the free energies of solvation calculated by OmniSol. In contrast to the $\text{LogD}_{6.5}$ feature where the distributions for acidic and basic compounds are nearly identical, values for the $\text{LogP}_{\text{Oct/NMF}}$ of the acidic compounds are substantially more negative than for the basic compounds. A sensitivity analysis of the correlation of the $\text{LogP}_{\text{Oct/NMF}}$ to the hERG IC_{50} was performed by manipulating the solvent parameters of *N*-methylformamide used in Omnisol. This analysis showed that the correlation was highly sensitive to the values of Abraham's H-bond acidity parameter,⁶⁵ $\sum \alpha_2^{\text{H}}$, and somewhat less sensitive to the H-bond basicity parameter.

The Volsurf feature, D7, captures the impact of hydrophobicity on affinity for the hERG channel. This may be a representation of the amount and character of hydrophobic surface in the ligand that could be buried in accordance with the hydrophobic effect. In any event, increasing lipophilicity is a well-recognized route to increased potency against hERG.³¹

We have clustered the 1679 validation compounds based on Daylight Fingerprint⁶⁶ Tanimoto using an average linkage method with a similarity cutoff of 0.7. Generally each cluster contains only a single chemotype, although a chemotype can be split across more than one cluster. This view of the data provides an analysis of how the model would perform within a group of compounds that would be generated within a single discovery project. Table 3 shows the prediction statistics for the 35 largest clusters of compounds, representing approximately 50% of the validation compounds. While the quality of the prediction results varies quite dramatically from cluster to cluster, the results support the use of the model in generating hypotheses of how to approach hERG optimization within many projects. In fact, the model has

Table 3. Model performance for clusters of related compounds

Cluster	R^2	RMSe	Spearman Rho	Range ^a	Num	Max. sim. to training data ^b	Average potency ^c	Most potent IC ₅₀ ^c
0	0.32	0.45	0.50	3.38	84	0.40	3.01	0.07
1	0.37	0.49	0.65	2.14	70	0.96	1.31	0.15
2	0.69	0.75	0.84	2.92	50	0.52	19.27	0.92
3	0.76	0.47	0.90	3.63	47	0.46	37.18	0.35
4	0.44	0.64	0.60	2.62	45	0.76	90.99	2.19
5	0.55	0.43	0.75	2.34	42	0.98	1.11	0.08
6	0.34	0.96	0.59	2.98	38	0.64	1.22	0.04
7	0.50	0.42	0.62	2.28	28	0.99	13.86	1.17
8	0.59	0.60	0.66	2.89	25	0.75	0.89	0.06
9	0.71	0.47	0.85	2.36	24	0.98	9.18	0.46
10	0.71	0.53	0.83	3.14	23	0.89	34.99	0.27
11	0.37	0.64	0.67	2.63	22	0.42	51.95	4.09
12	0.45	0.26	0.69	1.18	19	0.49	6.01	2.77
13	0.46	0.38	0.68	1.46	18	0.89	6.66	1.19
14	0.86	0.38	0.94	2.60	18	0.46	5.70	0.39
15	0.42	0.80	0.53	1.48	17	0.95	2.07	0.35
16	0.45	0.50	0.70	1.89	17	0.41	12.33	2.81
17	0.40	0.71	0.55	1.87	16	0.96	1.29	0.21
18	0.45	1.30	0.71	2.45	15	0.42	47.63	3.72
19	0.42	0.40	0.47	1.64	15	0.35	19.37	4.71
20	0.86	0.91	0.89	4.52	15	0.99	54.18	0.15
21	0.23	0.33	0.39	1.45	14	0.40	5.06	0.72
22	0.72	0.71	0.72	3.15	14	0.38	4.61	0.30
23	0.85	0.67	0.96	3.79	14	0.37	1.69	0.14
24	0.01	0.63	0.05	2.16	14	0.41	3.51	0.15
25	0.49	0.35	0.60	1.72	12	0.44	7.54	1.02
26	0.50	0.97	0.64	2.42	12	0.50	69.12	3.55
27	0.01	0.61	0.10	0.75	12	0.34	39.53	16.06
28	0.61	0.54	0.64	2.66	12	0.49	19.35	1.19
29	0.13	0.54	0.35	1.01	11	0.90	27.02	9.12
30	0.92	0.40	0.87	1.62	10	0.91	3.47	0.45
31	0.19	0.78	0.37	2.41	10	0.42	11.11	2.08
32	0.29	0.48	0.55	1.11	10	0.37	2.31	1.01
33	0.48	0.32	0.48	1.55	10	0.33	3.28	0.24
34	0.52	0.62	0.72	2.24	10	0.66	25.68	2.96
Average	0.49	0.58	0.63		813			
Median	0.46	0.54	0.65					
Std. dev.	0.23	0.22	0.21					

^a Range of Log₁₀(hERG IC₅₀) of compounds within each cluster.^b Highest Daylight Tanimoto between a compound in the cluster and any compound in the training data used to generate the model.^c hERG potency in μ M.

been used in the optimization efforts for several of the clusters of compounds in Table 3.

Figure 6 shows the prediction results for compounds in the largest cluster of compounds, cluster 0. The discovery project responsible for these compounds frequently used the model in the optimization of hERG potency. The graphical results shown in the figure appear much better than the R^2 of 0.32 implies. Circled in the figure are a group of compounds predicted to be substantially less active than the observed data. These compounds all contain either a pyrimidine or a pyrazine in place of a particular phenyl ring present in the well-predicted compounds. This particular error is also observed in other unrelated clusters of compounds. We believe that this is a result of the simplistic nature of the ESP feature used to encode the importance of Pi-stacking interactions. The view of the ESP used here was specifically conceived to capture Pi-stacking interactions for phenyl rings. There is ample evidence for Pi-stacking in other

aromatic rings, although the ESP plots for heterocycles discussed in the literature are often quite different from those for a phenyl ring. While the particular quantitative estimate from the model is subject to these errors the underlying concept of modulating the ESP of the ring has played an important role in optimizing hERG liabilities in several projects. Future model refinement will entail a more sophisticated treatment of these interactions.

Also evident from Table 3 is that there is no single statistic that adequately summarizes the utility of a model. For example, cluster 11 has a relatively low R^2 of 0.37, but an acceptable ability to rank order compounds based on the Spearman's Rho of 0.67. The same is true for cluster 1, although these compounds are much more similar to the training data. The predictions for cluster 20, however, correlate very highly ($R^2 = 0.86$) with the observed data. However, the RMSe of prediction of 0.91 is substantially higher than the overall RMSe of 0.63. While this is well known to those skilled in the

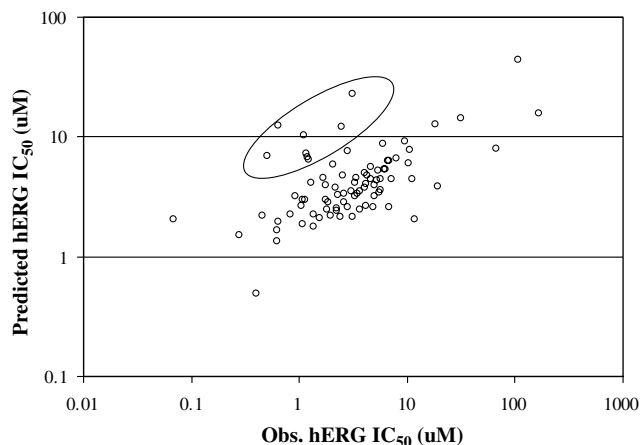


Figure 6. Predicted versus observed plot for the compounds in Cluster 0 (see Table 3). The circled compounds contain either a pyrimidine or a pyrazine in place of a particular phenyl ring present in the well-predicted compounds. This particular error is also observed in other unrelated clusters of compounds. See the text for more discussion.

art of modeling, it is frequently a barrier to the exploitation of the model by discovery projects. Often, users will latch onto whatever statistical quantity illustrates their own bias in favor of or against utilization within a project.

There is considerable debate in the modeling community about how to identify a compound as being similar to the training data, and thus likely to be predicted reliably. Table 3 shows the maximum Daylight fingerprint Tanimoto similarity of any compound in each cluster to a compound in the training data used to generate the model. Figure 7 shows that there is no correlation between the R^2 of prediction for the compounds in a cluster, and the similarity of the compounds in that cluster to the training data. In contrast, the overall correlation for the validation set improves to $R^2 = 0.72$ when

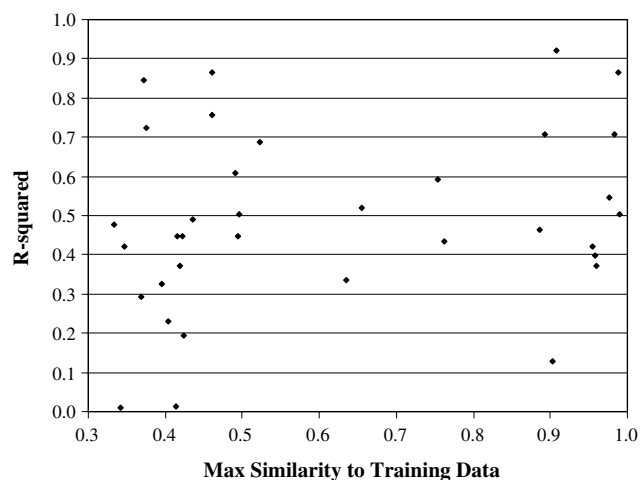


Figure 7. Relationship between topological similarity (Daylight Fingerprints) of validation compound clusters and training data and the correlation between observed and predicted IC_{50} s. There is no obvious relationship between similarity to the training data, as measured using Daylight Fingerprints, and the correlation between predicted and measured data.

considering only those compounds with a Tanimoto similarity greater than 0.65 to a training set member. This disparity arises because many of the clusters with low similarity to the training data have significantly different slopes or constants (intercepts) for the predicted versus observed plot than those of the high similarity clusters even while the correlation coefficient for the cluster may be acceptable.

3. Conclusions

We have presented an *in silico* model of hERG inhibition that is based upon a large diverse collection of discovery compounds. Much of the information used in the model is consistent with previously established literature. However, the model does include a number of factors that have not been considered explicitly previously. The AAL334 pharmacophore largely negates the impact of the common basic amine and aromatic system that is so prevalent among hERG inhibitors. Another is the use of electrostatic potentials to assess potential π -stacking interactions with Y652 and/or F656. The inclusion of the $\text{LogP}_{\text{Oct/NMF}}$ improves the prediction over the range of activities seen for acidic and other polar compounds. Each of these features suggests different avenues for medicinal chemistry optimization of hERG affinity.

Further, the ability of the model to predict across an array of different chemistries was demonstrated. While much room for improvement remains, the capability for quality predictions across several different chemotypes allows discovery projects to leverage knowledge from one project to advance another. By analyzing the model performance within clusters of compounds specific weaknesses of the model are much more easily brought to light. This facilitates the revision of the assumptions made during model generation, allowing for a continual improvement and retention of knowledge as it is generated within discovery.

The derivation of a model of this sort requires close analysis of the data throughout the process. Frequently in drug discovery, incomplete data are collected that, while it meets the needs of many users, presents significant challenges for use in modeling.⁶⁷ We presented a simple Monte Carlo method capable of providing insight on the impact of data manipulation and/or data reproducibility. Indeed, all too often in the QSAR literature data are fit to a greater degree than is reasonable given the quality of the data available. Simple methods such as that discussed here can provide guidance as to when a model has reached the degree of precision and accuracy supported by the underlying data.

There are a number of avenues by which the model could be refined and improved. Improvements to the electrostatic potential descriptors used to describe π -stacking interactions are needed to better account for heterocycles. It is also likely that different ESP features intended to differentially encode face-to-face or face-to-edge interactions may significantly improve our understanding of hERG inhibition. Other

improvements would include coupling the ligand-based model to a structure-based approach to better capture stereochemical differences and steric constraints, as well as more sophisticated pharmacophoric descriptors. All of these are currently areas of active research within our effort.

4. Methods

4.1. hERG assay

Human embryonic kidney (HEK293) cells were stably transfected with human Ether-a-go-go Related Gene (hERG) cDNA for use in the hERG assay. The biophysical and pharmacological properties of recombinant hERG channels expressed in HEK293 cells and of native I_{Kr} channels in human cardiac cells are nearly identical. Several known hERG blockers, including dofetilide, terfenadine, cisapride, and E-4031, inhibit recombinant hERG currents in our hERG stable cell line and I_{Kr} current in cardiac myocytes with the same potency.

Membrane current recordings were made with an Axopatch 200 series integrating patch-clamp amplifier (Axon Instruments, Foster City, CA) using the whole-cell variant of the patch-clamp technique. For hERG current recording the bath solution, which replaced the cell culture media during experiments, contained: 140 mM NaCl, 4 mM KCl, 1.8 mM CaCl_2 , 1 mM MgCl_2 , 10 mM glucose, and 10 mM Hepes (pH 7.4, NaOH). Borosilicate glass pipettes had tip resistances of 2–4 M Ω when filled with an internal solution containing: 130 mM KCl, 1 mM MgCl_2 , 1 mM CaCl_2 , 5 mM ATP- K_2 , 10 mM EGTA, and 10 mM Hepes (pH 7.2, KOH).

hERG-expressing cells were placed in a plexiglass bath chamber, mounted on the stage of an inverted microscope, and perfused continuously with bath solution. To determine potency of test agents for inhibiting hERG current, repetitive test pulses (0.05 Hz) were applied from a holding potential of -80 mV to $+20$ mV for 2 s. Tail currents were elicited following the test pulses by stepping the voltage to -65 mV for 3 s. After recording the steady state current for 2–5 min in the absence of test agent, the bath solution was switched to one containing the lowest concentration of the agent to be used. The peak tail current was monitored until a new steady-state in the presence of test agent was achieved. This was followed by the application of the next higher concentration of the agent to be tested, and this was repeated until all concentrations of test agent had been evaluated. Percent inhibition of tail currents was plotted as a function of test agent concentration to quantify hERG channel inhibition. Compound effects were calculated using tail currents because there are no endogenous tail currents in plasmid-transfected control HEK293 cells. Data were sampled at rates at least two times the low pass filter rate. The flow rate was kept constant throughout the experiments (~ 1 to 5 mL/min). All membrane currents were recorded at room temperature ($\sim 25^\circ\text{C}$).

4.2. Computational methods

4.2.1. Training and validation data. The final collection of molecular descriptors and coefficients were derived using a collection 1075 compounds with either IC_{50} s (289 compounds) or percent inhibition at a single concentration (786 compounds). These 1075 compounds were randomly divided into a training set of 925 compounds and test set of 150 compounds for the purposes of model derivation and testing.

In the time since the model was developed, an additional 1679 compounds have been tested in the manual whole cell patch-clamp hERG inhibition assay. These include 324 IC_{50} s and 1355 single-point percent inhibition measurements. This second data set was used as a true forward-looking validation set as there were no hERG inhibition data available for these compounds at the time of model development.

4.2.2. Estimation of IC_{50} s from percent inhibition. As noted above, the data set used in deriving the model of hERG inhibition was composed of both fully determined IC_{50} s and percent inhibition at a single concentration. To facilitate modeling, the percent inhibition data were transformed into estimated IC_{50} s using the Logit^{40–43} function

$$\text{Est. IC}_{50} \equiv \frac{100 - \%inh}{\%inh} \times \text{Conc.} \quad (1)$$

where $\%inh$ is the percent inhibition measured at the concentration Conc. The Logit function makes the implicit assumption that at zero concentration there is no inhibition, while at some infinite concentration there is complete inhibition of potassium ion flow through the hERG channel. It also assumes a universal slope to the estimated IC_{50} curve, which is a substantial source of error in estimating the IC_{50} .

As this estimation of the IC_{50} for the single-point measurements could introduce significant error, a brief analysis of the transformed data was performed. Figure 1a shows the relationship between the 613 IC_{50} measurements available when writing this manuscript and the estimated IC_{50} s that are derived from single-point inhibition data. Overall the relationship is quite good, with a root mean square error (RMSE) of 0.27 log units (μM) and an $r^2 \sim 0.83$. Figure 1b is a plot of the residuals of the measured IC_{50} and the estimated IC_{50} as a function of the percent inhibition at whatever concentration they were determined. A few important observations are drawn from this analysis. First, compounds with percent inhibition between 20% and 80% are well predicted using this approach (median error is 1.3-fold on a μM basis). Second, compounds with very low or very high inhibition are predicted fairly poorly, with median error of 2.7-fold on a μM basis. The maximum error for any transformation was approximately 45-fold.

These projected error rates were used in a Monte Carlo simulation to gauge their effect on the quality of the data used to derive the *in silico* model. The estimated IC_{50} s for the 786 compounds without true IC_{50} s were

perturbed by a normal distribution of error, using the moving averages shown in Figure 1b based on the %-inhibition used to estimate the IC_{50} as dictated above. The correlation coefficient (r^2) was then calculated between the original estimated data and the perturbed data. This process was repeated for 5000 iterations, building a distribution of correlations. The cumulative distribution plot shown in Figure 2 indicates the median r^2 is approximately 0.77 based on the expected error due to the estimated IC_{50} s. A correlation greater than 0.8 is extremely unlikely and is likely the upper bound on the performance of an in silico model. This estimate is probably overly optimistic as it assumes no error in the measured IC_{50} s and does not account for the peak errors of 45-fold for some single-point to IC_{50} estimates for the IC_{50} data.

4.2.3. Conformation selection. Initial conformations were generated using Omega⁴⁴ followed by a minimization using Batchmin⁴⁵ using OPLS2005. Conformers within 10 kJ/mol of the minimum energy conformation are retained for use in descriptor generation. This conservative cutoff was selected to limit the ability of molecules to adopt strained conformations in order to match potential pharmacophores.

4.2.4. Descriptor generation. A multitude of physico-chemical descriptors were calculated for consideration as predictors of hERG inhibition. Included among these were Volsurf^{46–49} descriptors using the Dry and H₂O probes. These descriptors encode hydrophobic and hydrophilic interaction volumes for compounds and have been suggested as useful features in developing predictive ADMET models. Also included were free energies of solvation in various solvents as calculated using the OmniSol program.⁵⁰ Finally, the calculated LogD values at pH = 6.5 and pH = 7.4 were obtained using the ACD/LogD software.⁵¹ For each of these features, only the lowest energy conformation generated as described above was utilized.

A number of binary interaction pair descriptors were also calculated. These descriptors are essentially two-point pharmacophore descriptors with through-bond distances. It is expected that they might identify structural groups important for binding along with approximate orientations captured by the bond distances.

Many potential three-dimensional pharmacophores were also generated. For the pharmacophoric descriptors, each of the retained conformers was evaluated for a match. A pharmacophore was considered to be present if it was present in any of the conformations generated above. A pool of pharmacophores was identified for further analysis by searching for pharmacophores present in highly potent blockers ($IC_{50} < 1 \mu M$) in the training set in a substantially higher proportion than found among the non-potent compounds ($IC_{50} > 30 \mu M$). Pharmacophores were eliminated from the pool either because of low frequency in the data set as a whole, or were determined to be generally uninteresting in that they contained only lipophilic moieties.

The electrostatic potential around aromatic rings was also utilized as a descriptor. The descriptor was calculated using the lowest energy conformation identified as described above following OPLS2005 minimization. AM1-BCC⁵² charges were calculated using the Molcharge program, a part of the QuACPAC⁵³ suite of tools. The compound was placed on a 1 Å grid, and the electrostatic potential was calculated at each grid point. The grid points closer to an aromatic atom than to any non-aromatic atom with an ESP below a threshold of either -1, -5 or -10 kcal/mol are counted. This count is an approximate volume of electrostatic potential proximal to aromatic groups.

4.2.5. Feature selection. The current model represents several cycles of hypothesis generation and testing. Consequently, the features present in the model were selected over time rather than in a single optimization method as might be typical for a ligand-based QSAR model. Substantial effort was directed at descriptor selection. Descriptors were selected using either an automated supervised selection algorithm or selected manually based on observed trends in discovery projects.

The supervised descriptor selection algorithm was developed in-house. This algorithm, similar to others in the literature, first removes descriptors with greater than 90% identical values. A complete pairwise correlation matrix is then generated for the entire remaining descriptor pool. Subsets of descriptors are selected with the constraint that no two included descriptors have a pairwise $r^2 > 0.9$.

The supervised algorithm uses simulated annealing with a leave-10%-out PRESS cross-validation function to evaluate potential models. Initially, a random subset of descriptors is chosen. Following this, a single descriptor is removed and replaced with another feature from the pool. This replacement is done randomly, with the exception that the new feature cannot be correlated with any other feature already present in the model. The PRESS score is then evaluated for this new model. If the score is an improvement over the previous model, the result is accepted and the procedure repeats. If the model is not an improvement, the model may still be accepted based on a probability derived from the Boltzmann distribution.

The algorithm used here fixes the initial temperature such that 80% of the detrimental steps are accepted originally.⁵⁴ The temperature is then decreased by 25% every 1000 iterations. The selection process halts when no model is accepted for 900 iterations.

The descriptors selected by the automated routine were each manually analyzed before being included in the final model. For example, descriptors are often highly correlated with other available descriptors. It is often difficult to determine which of several correlated descriptors should be included in a model. Descriptors that appeared to make more physical sense with respect to existing knowledge and/or literature for hERG

inhibition were preferred over correlated descriptors with unclear interpretations.

Several descriptors were selected manually in response to a particular trend seen within a discovery project. A primary example in the model would be the electrostatic potential descriptor. This descriptor was included based on the observation of a 5-fold increase in potency for a 4-fluoro substitution relative to an unsubstituted phenyl, which was then followed by 17-fold increase in potency for the 4-chloro substitution. In contrast, a 2-fold and 3-fold loss in activity was observed for the 4-amino and 4-methyl substitutions, respectively. A search of the corporate database indicated that this trend was fairly commonplace although does not appear to be prevalent in the literature.

4.2.6. Model training. Once descriptors were selected, they were utilized in a least median squares (LMS) regression.^{55–58} Descriptors were range scaled so that the minimum value for each descriptor was 0.0 and the maximum value was 1.0. LMS regression is a robust regression method that de-emphasizes leverage points in determining coefficients. The method searches for the set of coefficients that minimizes the median squared residual, rather than the sum of the squared residuals as in ordinary multiple regression. This is accomplished by a stochastic procedure in which small subsets of observations are selected to generate the model coefficients. The coefficients are then evaluated by determining the median squared residual for all of the observations. The implementation used here was programmed in our laboratory.

References and notes

- Sanguinetti, M. C.; Jiang, C.; Curran, M. E.; Keating, M. T. *Cell (Cambridge, Massachusetts)* **1995**, *81*, 299–307.
- Trudeau, M. C.; Warmke, J. W.; Ganetzky, B.; Robertson, G. A. *Science (Washington, DC)* **1995**, *269*, 92–95.
- Keating, M. T.; Sanguinetti, M. C. *Cell (Cambridge, MA, United States)* **2001**, *104*, 569–580.
- Viskin, S. *Lancet* **1999**, *354*, 1625–1633.
- Recanatini, M.; Cavalli, A.; Masetti, M. *Novartis Found. Symp.* **2005**, *266*, 171–185.
- Aronov, A. M. *Drug Discov. Today* **2005**, *10*, 149–155.
- Recanatini, M.; Poluzzi, E.; Masetti, M.; Cavalli, A.; de Ponti, F. *Med. Res. Rev.* **2005**, *25*, 133–166.
- Vaz, R. J.; Li, Y.; Rampe, D. *Prog. Med. Chem.* **2005**, *43*, 1–18.
- Li, Y.; Cianchetta, G.; Vaz, R. J. *Methods Princ. Med. Chem.* **2006**, *29*, 428–443.
- Stansfeld, P. J.; Sutcliffe, M. J.; Mitcheson, J. S. *Expert Opin. Drug Metab. Toxicol.* **2006**, *2*, 81–94.
- Sanguinetti, M. C.; Mitcheson, J. S. *Trends Pharmacol. Sci.* **2005**, *26*, 119–124.
- Aptula, A. O.; Cronin, M. T. D. *SAR QSAR Environ. Res.* **2004**, *15*, 399–411.
- Aronov, A. M.; Goldman, B. B. *Bioorg. Med. Chem.* **2004**, *12*, 2307–2315.
- Aronov, A. M. *J. Med. Chem.* **2006**, *49*, 6917–6921.
- Bains, W.; Basman, A.; White, C. *Prog. Biophys. Mol. Biol.* **2004**, *86*, 205–233.
- Becker, O. M.; Dhanoa, D. S.; Marantz, Y.; Chen, D.; Shacham, S.; Cheruku, S.; Heifetz, A.; Mohanty, P.; Fichman, M.; Sharadendu, A.; Nudelman, R.; Kauffman, M.; Noiman, S. *J. Med. Chem.* **2006**, *49*, 3116–3135.
- Cavalli, A.; Poluzzi, E.; De Ponti, F.; Recanatini, M. *J. Med. Chem.* **2002**, *45*, 3844–3853.
- Choe, H.; Nah, K. H.; Lee, S. N.; Lee, H. S.; Lee, H. S.; Jo, S. H.; Leem, C. H.; Jang, Y. J. *Biochem. Biophys. Res. Commun.* **2006**, *344*, 72–78.
- Coi, A.; Massarelli, I.; Murgia, L.; Saraceno, M.; Calderone, V.; Bianucci, A. M. *Bioorg. Med. Chem.* **2006**, *14*, 3153–3159.
- Cianchetta, G.; Li, Y.; Kang, J.; Rampe, D.; Fravolini, A.; Cruciani, G.; Vaz, R. J. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3637–3642.
- Ekins, S.; Crumb, W. J.; Sarazan, R. D.; Wikel, J. H.; Wrighton, S. A. *J. Pharmacol. Exp. Ther.* **2002**, *301*, 427–434.
- Ekins, S.; Balakin, K. V.; Savchuk, N.; Ivanenkov, Y. J. *J. Med. Chem.* **2006**, *49*, 5059–5071.
- Farid, R.; Day, T.; Friesner, R. A.; Pearlstein, R. A. *Bioorg. Med. Chem.* **2006**, *14*, 3160–3173.
- Fernandez, D.; Ghanta, A.; Kauffman, G. W.; Sanguinetti, M. C. *J. Biol. Chem.* **2004**, *279*, 10120–10127.
- Fioravanzo, E.; Cazzolla, N.; Durando, L.; Ferrari, C.; Mabilia, M.; Ombrato, R.; Parenti, M. D. *Internet Electronic J. Mol. Des.* **2005**, *4*, 625–646.
- Gepp, M. M.; Hutter, M. C. *Bioorg. Med. Chem.* **2006**, *14*, 5325–5332.
- 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. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1829–1835.
- Rajamani, R.; Tounge, B. A.; Li, J.; Reynolds, C. H. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1737–1741.
- Song, M.; Clark, M. J. *Chem. Inf. Model.* **2006**, *46*, 392–400.
- Seierstad, M.; Agrafiotis, D. K. *Chem. Biol. Drug Des.* **2006**, *67*, 284–296.
- Jamieson, C.; Moir, E. M.; Rankovic, Z.; Wishart, G. J. *J. Med. Chem.* **2006**, *49*, 5029–5046.
- Mitcheson, J. S.; Chen, J.; Lin, M.; Culberson, C.; Sanguinetti, M. C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12329–12333.
- Witchel, H. J.; Dempsey, C. E.; Sessions, R. B.; Perry, M.; Milnes, J. T.; Hancox, J. C.; Mitcheson, J. S. *Mol. Pharmacol.* **2004**, *66*, 1201–1212.
- Milnes, J. T.; Crociani, O.; Arcangeli, A.; Hancox, J. C.; Witchel, H. J. *Br. J. Pharmacol.* **2003**, *139*, 887–898.
- Sanchez-Chapula, J. A.; Navarro-Polanco, R. A.; Culberson, C.; Chen, J.; Sanguinetti, M. C. *J. Biol. Chem.* **2002**, *277*, 23587–23595.
- Ferrer-Villada, T.; Navarro-Polanco, R. A.; Rodriguez-Menchaca, A. A.; Benavides-Haro, D. E.; Sanchez-Chapula, J. A. *Eur. J. Pharmacol.* **2006**, *531*, 1–8.
- Tang, W.; Kang, J.; Wu, X.; Rampe, D.; Wang, L.; Shen, H.; Li, Z.; Dunnington, D.; Garyantes, T. *J. Biomol. Screen.* **2001**, *6*, 325–331.
- Finlayson, K.; Turnbull, L.; January, C. T.; Sharkey, J.; Kelly, J. S. *Eur. J. Pharmacol.* **2001**, *430*, 147–148.
- Wood, C.; Williams, C.; Waldron, G. J. *Drug Discov. Today* **2004**, *9*, 434–441.
- Gao, F.; Johnson, D. L.; Ekins, S.; Janiszewski, J.; Kelly, K. G.; Meyer, R. D.; West, M. J. *Biomol. Screen.* **2002**, *7*, 373–382.
- Kopman, A. F.; Klewicka, M. M.; Neuman, G. G. *Anesth. Analg. (Baltimore)* **2000**, *90*, 1191–1197.
- Erratum *Anesth. Analg.* **2000**, *91*, 67.

43. Yoo, S.-E.; Cha, O. J. *Bulletin Korean Chem. Soc.* **1995**, *16*, 110–112.
44. Omega; 2.1 ed.; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2006.
45. MacroModel; 9.1 ed.; Schrodinger, Inc.: New York, NY, 2006.
46. Cruciani, G.; Pastor, M.; Guba, W. *Eur. J. Pharm. Sci.* **2000**, *11*, S29–S39.
47. Cruciani, G.; Crivori, P.; Carrupt, P. A.; Testa, B. *Theochemistry* **2000**, *503*, 17–30.
48. Crivori, P.; Cruciani, G.; Carrupt, P.-A.; Testa, B. *J. Med. Chem.* **2000**, *43*, 2204–2216.
49. Volsurf; 4.1.4 ed.; Molecular Discovery, Ltd., 2004.
50. Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1998**, *63*, 4305–4313.
51. ACD/LogD; 4.76 ed.; Advanced Chemistry Development, Inc.: Toronto, ON, Canada, 2001.
52. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623–1641.
53. QuACPAC; 1.1 ed.; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2004.
54. Sutter, J.M., personal communication.
55. Massart, D. L.; Kaufman, L.; Rousseeuw, P. J.; Leroy, A. *Anal. Chim. Acta* **1986**, *187*, 171–179.
56. Cruz Ortiz, M.; Sarabia, L. A.; Herrero, A. *Talanta* **2006**, *70*, 499–512.
57. Igumenova, T. I.; Lee, A. L.; Wand, A. J. *Biochemistry* **2005**, *44*, 12627–12639.
58. Johnson, S. R.; Jurs, P. C. *Computer-Assisted Lead Finding and Optimization: Current Tools for Medicinal Chemistry*, [European Symposium on Quantitative Structure–Activity Relationships], 11th, Lausanne, Sept. 1–6, 1996 **1997**, 31–48.
59. Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2003**, *107*, 8377–8379.
60. Ringer, A. L.; Sinnokrot, M. O.; Lively, R. P.; Sherrill, C. D. *Chem.—A Eur. J.* **2006**, *12*, 3821–3828.
61. Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Am. Chem. Soc.* **2005**, *127*, 8594–8595.
62. Lee, E. C.; Hong, B. H.; Lee, J. Y.; Kim, J. C.; Kim, D.; Kim, Y.; Tarakeshwar, P.; Kim, K. S. *J. Am. Chem. Soc.* **2005**, *127*, 4530–4537.
63. Perry, M.; Stansfeld, P. J.; Leaney, J.; Wood, C.; de Groot, M. J.; Leishman, D.; Sutcliffe, M. J.; Mitcheson, J. S. *Mol. Pharmacol.* **2006**, *69*, 509–519.
64. Roux, B.; MacKinnon, R. *Science (Washington, DC)* **1999**, *285*, 100–102.
65. Abraham, M. H. *Chem. Soc. Rev.* **1993**, *22*, 73–83.
66. FingerprintToolkit; 4.62 ed.; Daylight Chemical Information Systems, Inc.: Santa Fe, NM.
67. Stouch, T. R.; Kenyon, J. R.; Johnson, S. R.; Chen, X.-Q.; Doweyko, A.; Li, Y. *J. Comput. Aided Mol. Des.* **2003**, *17*, 83–92.