Evolving hERG Inhibition Model

Adam C. Lee, Grazyna Fraczkiewicz, Robert Fraczkiewicz, Robert Clark, and Walter S. Woltosz

[1]Simulations Plus, Inc., 42505 10th Street West, Lancaster, CA 93534, USA (www.simulations-plus.com)

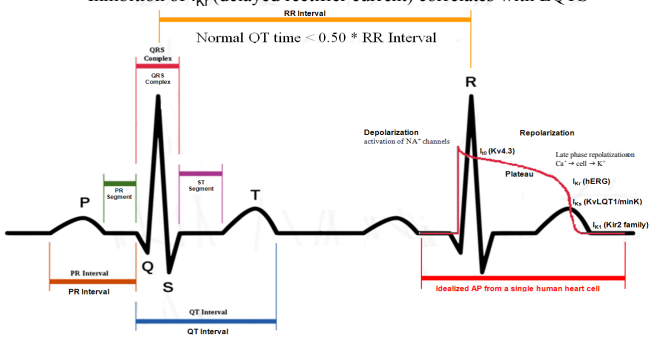
Introduction

Modeling hERG inhibition has significantly gained popularity since 2005, when the FDA recognized the correlation between hERG inhibition and a prolonged QT interval by issuing guidance for the evaluation of new non-antiarrythmic drugs against the hERG channel. Long QT syndrome or LQTS is a risk factor for ventricular tachyarrhythmias and sudden death.

Here we present the evolution of our hERG inhibition model in consecutive releases of ADMET Predictor™. Examples detailing the impact of new and evolving descriptors on the "TOX_hERG" model's applicability domain and performance on internal and external data are provided. Focus is given to a particularly interesting case where an earlier release of ADMET Predictor outperformed its successor on a client's proprietary data. Finally, we discuss how we are improving model selection criteria through the use of descriptor sensitivity analysis with artificial neural network ensembles in combination with a better understanding of the model's applicability domain, based on the World Drug Index.

ECG and Cardiac Action Potential

Inhibition of I_{Kr} (delayed rectifier current) correlates with LQTS

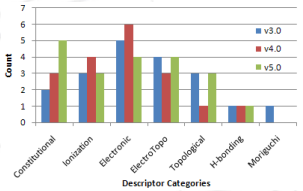


Modified from Pearlstein, R., et al. *J. Med. Chem.*, **2003**, *46*, 2017-2022.

Comparison of Descriptors and Model Architectures v3.0–5.0

Although different descriptor selection methods were used when building the models with ADMET Predictor v3.0 (genetic algorithm) and v4.0-5.0 (input gradient), five descriptors were common to each model and the first four ranked among the top 8 descriptors according to our descriptor sensitivity analysis.

1. (Topological) Distance between the center of mass and the most distant atom
2. (Ionization) Fraction of zwitterionic species in ionization states with no net charge
3. (Ionization) Absolute value of the average across all ionized species of the net formal negative charge
4. (Constitutional) Number of amide groups
5. (Constitutional) Number of distinct π-systems, excluding lone pairs



Model	Neurons	Common Descriptors		
		v3.0	v4.0	v5.0
v3.0	3	19	7	6
v4.0	2	7	18	11
v5.0	2	6	11	19

Network architectures are defined by the number of neurons present and the number of inputs (descriptors) used. # of weights = neurons*(inputs+2) +1

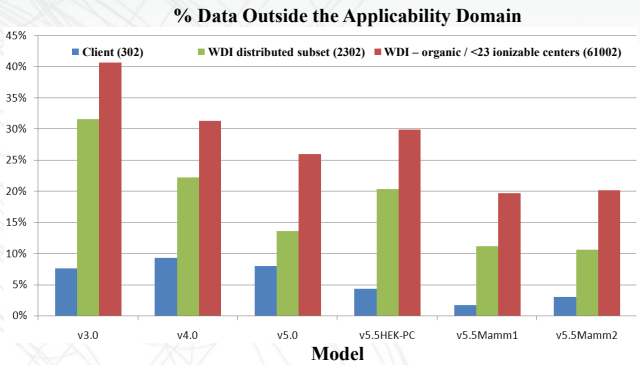
Training Set Effects

The data used to construct our earlier hERG inhibition models (v3.0-5.0) were restricted to values obtained by patch clamp studies on mammalian cell lines composed primarily of human embryonic kidney (HEK) cell lines. We hypothesized that expanding that base would improve the model performance on external data (HEK293).

That hypothesis was tested by building new models for version 5.5 using data sets consisting of measurements for patch clamp assays alone or in combination with values obtained by displacement assay (e.g., of [3H]-dofetilide), which have been shown to correlate well with patch clamp results (Murphy, SM. *J. Pharm. Tox. Meth.* **2006**, *54*, 42-55).

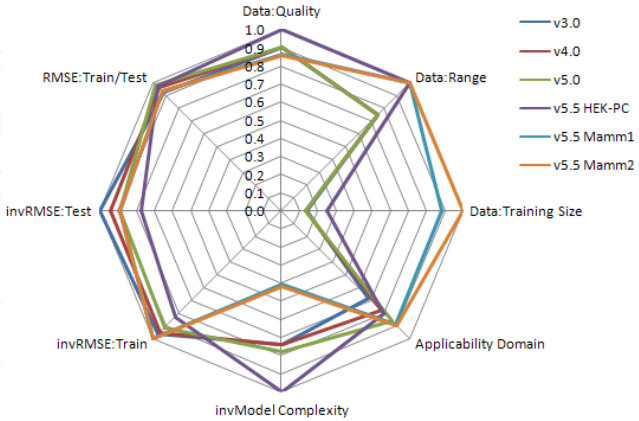
The figures and table below show results for models based on earlier versions of ADMET Predictor as well as for three different combinations of data types, denoted "HEK Patch Clamp," "Mammalian 1" and "Mammalian 2": HEK Patch Clamp ⊂ Mammalian 1 ⊂ Mammalian 2.

TOX_hERG Applicability Domain and WDI



Model Selection & Conclusions

Model selection occurs using a combination the following normalized elements:

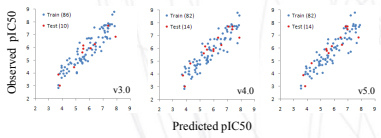


Take home messages:

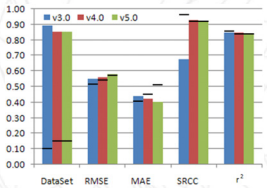
1. Model selection criteria play a vital part when selecting a useful model
2. Two criteria that often go overlooked:
   a. balanced statistics between the training/test sets
   b. understanding of the overall applicability domain
3. When selecting models, especially those built from a small training set, it pays dividends to focus on and maximize the applicability domain.
4. Manual patch clamp measurements produce the highest quality hERG data, but useful insights were gained by including data from displacement assays.

Model Performance for ADMET Predictor v. 3.0–5.5

Internal Data Set: 96 Patch clamp measurements (IC50) on mammalian cells transfected with hERG

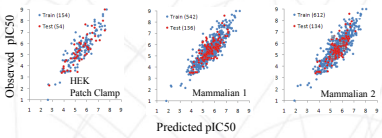


The training and test sets of v4.0 and v5.0 are equivalent. Nine of the ten molecules making up the test set in v3.0 are common to v4.0 and v5.0.



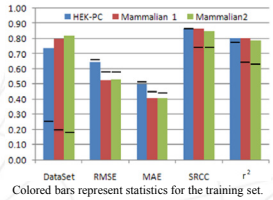
Colored bars represent statistics for the training set.
Black lines represent statistics for the test set.

New Data: 356 patch clamp (mammalian) and 410 displacement (HEK) cell lines transfected with hERG



The HEK Patch Clamp test set was selected using every nth compound, while k-means clustering was used to select the test sets for the mammalian models.

HEK Patch Clamp ⊂ Mammalian 1 ⊂ Mammalian 2



Colored bars represent statistics for the training set.
Black lines represent statistics for the test set.

Client's External Data Set: Patch Clamp measurements on HEK293 cells transfected with hERG

Performance on External Data	RMSE on 191 (dose response)	Category 1μM cutoff		Predicted Category Non	Predicted Category Tox	Sensitivity	Specificity	Precision Toxic	Precision nonToxic
v3.0	0.82	Observed	Tox	14	154	0.917	0.688	0.782	0.872
			Non	95	43				
v4.0	0.92		Tox	23	145	0.863	0.688	0.771	0.805
			Non	95	43				
v5.0	0.95		Tox	42	126	0.750	0.746	0.783	0.710
			Non	103	35				

Performance on External Data	RMSEon 191 (dose response)	Category 1μM cutoff		Predicted Category Non	Predicted Category Tox	Sensitivity	Specificity	Precision Toxic	Precision nonToxic
V5.5 HEK/Patch Clamp	0.78	Observed	Tox	20	148	0.881	0.580	0.718	0.800
			Non	80	58				
V5.5 Mammalian 1	0.69		Tox	26	142	0.845	0.833	0.861	0.816
			Non	115	23				
V5.5 Mammalian 2	0.67		Tox	32	136	0.81	0.849	0.866	0.785
			Non	117	21				

ACKNOWLEDGEMENTS

Our collaboration is ongoing and we look forward to enhancing our understanding of and building a more useful hERG model. Special thanks to Jinhua Zhang of Simulations Plus for providing the WDI subsets.

simulationsplus, inc.