Prediction of I_{Kr} Blocker Channel State Preference Based on Voltage Clamp Simulations Using Machine Learning Techniques

Fernando Escobar[1], Julio Gomis-Tena[1], Javier Saiz[1], Lucía Romero[1]

[1]Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, Valencia, Spain

Abstract

Assessment of drug cardiotoxicity is crucial in the development of new compounds and is typically addressed by evaluating the blockade they cause in the potassium human ether-à-go-go related gene (hERG) channels. Our objective is to develop a classifier to determine the preference for binding to the different states of a drug.

We created a set of 2600 virtual blockers with different affinities and kinetics to the conformational states of the channel divided into 13 classes. Simulations were carried out using three stimulation protocols that enhance the probabilities of the channel to occupy a certain state. Three measurements were taken for each of the simulations: IC_{50}, the recovery constant of the I_{Kr} potassium current and an estimation of the time required for the simulation to be stable. Therefore, we obtained 9 variables for each of the blockers studied. A two-step classifier was developed, trained and evaluated. First, we used support vector machines on the IC_{50} to separate the 13 classes into three groups with 4, 5 and 4 classes respectively. Secondly, we used neural networks on each group with all the variables to finally classify the blockers.

The three classifiers obtained an overall accuracy on the test group of 90.83, 88.66 and 89.16% for each of the groups respectively.

1. Introduction

Preclinical assessment of cardiac toxicity of new drugs has become a priority for pharmaceutical companies due to potentially life-threatening side effects involving ventricular arrythmias [1]. Nowadays, the most commonly used tests to assess cardiac safety are based on the in vitro blockade that a drug causes in the human ether-à-go-go-related gene (hERG) and the in vivo prolongation of the QT interval. Both these phenomena have been linked to the appearance of Torsades de Pointes (TdP).

While these safety tests have proved successful in preventing harmful drugs from reaching the market, they have also stopped the development of potentially useful drugs. In fact, there are drugs such as verapamil, a well know hERG blocker, that do not lead to the development of TdP despite of being potent I_{Kr} blockers.

Recent studies have shown that assessment of cardiac safety improves when drug dynamics and kinetics are taken into account [2]. State dependent drug binding can significantly alter the IC_{50} values, which is the drug concentration at which the value of the ionic current is halved, depending on the voltage clamp protocols used [3]. Thus, consideration of drug behaviour and standardization of the voltage clamp protocols are important aspects for the improvement of the assessment of cardiac safety. A recent study proposed a set of three protocols enhancing the probability of the channel to occupy a certain state [4]. In this work, we will use these stimulation protocols with the aim of elucidating the channel state preference of a drug by using machine learning techniques.

2. Material and Methods

2.1. Models

Six variants of human ventricle the Fink et al. [5] I_{Kr} Markov model were used in order to simulate drug-channel interactions, as shown in Figure 1. The original drug free model has five conformational states, three closed (C_1, C_2 and C_3), open (O) and inactivated (I). The five drug bound states (C_{1d}, C_{2d}, C_{3d}, O_d and I_d) were incorporated to simulate the states of the channel when the drug is bound. Transition from the drug free state to the drug bound one is regulated by a constant named $k_{(C,\ O,\ I)}$ which is multiplied by the drug concentration (D) to obtain the transition rate. Drug unbinding from a certain state is regulated by a constant $r_{(C,\ O,\ I)}$.

2.2. Protocols

Three previously developed voltage clamp protocols were used to simulate the effects of the drugs [4]. These protocols, which were called P40, P0 and P-80, maximized the time the channel are at 40, 0 and -80 mV, respectively. P40 consist of a 5 s conditioning step at 40 mV, followed by 0.2 s test pulse at -60 mV and a 0.2 s recovery time at -80 mV resting potential. P0 is equal to P40 but the conditioning step is applied at 0 mV. P-80 consist of a 0.5

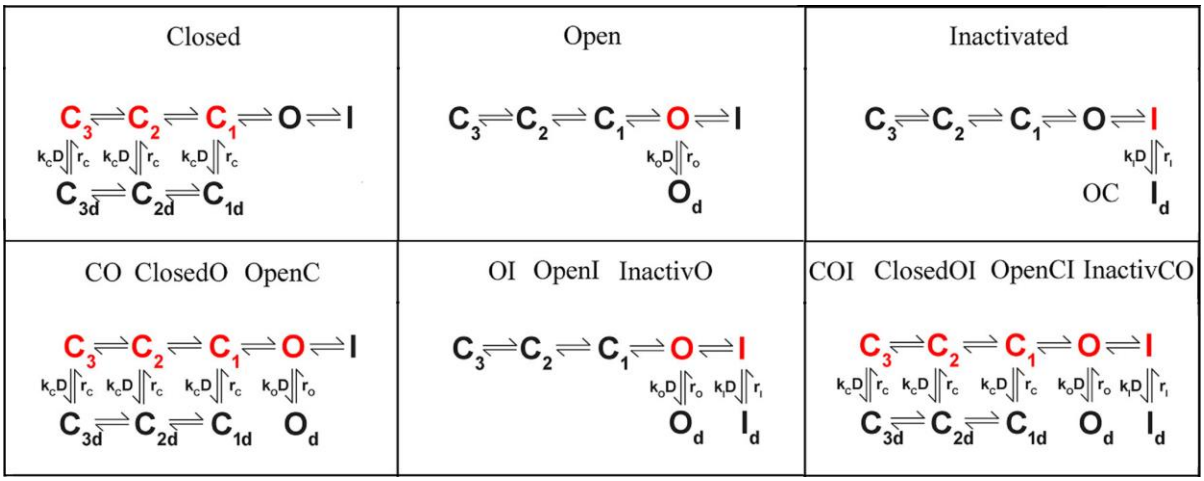

Figure 1: Markov models with drug free (C_3, C_2, C_1, O and I) and drug bound (C_{3d}, C_{2d}, C_{1d}, O and I) states. D is the drug concentration, k_C, ko and k_I are multiplied by D to calculate the association rate constants in the closed, open and inactivated state**s** respectively, and r_C, r_o and r_I are the dissociation rate constants. Red letters indicate binding states. The name of the classes appear with the corresponding the Markovian model.

s test step at 20 mV, followed by a 0.2 s step at -50 mV and a 4.5 s conditioning step at -80 mV. The protocols enhance the probability of the channel to occupy the inactivated, open and close state respectively. Protocols shown in figure 2. [6]

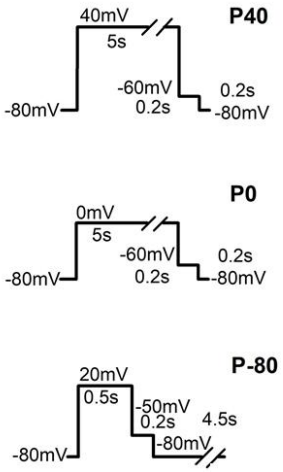

Figure 2: Voltage clamp protocols used to simulate drug effects, namely P-80, P0 and P40.

2.3. Drugs

A set of 2600 virtual drugs with different dynamics and kinetics for the conformational states of the channel was created. Considering all affinities, there are thirteen possible classes of drug. The names of these classes appear with their corresponding Markov model in Figure 1 and are defined as follows. Drugs that bind exclusively the closed state (Closed), the open state (Open) and the inactivated state (Inactivated). Drugs that bind the open and close states simultaneously with equal affinities (CO), preference for the closed state (ClosedO) or the open state (OpenC). Drugs which binds the open and inactivated states without preference (OI), with higher affinity for the open state (OpenI) or the inactivated state (InactivO). Finally, drugs that bind simultaneously to all three states, equally (COI), with closed state preference (ClosedOI), open state preference (OpenCI) and inactivated state preference (InactivCO). In all cases, the association rate constant is equal for all blocked states. Higher affinity to a certain state is simulated by a 100-fold reduction in the dissociation rate constant of the corresponding state. 200 drugs of each class were simulated, randomly generating the transition rates values using Matlab.

3. Results

The effects of the virtual drugs were simulated using the three aforementioned voltage clamp protocols. For each simulation we calculated three variables, 1) IC_{50} for each of the protocols, the most commonly used parameter to assess cardiac safety, 2) a derivative based estimation of the number of pulses required to reach steady state and 3) the time constant of the evolution of the inhibition of the current at the IC_{50} concentration, which was estimated using an exponential fitting. Therefore, a total of nine variables per drug were calculated.

After all variables were obtained, a two-step classifier was developed to try to classify drugs intro the thirteen target classes (shown in Figure 1). All data analysis was conducted using Matlab (Matworks Inc.)

3.1. Support Vector Machine Classifier

An exploratory analysis of the data was conducted to visualize all variables. 3D plots were constructed by plotting the values of the variable obtained with each protocol in each axis. A 3D representation of the IC_{50} values is shown in Figure 3 (plots for the other two variables are not shown).

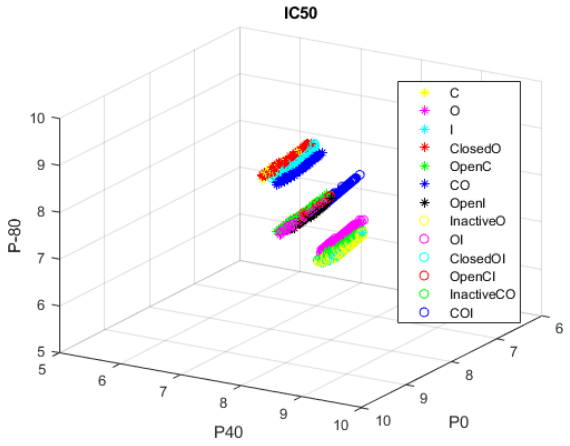

Figure 3: 3D representation of the IC_{50} values of the 2600 virtual blockers. Axis represent the –log([IC_{50}]) obtained with the three voltage clamp protocols (P40, P0 and P-80).

As seen in Figure 3, there are three clearly distinguishable groups among the drugs, which can be separated by drawing a hyperplane between them. For this reason, we decided to use support vector machines (SVM) to create the frontier between groups. Thus, two SVM were trained using a one-vs-all approach. The first SVM separated the top group, considering all classes present in this group as one, from the rest of drugs (Figure 4, top panel). The second SVM did so with the bottom group, again considering all classes in that group as one (Figure 4, bottom panel). All drugs not fitting to any of the two groups were assigned to the middle one.

This first step of the classifier was able to fully separate drugs into three groups based only on IC_{50} values. The first group (top, blue coloured in the top panel of Figure 4) corresponded with drugs that preferably binded to the closed state and drugs that equally binded the closed and open state (Closed, ClosedO, ClosedOI, CO). The second group (middle) contained drugs with higher affinity for the open state and those that equally blocked all three states (Open, OpenC, OpenI, OpenCI, COI). Finally, the bottom group (blue coloured in the bottom panel of Figure 4) corresponds to inactivated state preference drugs and drugs simultaneously binding the open and inactivated states with equal affinity (Inactivated, InactivO, InactivCO, OI).

Thus, using only the IC_{50} values obtained with the three protocols we were able to separate drugs based on their preferred binding state.

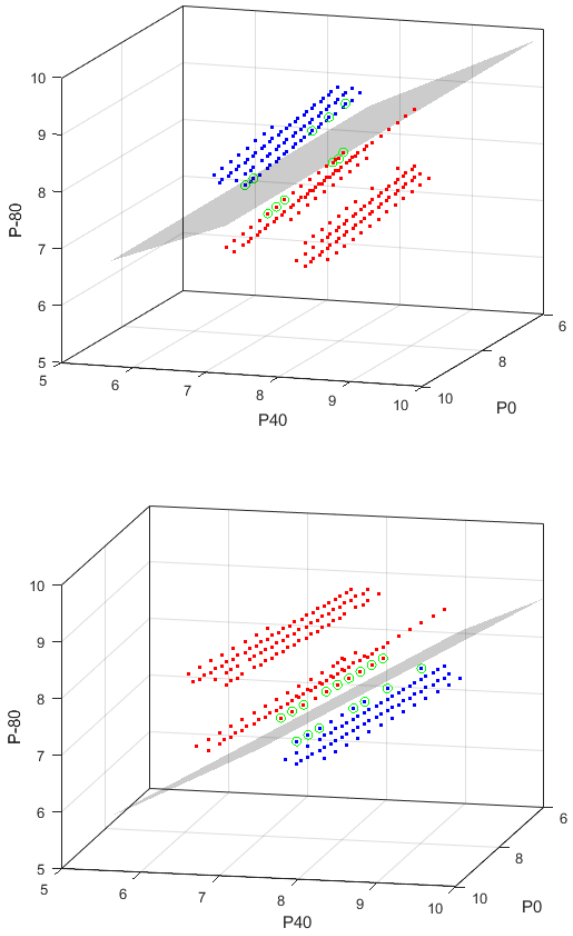

Figure 4: Hyperplanes created by the SVM to separate the drugs into three groups based on the IC_{50} values obtained with the protocols P-80, P0 y P40. (A) SVM used to separate the top (Closed state preference) group. (B) SVM used to separate the bottom (Inactivated state preference). Blue dots represent drugs belonging to the separated group and red dots correspond to the ones that do not.

3.2. Neural Networks Classifier

After the first step of the classification procedure, a set of three neural networks was trained using Matlab toolbox. These networks were used to classify the remaining drugs in each of the previously separated groups into the final target classes. Training and test groups were automatically selected by the toolbox. 120, 126 and 120 drugs were selected for the test groups, respectively. All nine available variables were considered in this step. A range of 1-18 (double the number of neurons in the input layer) neurons in the hidden layer was evaluated, keeping the combination providing best results on the test group [7]. The best results were obtained with 11, 7 and 9 hidden neurons for the groups of closed, open and inactivated state preference, respectively. Results for the test group are shown in Figure

5. This figure shows an overall accuracy of 90.83, 88.66 and 89.16% for the test groups.

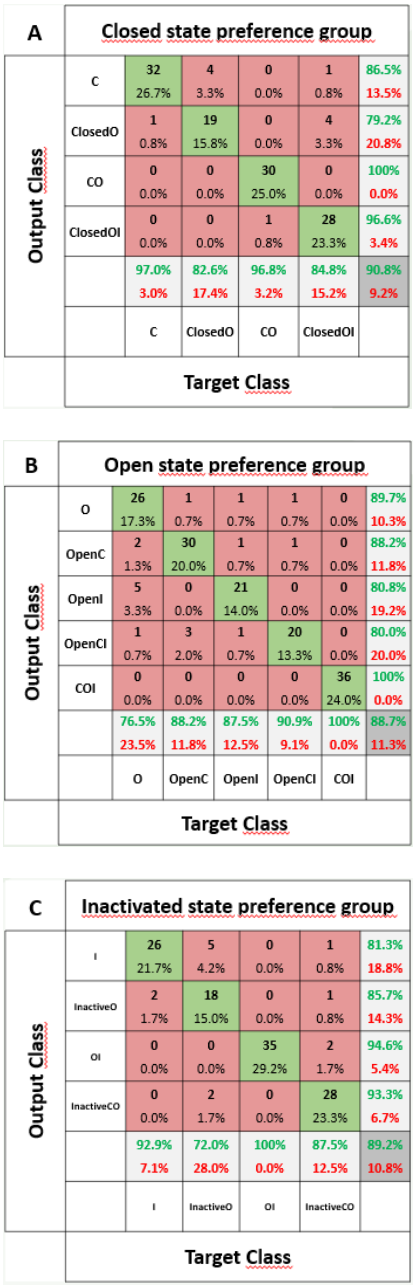

Figure 5: Confusion matrices of the test groups for the neural networks. The top number in the cells represents the absolute number of drugs and the bottom number the percentage. (A) Closed state preference group. (B) Open state preference group. (C) Inactivated state preference group. Rows correspond to the predicted class and columns to the real class. Green cells show correctly classified drugs and red cells missclassified ones. The far right column summarizes the positive predictive value (green) and the false discovery rate (red) and the bottom row shows the true positive rate (green) and the false negative rate (red). Dark grey colored cells show the overall accuracy.

4. Conclusion

We have developed a tool for classifying drugs according to their affinities based on measurements obtained from voltage clamp protocols with high accuracy.

5. Future Work

We intend to further increase the accuracy of our neural networks. We will also use newly generated virtual drugs sets, containing drugs with different kinetics and dynamics. An additional step will be the validation of our in-silico results with experimental data using the same voltage clamp protocols.

6. References

[1] Noble, D. Computational Models of the Heart and Their Use in Assessing the Actions of Drugs. J. Pharmacol. Sci. (Tokyo, Jpn.). 2008; 107(2): 107−117.

[2] Li Z, Dutta S, Sheng J, Tran P, Wu W, Chang K et al. Improving the In Silico Assessment of Proarrhythmia Risk by Combining hERG (Human Ether-à-go-go-Related Gene) Channel–Drug Binding Kinetics and Multichannel Pharmacology. Circulation: Arrhythmia and Electrophysiology. 2017; 10(2).

[3] Lee, W.; Windley, M. J.; Perry, M. D.; Vandenberg, J. I.; Hill, A. P. Protocol-Dependent Differences in IC50 Values Measured in Human Ether-Á-Go-Go-Related Gene Assays Occur in a Predictable Way and Can Be Used to Quantify State Preference of Drug Binding. Mol. Pharmacol. 2019; 95: 537−550.

[4] Gomis-Tena J. When Does the IC50 Accurately Assess the Blocking Potency of a Drug?. Journal of Chemical Information and Modeling. 2020; 60(3): 1779-1790.

[5] Fink M, Noble D, Virag L, Varro A, Giles W. Contributions of HERG K+ current to repolarization of the human ventricular action potential. Progress in Biophysics and Molecular Biology. 2008; 96(1-3): 357-376.

[6] Milnes, J. T.; Witchel, H. J.; Leaney, J. L.; Leishman, D. J.; Hancox, J. C. Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37 °C: Cisapride versus dofetilide. J. Pharmacol. Toxicol. Methods. 2010; 61: 178−191.

[7] Heaton J. The Number of Hidden Layers [Internet]. 2017 [cited 26 August 2020]. Available from: https://www.heatonresearch.com/2017/06/01/hidden-layers.html

Address for Correspondence:

Fernando Escobar Ropero

fescobar@ci2b.upv.es