ARTICLE

A Multiscale Simulation System for the Prediction of Drug-Induced Cardiotoxicity

Cristian Obiol-Pardo,[†] Julio Gomis-Tena,[‡] Ferran Sanz,[†] Javier Saiz,[‡] and Manuel Pastor*,[†]

[†]Research Programme on Biomedical Informatics (GRIB), IMIM, Universitat Pompeu Fabra, PRBB, Dr. Aiguader 88, E-08003 Barcelona, Spain

[‡]Grupo Bioelectronica I3BH, Universitat Politecnica de Valencia, Camino de Vera s/n, E-46022 Valencia, Spain

Ⓢ *Supporting Information*

ABSTRACT: The preclinical assessment of drug-induced ventricular arrhythmia, a major concern for regulators, is typically based on experimental or computational models focused on the potassium channel hERG (human ether-a-go-go-related gene, $K_v11.1$). Even if the role of this ion channel in the ventricular repolarization is of critical importance, the complexity of the events involved make the cardiac safety assessment based only on hERG has a high risk of producing either false positive or negative results. We introduce a multiscale simulation system aiming to produce a better cardiotoxicity assessment. At the molecular scale, the proposed system uses a combination of docking simulations on two potassium channels, hERG and KCNQ1, plus three-dimensional quantitative structure—activity relationship modeling for predicting how the tested compound will block the potassium currents IK_r and IK_s. The obtained results have been introduced in electrophysiological models of the cardiomyocytes and the ventricular tissue, allowing the direct prediction of the drug effects on electrocardiogram simulations. The usefulness of the whole method is illustrated by predicting the cardiotoxic effect of several compounds, including some examples in which classic hERG-based models produce false positive or negative results, yielding correct predictions for all of them. These results can be considered a proof of concept, suggesting that multiscale prediction systems can be suitable for being used for preliminary screening in lead discovery, before the compound is physically available, or in early preclinical development when they can be fed with experimentally obtained data.

■ INTRODUCTION

The occurrence of several cases of serious, potentially life-threatening, ventricular arrhythmia episodes after the administration of noncardiovascular agents in the late 1990s raised a remarkable and understandable concern in regulators.[1] Even though the mechanisms of such drug-induced arrhythmia remain unclear, there is a wide agreement that it is frequently associated with prolongation of the QT interval. The QT interval on the surface electrocardiogram (ECG) is the time interval between the onset of the QRS complex and the end of the T wave and is a measure of the duration of the ventricular depolarization and repolarization cycle. A delayed ventricular repolarization is associated to an increased risk of ventricular tachiarrhythmia[2] and nowadays the QT interval prolongation has become a de facto surrogate biomarker of the arrhythmogenic potential. In practice, the detection of QT prolongation can justify halting a candidate compound development or withdrawing a drug from the market. Strategies for assessing risk for QT interval prolongation typically involve both in vivo and in vitro assays, as it is reflected in preclinical guidelines like ICH S7B. In vivo assays are usually based on electrophysiology studies carried out in dog, monkey, or other nonrodent species in which the QT interval data are obtained after the administration of the tested compound and compared with control and baseline data in small samples. The most used in vitro assays are electrophysiologic patch clamp inhibition assays of the hERG (human ether-a-go-go-related gene, $K_v11.1$) potassium channel, since it is responsible for the rapid delayed rectifier current (IK_r), one the most important components of repolarization.[3]

In this scenario, there is a pushing need in the pharmaceutical industry for methods that provide an early and reliable identification of potential QT prolongation issues. Combinations of functional hERG inhibition assays and in vivo ECG experiments have been reported to predict the risk of human QT prolongation,[4] although the assays involved are time and resource consuming and not suitable for being applied in medium- or high-throughput scale during the discovery stages. Screening methods based on competitive binding with radiolabeled channel blockers can provide a faster assessment, but their usefulness have been criticized and are not considered to be substitutive of electrophysiology patch clamps assays (ICH S7B). Alternatively, diverse in silico methods have been developed, mainly aimed at the prediction of drug binding affinities for the hERG potassium channel.[5−7]

In general, all the experimental or computational methods mentioned above and based solely on hERG provide a limited picture of the drug effects on the ventricular repolarization, since the final biological outcome depends on the interaction of the drug with several ion channels present at the cardiomyocyte surface and on how these multiple interactions are translated into alterations of the whole heart electrophysiology. Recent studies suggest that considering multiple ions channels provide better cardiotoxicity predictions.[8,9]

For example, several studies[10−14] have reported that compounds able to inhibit contemporaneously the rapid delayed rectifier

current (IK_r), produced by hERG, and the slow delayed rectifier current (IK_s), produced by the KCNQ1 potassium channel ($K_v7.1$), are more likely to trigger arrhythmogenic processes by decreasing the so-called repolarization reserve. Therefore, cardiotoxicity prediction systems based solely on drug—hERG interaction can predict safe compounds as potentially arrhythmogenic (false positives) or, even worse, can fail to detect toxic compounds (false negatives) as, for example, the recently reported sulphonamide JNJ303 that caused sudden death in animal models.[15]

For all these reasons we consider that any prediction system aiming to produce a reliable assessment for the liability of a drug to prolong the QT interval should not be limited to hERG and must reflect to a higher degree the complexity of the phenomena involved. In the present work we introduce an innovative system for the in silico prediction of the QT interval prolongation that integrates predictions done with two major repolarization currents and that couples simulations at different scales. At the molecular scale, the system applies ion channel docking simulations and three-dimensional quantitative structure—activity relationship (3D-QSAR) modeling to predict the binding affinity of the drug for two of the main actors in the phase III of the repolarization, the hERG and KCNQ1 potassium channels. Then, the results of this prediction have been integrated into a computational model of the electrophysiology of the guinea pig cardiomyocyte using the Hodgkin—Huxley formalism[16] for simulating the drug effects on the cardiac cell. Going one step further, the so-modeled cells were used to build a simplified model of the heart tissue and simulate the propagation of the cardiac beat, obtaining a simulated ECG wave representation, from which parameters like changes in QT interval and transmural dispersion of repolarization (TDR)[17] can be measured.

Some of the concepts applied to the present system have been previously reported. Thus, Silva et al.[18] reported the use of a multiscale approach, which coupled molecular dynamics simulations with electrophysiological models for the prediction of changes in the action potential generated by mutations in the KCNQ1 channel. Moreover, Bottino et al.[19] proposed the integration of experimentally obtained data on ion channel inhibition into a simple pore block model, and Suzuki et al.[17] proposed the joint consideration of the IK_s and IK_r currents, represented in 2D plots, as major biomarkers of drug cardiotoxicity. However, in the present work we are presenting for the first time a computational framework which integrates in silico predictions of drug binding affinities for two ion channels and electrophysiological models for generating outputs which are relevant for the assessment of the arrhythmogenic risk of test compounds. The advantages of this method over hERG-only based systems are illustrated by predicting the effect of several drugs with known QT interval prolongation effects, some of which are mispredicted by standard methodologies. In addition, the method has been applied to produce 2D maps representing the predicted ECG effects of hypothetical drugs with any given combination of inhibitory potencies for hERG and KCNQ1, at several concentrations. Although the model presented here is only a first prototype, the results obtained so far illustrate the advantages of multiscale prediction systems and may encourage further developments in this direction.

■ METHODS

Ion Channel Docking Simulations. Initial 3D coordinates for the drugs were obtained from their 2D structures using the CORINA program.[20] Their protonation state at physiological pH was assigned using MoKa.[21,22] The so-obtained structures were submitted to docking simulations into the binding sites of homology models of hERG and KCNQ1 representing the open state of the ion channels, based on the original models described by Farid et al.[23] and Lerche et al.,[24] respectively. The docking simulations were carried out with the GOLD software[25] using default parameters. The top-scored poses, as evaluated with the ASP statistical function,[26] were considered to be representative of the bioactive drug conformation and were used for the 3D-QSAR modeling.

3D-QSAR Modeling. We used the 3D-QSAR methodology for predicting the IC_{50} values of the studied drugs for hERG and KCNQ1. The first step was to compile from public sources a database of compounds with known inhibitory properties against hERG and KCNQ1. For each compound we collected the 2D structure and the experimental functional inhibition measured in patch clamp experiments, expressed as IC_{50}. The final databases contained 355 hERG blocker compounds, an additional database of 45 hERG blockers compounds to be used in external validation, and 162 KCNQ1 blocker compounds, including 13 compounds used for external validation (Supporting Information, Tables S1 to S3).

It is worth mentioning that the quality of this data, compiled from public and bibliographic sources, is limited by the disparity of the experimental conditions used for its determination, a fact that hampers the quality of the resulting 3D-QSAR models.

All the compounds included in these databases were submitted to the protocols detailed above, obtaining two structures, one for hERG and another for KCNQ1, which were described using alignment-independent 3D descriptors GRIND-2 as implemented in the Pentacle software (Pentacle v 1.03, Molecular Discovery Ltd.) with default settings. A detailed description of the GRIND-2 is out of the scope of this work, but basically they are equivalent to the GRIND,[27] using the AMANDA[28] algorithm for extracting the most relevant nodes (hot spots) from the molecular interaction fields. The same Pentacle software was used for building projection to latent structures (PLS) models describing the relationship between the GRIND-2 (dependent variables) and the pIC_{50} ($-\log IC_{50}$) for each ion channel (independent variable). The resulting models were stored and used for predicting the IC_{50} of new compounds, not incorporated into the model training set. Further details about the PLS models building, validation, and interpretation are provided as Supporting Information Table S4.

Cellular and Tissue Levels: Ventricular Electrophysiological Models. The IC_{50} values for hERG and KCNQ1 predicted in the previous step were used as input in an electrophysiological model constituted by guinea pig cardiac cells coupled to simulate a ventricular tissue. Electrical activity of isolated ventricular cells was simulated using a modified version of Luo—Rudy dynamic action potential (AP) model.[29] Three different cell types were obtained maintaining IK_r constant and setting the IK_s:IK_r ratio to 23:1 for endocardial cells (endo),[29] 7:1 for midmyocardial cells (mid-type),[29] and 35:1 for epicardial cells (epi).[30] The effect of the drug concentration (D) on the ionic current I_x was represented by introducing the factor $(1 - b)$ as the fraction of unblocked channels, in agreement with the simple pore block model.[19,31] The drug I_x model, $I_x(D)$, is based on the standard sigmoid dose—response curve (eq 1), parametrized using the half-maximal response dose IC_{50} predicted at the molecular level as described above (the

Hill coefficient considered here is 1, which is a reasonable approximation):[19]

$$\frac{I_x(D)}{I_x} = \frac{1}{1 + \dfrac{D}{IC_{50}}} = 1 - b \qquad (1)$$

The action potential duration (APD) was measured at 90% (APD_{90}) and 30% (APD_{30}) of repolarization. Triangulation is defined as the repolarization time from APD_{30} to APD_{90}.

To simulate the electrical propagation in ventricular tissue, we built a model representing a one-dimensional (1D) heterogeneous transmural fiber composed of 600 ventricular transversally coupled cells (100 μm in length and 22 μm of diameter). The model comprises an endocardial region (cells 0−359), a mid-myocardial cells region (cells 360−539), and an epicardial region (cells 540−599), representing a transmural width of 1.3 cm.[32] An effective intercellular resistance R_i of 0.3 Ωcm^2 was considered, and the presence of a resistive barrier between epicardial and mid-type cells[33] was simulated by a five-fold increase in the intracellular resistance at the mid-to-epicardium border region. In this model, the propagation of the action potential (V_m) is described by the nonlinear reaction diffusion equation (eq 2):

$$C_m \frac{\partial V_m(x,t)}{\partial t} + \sum I_{ion} + \frac{a}{2}\frac{\partial}{\partial x}\left(\frac{1}{R_i(x)}\frac{\partial V_m(x,t)}{\partial x}\right) = 0 \quad (2)$$

where C_m is the membrane capacitance, a is the radius of the fiber, and $\sum I_{ion}$ is the sum of all the ionic currents that cross the cellular membrane. The reaction diffusion equation is solved using a central difference scheme in space and the Crank−Nicholson method in time, with a space step of Δx equal to the diameter of the cell. Transmural dispersion of repolarization (TDR) was defined as the difference between the longest and the shortest repolarization time (activation time + APD_{90}) of transmembrane action potentials simulated across the ventricular wall. Then, we generate a pseudo-ECG resulting from the propagation of beats from endocardium to epicardium by calculating the potential at a "virtual electrode" situated 20 mm from the epicardial surface.[30,34] Only cells from 10 to 590 are included in pseudo-ECG computations in order to minimize border effects. The QT interval in the pseudo-ECG was defined as the time between QRS onset and the point at which the end of the T wave crossed the baseline. For the ease of interpretation, values of the APD_{90}, $APD_{90} - APD_{30}$, QT interval duration, and TDR have been represented in 2D color-coded maps confronting the pIC_{50} in hERG (horizontal axis) vs the pIC_{50} in KCNQ1 (vertical axis).

RESULTS

Development of the Multiscale Simulation System. The prediction method that we have developed is represented schematically in Figure 1 and will be briefly described here (further details are provided in the Methods Section). The only input required by the method is the 2D structure of the compound, and the final output of the method is a simulation of the AP and ECG (pseudo-ECG), in which standard parameters, like QT interval prolongation and AP triangulation, can be easily estimated.

Molecular Level. Once the 2D structure of the test compound is entered, the system generates a preliminary 3D structure that is used for carrying out docking simulations into the binding sites of the ion channels hERG and KCNQ1, by using previously published structures (see Methods Section). The aim of this simulation is to obtain a single pose, representative of the compound bioactive conformation (the conformation adopted by the drug at the binding site) that is used in the next step. This structure is then processed and transformed into a variety of 3D, molecular interaction field descriptors (GRIND-2), which are projected into precomputed QSAR models producing two predicted IC_{50} values, one for hERG and another for KCNQ1. These IC_{50} values reflect the ability of the target compound to block ion channels playing a critical role in the ventricular repolarization. Obviously, compounds with very low IC_{50} are likely to produce a QT interval prolongation, but for compounds with intermediate IC_{50} values, its final cardiotoxicity will depend on the balance between both values, something that will be dealt with in the next level.

Electrophysiological Level. The values of IC_{50} obtained at the molecular level are inserted into a first electrophysiology model representing an isolated cardiomyocyte, where they describe the current pore blocking produced by the compound (see Methods Section). In this model the electrical activity was simulated using a modified version of Luo−Rudy dynamic action potential for diverse drug concentrations, from which we can obtain parameters, such as the APD_{90} (the action potential duration measured at 90% repolarization), integrating in a single value the blocking effect computed for both ion channels.

However, not all myocytes in the ventricular tissue are equally sensitive to the IK_r and IK_s current blocking, since they have different hERG/KCNQ1 ratios (see Methods Section). A realistic representation must account for these differences at diverse tissue levels (endocardial, midmyocardial and epicardial) and their implication on propagation of the action potential. This is possible by carrying out a simulation in a second electrophysiological model representing a 1D section of ventricular tissue, populated with a representative proportion of the diverse myocyte types and in which we can simulate the current propagation. In this manner, the simulation allows to estimate a pseudo-ECG from which parameters like the QT interval can be extracted.

Validation and Example Application. The two 3D-QSAR models (see Methods Section) were validated using both leave-one-out cross-validation and external test sets. The quality of the QSAR model obtained for hERG ($r^2 = 0.52$, $q^2 = 0.46$, SDEP = 0.92, SDEP$_{ext}$ = 0.89) is acceptable, and the main GRIND contributions can be interpreted in structural terms, overlapping actual ligand−receptor interactions (Supporting Information, Table S4). In line with these results, the QSAR model for KCNQ1 also shows a lower but still acceptable quality ($r^2 = 0.69$, $q^2 = 0.41$, SDEP = 0.99, SDEP$_{ext}$ = 0.82,) and the model can be also interpreted structurally (Supporting Information, Table S4).

To illustrate the proposed system, we are presenting here the predictions obtained for four compounds (see Table 1), representing extremely diverse behaviors with respect to IK_r and IK_s blocking.[15,35−39] All the compounds show QT interval prolongation in relevant animal models[12,15] except ebastine for which no clear QT effect was observed.[40] Dofetilide is highly selective for hERG, while HMR1556 and JNJ303 are selective for KCNQ1. Ebastine shows a significant hERG blocking but has low KCNQ1 effects. The sulphonamide JNJ303 was selected as a recent case of hERG false negative compound producing a marked prolongation of QT, while ebastine was chosen for being a compound lacking cardiotoxicity, which appears as a false positive in hERG assays.
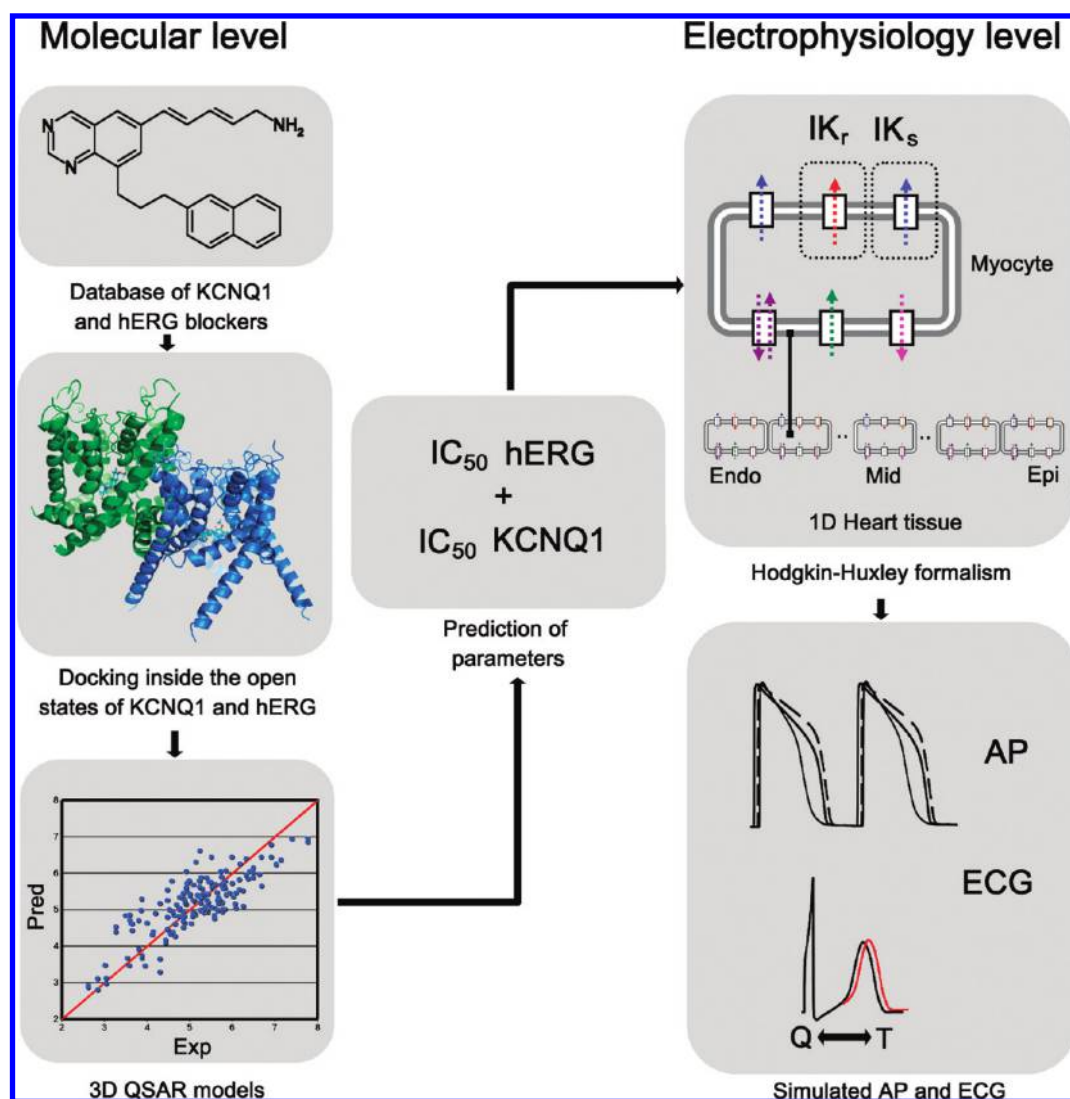
**Figure 1.** Schematic representation of the proposed multiscale simulation system.

Table 1. Predicted and Experimental Data of the hERG and KCNQ1 Blocking Profiles for Dofetilide, HMR1556, JNJ303, and Ebastine[a]

	pIC_{50} hERG		pIC_{50} KCNQ1		QT interval prolongation	pharmacology
	pred	exp	pred	exp		
dofetilide	7.05	7.92 [b]	4.95	<4 [c]	20% (dog)	antiarrhythmic
HMR1556	5.03	<5 [d]	6.68	6.92 [e]	12% (dog)	antiarrhythmic
JNJ303	5.04	4.90 [f]	6.64	7.19 [f]	25% (guinea pig)	antiobesity/diabetes
ebastine	6.54	∼6.52 [g]	4.55	∼5.52 [g]	no	antiallergic

[a] $pIC_{50} = -\log IC_{50}$. [b] Data from HEK cells.[35] [c] Data from GP cells.[36] [d] Data from GP cells.[37] [e] Data from XO cells.[38] [f] Data from mammalian cells.[15] [g] Data from GP cells.[39]

All the compounds were submitted to the multiscale simulation system described earlier. The intermediate results obtained after completing the molecular level (Table 1) show a rather good agreement with experimental values, also because patch clamp experiments are highly dependent on the cell expression system and on experimental conditions of the assay,[41] such as temperature or concentration of potassium ions. Even so, the predicted values clearly identify dofetilide and ebastine as selective hERG

blockers and HMR1556 and JNJ303 as selective KCNQ1 blockers. In the next step, these values were used for simulating the effect of blocking KCNQ1 (IK_s) and hERG (IK_r) in three diverse types of myocytes, representing the relative population of such ion channels in endo, mid-type and epi cells. Our model (see Methods Section) can simulate the effect of the drugs at diverse concentrations (10, 100, and 200 nM) in the AP, representing parameters like APD_{90} or triangulation ($APD_{90} - APD_{30}$), which
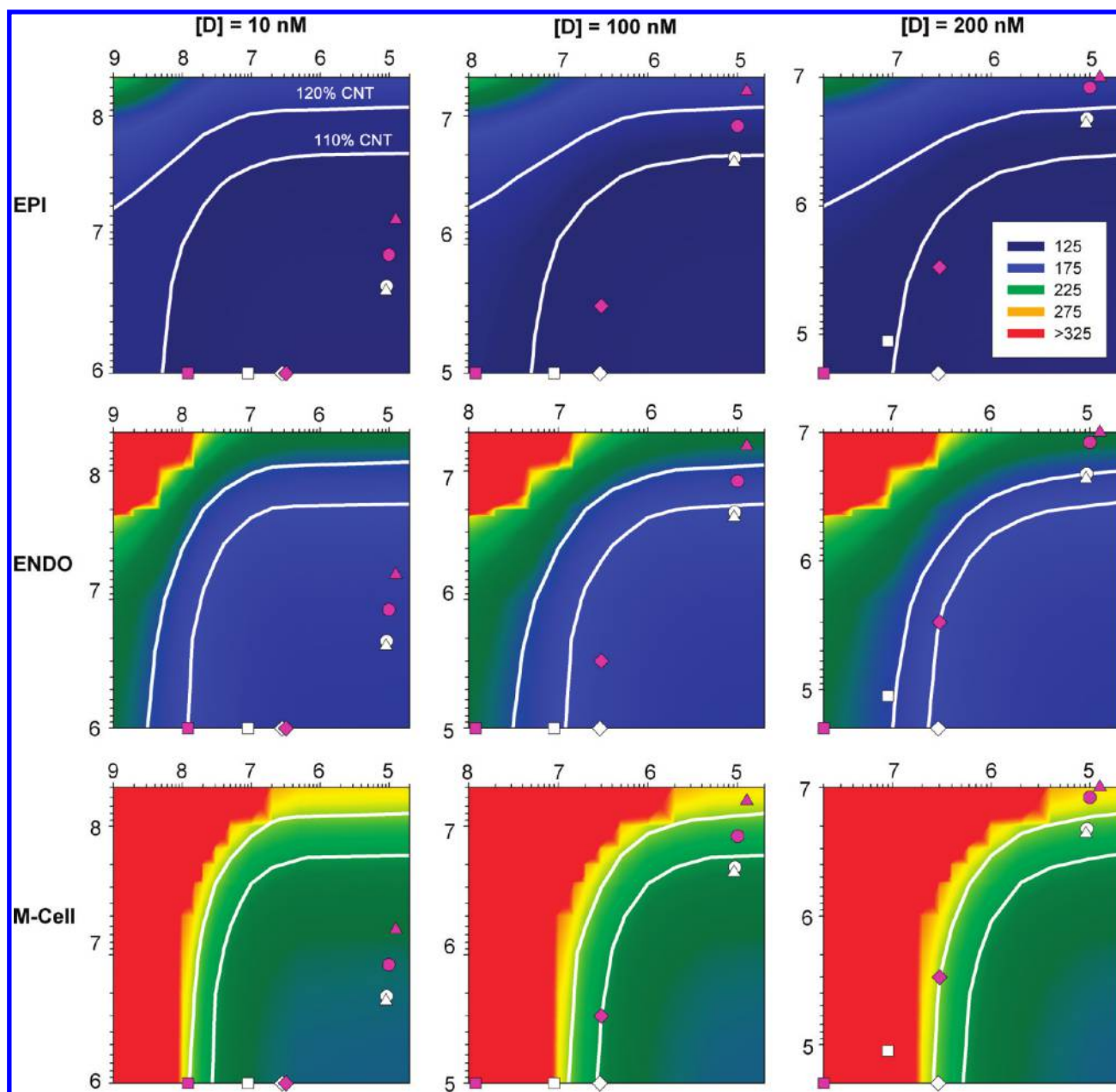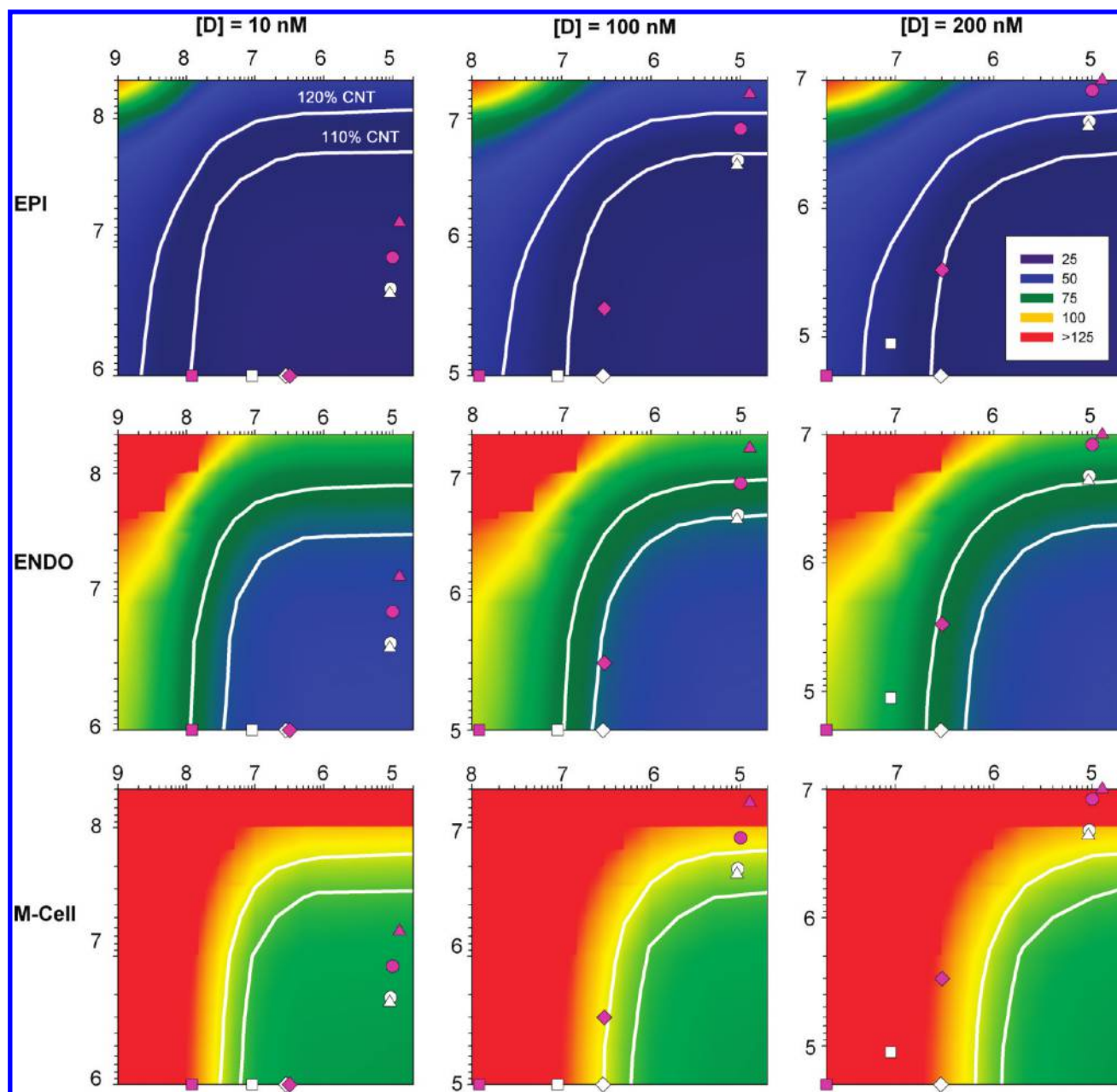
Figure 2. The 2D maps of APD$_{90}$ as a function of pIC$_{50}$ to IK$_r$ (horizontal axis) and to IK$_s$ (vertical axis) at 10, 100, and 200 nM drug concentrations. Upper plots correspond to epi-type cells, middle to endo-type cells, and bottom to mid-type cells. Experimental (purple) and predicted data (white) of dofetilide (square), HMR1556 (circle), JNJ303 (triangle), and ebastine (diamond) are shown. Color legend is expressed in ms.

have been proposed to be highly representative of the arrythmogenic potential.[19] Here, instead of tabulating these parameters for different conditions, we used the model to produce color-coded 2D maps representing the APD$_{90}$ (Figure 2) and triangulation values (Figure 3) obtained for every possible combination of KCNQ1 and hERG pIC$_{50}$ within a wide value range (from 1 nM, pIC$_{50}$ = 9 to more than 10 μM, pIC$_{50}$ = 5). In the same maps we have represented two lines representing an increase of 10% and 20% over the baseline values, and we have located the tested compounds according to both the computed and the experimental pIC$_{50}$ values. In general, when the predicted values for a certain drug at a given concentration are below the safety margin represented by the 10% line, the model predicts no significant toxic effect. On the contrary, if the predicted values are over the

10% change margin, then the system predicts that some cardiotoxicity can be expected, which can be more relevant if the predicted values are beyond the 20% change margin.

Focusing first on the APD$_{90}$ (Figure 2), at 10 nM no compound produced significant cardiotoxicity in any cell type. At a more realistic 100 nM concentration, only ebastine was safe for all cell types, and all the rest were predicted to be cardiotoxic, with the only exception of HMR1556 and JNJ303 in mid-type cells, for which the predicted pIC$_{50}$ values locate them in the safety borderline. At a concentration of 200 nM, all compounds are predicted to be cardiotoxic but ebastine, which is only predicted as cardiotoxic in mid-type cells.

In summary, the results for the four compounds were consistent with the experimental results shown in Table 1 in the
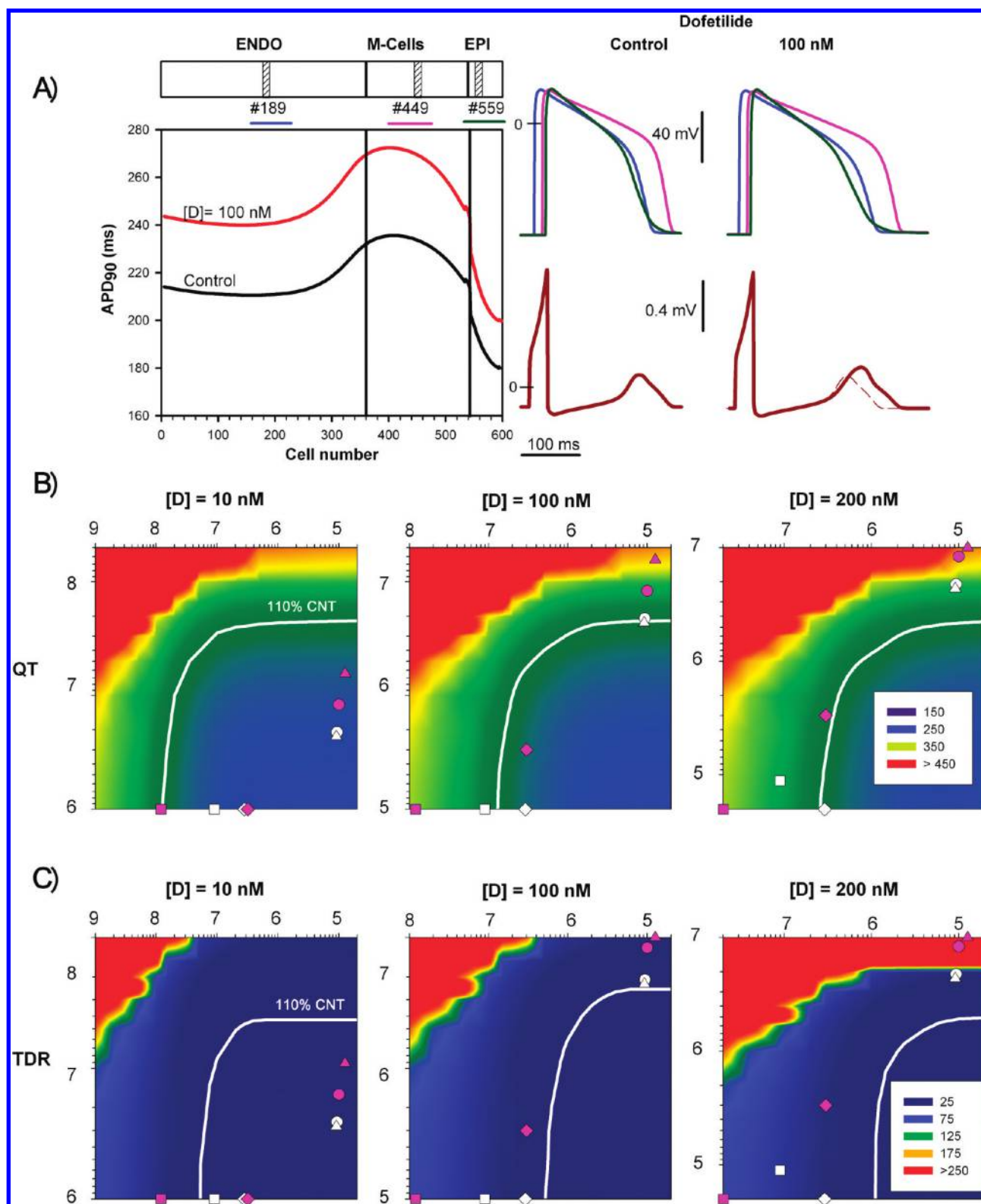
Figure 3. 2D maps of $APD_{90} - APD_{30}$ as a function of pIC_{50} to IK_r (horizontal axis) and IK_s (vertical axis) at 10, 100, and 200 nM drug concentrations. Upper plots correspond to epi-type cells, middle to endo-type cells, and bottom to mid-type cells. Experimental (purple) and predicted data (white) of dofetilide (square), HMR1556 (circle), JNJ303 (triangle), and ebastine (diamond) are shown. Color legend is expressed in ms.

sense that all tested compounds but ebastine show potential cardiotoxicity. Notice that these results are in direct contradiction with predictions obtained using hERG-only based methods, which will not find reasons to predict HMR1556 and JNJ303 as cardiotoxic. Other relevant findings are that mid-type cells are more sensitive to potassium channel blockade (especially to hERG) as well as the importance of the concentration for the predictions. Triangulation of the action potential, calculated as the difference between the action potential duration at 90% and at 30% of repolarization ($APD_{90} - APD_{30}$), has been proposed to be a useful biomarker of cardiotoxicity,[17] and hence we used it for generating the 2D maps shown in Figure 3. The results are remarkably similar to those reported above, even though triangulation in mid-type cells is a very sensitive biomarker, which

predicts changes over 10% for all tested drugs even at 100 nM, including ebastine in mid-type cells.

The information provided by the cell models is highly interesting and reflects the importance of considering more than one ion channel blocking. However, the high differences observed for the diverse cell types studied make the interpretation difficult. The most interesting end point (the drug-induced cardiotoxicity) will be linked to the effect produced in diverse cell types as the simple comparison of the graphics does not provide any hint with respect to how these effects could be integrated. With this aim, we carried out simulations at tissue level, as described in the Methods Section, in which a section of the heart was represented by a row of myocytes containing the three aforementioned cell types in representative proportions. A scheme of the simulation

Figure 4. (A) Schematic representation of the 1D heart tissue simulated together with the obtained AP and ECG when using the predicted IC_{50} of dofetilide. (B) 2D maps of 1D ECG QT prolongation as a function of pIC_{50} to IK_r (horizontal axis) and IK_s (vertical axis) at 10, 100, and 200 nM drug concentrations. (C) 2D maps of 1D ECG TDR as a function of pIC_{50} to IK_r (horizontal axis) and IK_s (vertical axis) at 10, 100, and 200 nM drug concentrations. Experimental (purple) and predicted data (white) of dofetilide (square), HMR1556 (circle), JNJ303 (triangle), and ebastine (diamond) are shown. Color legend is expressed in ms.

system and the results obtained for dofetilide at 100 nM are shown in Figure 4A. In addition, as in the previous case, the results have been represented in color-coded 2D maps (Figure 4B and C), which shows the simulation output parameters (in this case, QT
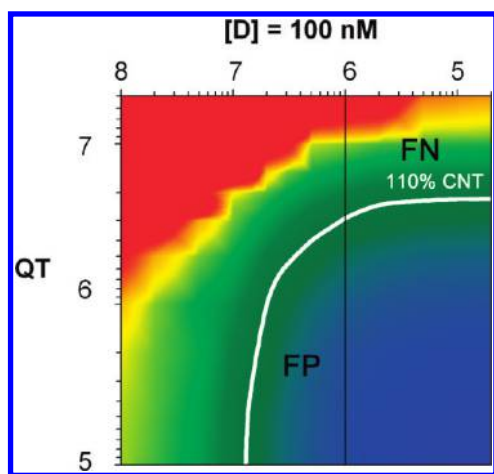
Figure 5. 2D map of QT prolongation as function of pIC_{50} to IK_r (horizontal axis) and IK_s (vertical axis) at 100 nM drug concentration. FN and FP areas when considering hERG blocking alone (vertical line) are shown.

interval and TDR) for any combination of KCNQ1 and hERG IC_{50} values within a wide range. In this case, the graphics are far simpler to interpret because the information has not only been integrated into a single 2D plot for every concentration but also because the parameters represented are closer to the physiological end points. As in the previous simulated models, we have also incorporated lines representing a 10% increase over the control and the predicted effects of the tested set, using either experimental or predicted IC_{50} values.

The analysis of the 2D maps shown in Figure 4B shows even a clearer agreement with the experimental results. At very low concentrations (10 nM) no compound prolongs the regular QT interval, with the exception of dofetilide, which is at the borderline. At the 100 nM concentration, the compounds with high affinity for hERG (dofetilide) or KCNQ1 (HMR1566 and JNJ303) are predicted to be cardiotoxic, and the predictions are less sensitive to the use of experimental or calculated IC_{50} and are not interfered by cell-type considerations. On the other hand, ebastine can be considered as safe, in agreement again with the experimental results. Remarkably, the same results are obtained even for the highest concentration (200 nM), thus making evident the robustness of the predictions. To conclude, we obtained similar maps representing TDR shown in Figure 4C, another biomarker that has been proposed as an alternative to QT interval for the prediction of cardiotoxicity.[17] Interestingly, TDR seems more sensitive than the QT because all the compounds produce TDR alterations over 10% at 100 nM. However, in this case we have no evidence that the 10% alteration represents a suitable safety margin, and probably an alternative cutoff value could be applied. In fact, no clear indications of safety margins concerning TDR have been proposed so far (ICH S7B).

■ DISCUSSION AND CONCLUSIONS

Current state-of-the-art cardiotoxicity prediction methodologies, including experimental ones, show important limitations as a consequence of being focused on hERG, which is probably a too sensitive and incomplete biomarker. Now it is widely accepted that compounds exhibiting high affinity for hERG are not always cardiotoxic or, even worse, compounds with low affinity for hERG could be so. On the other hand, methods based on in vivo

experiments require physical access to significant quantities of the tested compounds, which can be not the case at the initial stages of drug discovery projects. In this context, we propose an alternative in silico methodology that aims to produce more accurate predictions by means of a more complete depiction of the physiological events involved in the drug-induced ventricular arrhythmia. The procedure involves the prediction of the blocking ability of the drugs on two ion channels that are known to play a major role, using either in silico prediction methods or classical experimental determinations. Parameters describing the compound blocking properties are then incorporated into an electrophysiological model which adds two levels of integration: First we compute the drug effects in several types of heart cells with different relative population of both ion channels, and then, the relative importance of the effect on this level is taken into account by carrying out an ECG simulation on a simplified section of the heart tissue in which all cell types are represented in realistic proportions. At the end, the model yields a prediction of the final physiological outcome by means of straightforward parameters (QT interval or TDR).

From a purely formal point of view, our proposed methodology points in the direction which we consider desirable for in silico prediction methods: to provide a sufficiently complete representation of the phenomena modeled. We believe that advances in this field will not arrive by the use of complex or intensive computational methods but are linked to a deeper understanding of the reality we intend to model. The results of our example application confirm the usefulness of our approach by predicting correctly all the compounds in the test set, including not only compounds like HMR1556 and JNJ303, which appear as false negative results in hERG test, but also ebastine, which hERG assays predict as cardiotoxic in contradiction with in vivo results. Indeed, the advantage of the proposed prediction system over hERG blocking models can be easily illustrated. Focusing on the QT interval results only, we have inserted in Figure 5 a vertical line that represents a hERG IC_{50} of 1 μM ($pIC_{50} = 6$), a value typically considered as indicative of potential cardiotoxicity. In this map, all the compounds with pIC_{50} values enclosed in the region between the vertical line and the 110% CNT will be false positives (FP) and predicted incorrectly as cardiotoxic by hERG models, while the compounds on the region on the right top corner would be false negatives (FN) and pass the hERG filters only to be detected as cardiotoxic at latter stages. Examples of both categories are not rare and were included in the test set presented here.

The color-coded 2D graphics shown have been included for illustrating the parameters predicted by the electrophysiological models for diverse combinations of hERG and KCNQ1 blocking. These graphics have been obtained by precomputing a large number of models and have a high value on their own, since they can be used directly for estimating the effects of any drug with IC_{50} values within the explored range at any of the given concentrations (10, 100, or 200 nM). In particular, they can be used to produce a fast estimation of QT interval effect for screening a large collection of compounds, without the need to compute long and computationally demanding electrophysiological models. Also in respect to the method applicability, it must be mentioned that the model that we propose can make use of the IC_{50} computed at the molecular level or use experimental IC_{50} values obtained by classic patch clamp methods. Hence, this method can be applied at discovery stages, as a fast screening, for compounds still not synthesized or available in

low quantities as well as at preclinical stages, when the experimentally obtained IC_{50} for one or several ion channels can be integrated to obtain a prediction of a clinically relevant physiological end point.

In a broader perspective, the realistic prediction of clinically relevant outcomes is an attractive but extremely challenging objective. Recent efforts in this direction, like the early results of the Virtual Physiological Human project,[42] have shown that in most of the cases these predictions will require complex and multiscale modeling frameworks integrating different levels of description of the physiopathological phenomena, ranging from the biochemical reactions and interactions that are described at the molecular scale to adequate simulations of the tissues and organs that are involved. The multiscale simulation frameworks are relatively easy to envisage but extremely difficult to implement.

In this context, it is important to develop computational systems like the one presented here that, even still considering only in part the complexity of the biological phenomena under study, constitute proofs of concept of how multiscale predictive systems could work and what are their advantages in comparison to the classical single scale modeling approaches. Obviously, at present, the system is only a prototype; the quality of which can be extended in several different directions. At the molecular level, the quality of the predicted IC_{50} depends critically on the quality of the training set, which has been compiled from public heterogeneous sources. In corporate environments, the database can be easily extended, incorporating more homogeneous and accurate data that will necessarily produce more reliable predictions. At the electrophysiological level, our model included the two major ion channels involved in repolarization, but it could be extended by incorporating other actors (i.e., the Nav1.5 or the L-type Cav1.2 channels). In other order, the simple pore block model presented here can be replaced by models in which the interaction of the drugs with the diverse states of the ion channels are represented, thus incorporating the frequency-related effects of some drugs.[17] Moreover, the simulations could be improved by incorporating more computationally demanding but affordable 2D or 3D representations of cardiac tissue. The whole method is highly suitable for being implemented as a fully automatic prediction system accessible through a Web interface, in which the user can introduce the 2D structure, receiving in a few minutes the prediction of the effect of the compound on the QT interval.

In practice, a system based on the aforementioned principles can be integrated with other in silico prediction systems. Ideally, pharmacokinetics/pharmacodynamics simulation systems can provide information about the relevant drug concentrations which must be considered in the cardiotoxicity assessment, and metabolic predictions can suggest a set of metabolites that must be evaluated, together with the unaltered drug, in order to obtain more realistic predictions of the in vivo cardiotoxicity. In this sense, the predicted results obtained for ebastine in the aforementioned example are likely to be influenced by the metabolism of the original compound to carebastine, a well-known metabolite with reduced affinity for hERG.[40] The introduction of the ebastine metabolism in the system would probably predict a shorter QT elongation, in closer agreement with in vivo values. This highly desirable integration can be afforded only within computational frameworks like the one presented here, which has shown to be more useful than other current state-of-the-art approaches.

ASSOCIATED CONTENT

Ⓢ Supporting Information.   Detailed information of the databases of hERG and KCNQ1 blocker compounds and the statistics of the QSAR models is provided. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: manuel.pastor@upf.edu.

ACKNOWLEDGMENT

REFERENCES

(1) Hammond, T. G.; Carlsson, L.; Davis, A. S.; Lynch, W. G.; MacKenzie, I.; Redfern, W. S.; Sullivan, A. T.; Camm, A. J. Methods of collecting and evaluating non-clinical cardiac electrophysiology data in the pharmaceutical industry: results of an international survey. *Cardiovasc. Res.* **2001**, *49*, 741–750.

(2) Yap, Y. G.; Camm, A. J. Drug induced QT prolongation and torsades de pointes. *Heart* **2003**, *89*, 1363–1372.

(3) Hancox, J. C.; McPate, M. J.; El Harchi, A.; Zhang, Y. H. The hERG potassium channel and hERG screening for drug-induced torsades de pointes. *Pharmacol. Ther.* **2008**, *119*, 118–132.

(4) Yao, X.; Anderson, D. L.; Ross, S. A.; Lang, D. G.; Desai, B. Z.; Cooper, D. C.; Wheelan, P.; McIntyre, M. S.; Bergquist, M. L.; MacKenzie, K. I.; Becherer, J. D.; Hashim, M. A. Predicting QT prolongation in humans during early drug development using hERG inhibition and an anaesthetized guinea-pig model. *Br. J. Pharmacol.* **2008**, *154*, 1446–1456.

(5) Recanatini, M.; Poluzzi, E.; Masetti, M.; Cavalli, A.; De Ponti, F. QT prolongation through hERG K(+) channel blockade: current knowledge and strategies for the early prediction during drug development. *Med. Res. Rev.* **2005**, *25*, 133–166.

(6) Inanobe, A.; Kamiya, N.; Murakami, S.; Fukunishi, Y.; Nakamura, H.; Kurachi, Y. In silico prediction of the chemical block of human ether-a-go-go related gene (hERG) K^+ current. *J. Physiol. Sci.* **2008**, *58*, 459–470.

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

(8) Christ, T.; Wettwer, E.; Wuest, M.; Braeter, M.; Donath, F.; Champeroux, P.; Richard, S.; Ravens, U. Electrophysiological profile of propiverine--relationship to cardiac risk. *Naunyn Schmiedebergs Arch. Pharmacol.* **2008**, *376*, 431–440.

(9) Lee, H.; Kim, K. S.; Park, S. J.; Kim, E. J. Cellular mechanism of the QT prolongation induced by sulpiride. *Int. J. Toxicol.* **2009**, *28*, 207–212.

(10) Curtis, M. J. Is cardiac IKs a relevant drug target? *Cardiovasc. Res.* **2004**, *61*, 651–652.

(11) Jost, N.; Papp, J. G.; Varró, A . Slow delayed rectifier potassium current (IKs) and the repolarization reserve. *Ann. Noninvasive Electrocardiol.* **2007**, *12*, 64–78.

(12) Lengyel, C.; Varró, A.; Tábori, K.; Papp, J. G.; Baczkó, I. Combined pharmacological block of I(Kr) and I(Ks) increases short-term QT interval variability and provokes torsades de pointes. *Br. J. Pharmacol.* **2007**, *151*, 941–951.

(13) Michael, G.; Kane, K. A.; Coker, S. J. Adrenaline reveals the torsadogenic effect of combined blockade of potassium channels in anaesthetized guinea pigs. *Br. J. Pharmacol.* **2008**, *154*, 1414–1426.

(14) So, P. P.; Backx, P. H.; Dorian, P. Slow delayed rectifier K+ current block by HMR 1556 increases dispersion of repolarization and promotes Torsades de Pointes in rabbit ventricles. *Br. J. Pharmacol.* **2008**, *155*, 1185–1194.

(15) Towart, R.; Linders, J. T. M.; Hermans An, H.; Rohrbacher, J.; van der Linde, H. J.; Ercken, M.; Cik, M.; Roevens, P.; Teisman, A.; Gallacher, D. J. Blockade of the IKs potassium channel: An overlooked cardiovascular liability in drug safety screening? *J. Pharmacol. Toxicol. Methods* **2009**, *60*, 1–10.

(16) Luo, C. H.; Rudy, Y. A dynamic model of the cardiac ventricular action potential. I. Simulations of ionic currents and concentration changes. *Circ. Res.* **1994**, *74*, 1071–1096.

(17) Suzuki, S.; Murakami, S.; Tsujimae, K.; Findlay, I.; Karachi, Y. In silico risk assessment for drug-induction of cardiac arrhythmia. *Prog. Biophys. Mol. Biol.* **2008**, *98*, 52–60.

(18) Silva, J. R.; Pan, H.; Wu, D.; Nekouzadeh, A.; Decker, K. F.; Cui, J.; Baker, N. A.; Sept, D.; Rudy, Y. A multiscale model linking ion-channel molecular dynamics and electrostatics to the cardiac action potential. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 11102–11106.

(19) Bottino, D.; Penland, R. C.; Stamps, A.; Traebert, M.; Dumotier, B.; Georgiva, A.; Helmlinger, G.; Lett, G. S. Preclinical cardiac safety assessment of pharmaceutical compounds using an integrated systems-based computer model of the heart. *Prog. Biophys. Mol. Biol.* **2006**, *90*, 414–443.

(20) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Comput. Methodol.* **1990**, *3*, 537–547.

(21) Milletti, F.; Storchi, L.; Sforna, G.; Cruciani, G. New and original pKa prediction method using grid molecular interaction fields. *J. Chem. Inf. Model.* **2007**, *47*, 2172–2181.

(22) Milletti, F.; Storchi, L.; Sforna, G.; Cross, S.; Cruciani, G. Tautomer enumeration and stability prediction for virtual screening on large chemical databases. *J. Chem. Inf. Model.* **2009**, *49*, 68–75.

(23) Farid, R.; Day, T.; Friesner, R. A.; Pearlstein, R. A. New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem.* **2006**, *14*, 3160–3173.

(24) Lerche, C.; Buhova, I.; Lerche, H.; Steinmeyer, K.; Wei, A. D.; Strutz-Seebohm, N.; Lang, F.; Busch, A. E.; Zhorov, B. S.; Seebohm, G. Chromanol 293B binding in KCNQ1 (Kv7.1) channels involves electrostatic interactions with a potassium ion in the selectivity filter. *Mol. Pharmacol.* **2007**, *71*, 1503–1511.

(25) Verdonk, M. L.; Cole, J. V.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.

(26) Mooij, W. T. M.; Verdonk, M. L. General and targeted statistical potentials for protein-ligand interactions. *Proteins* **2005**, *61*, 272–287.

(27) Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J. Med. Chem.* **2000**, *43*, 3233–3243.

(28) Durán, A.; Martínez, G. C.; Pastor, M. Development and validation of AMANDA, a new algorithm for selecting highly relevant regions in Molecular Interaction Fields. *J. Chem. Inf. Model.* **2008**, *48*, 1813–1823.

(29) Faber, G. M.; Rudy, Y. Action potential and contractility changes in [Na(+)](i) overloaded cardiac myocytes: a simulation study. *Biophys. J.* **2000**, *78*, 2392–2404.

(30) Gima, K.; Rudy, Y. Ionic current basis of electrocardiographic waveforms: a model study. *Circ. Res.* **2002**, *90*, 889–896.

(31) Brennan, T.; Fink, M.; Rodríguez, B. Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *Eur. J. Pharm. Sci.* **2009**, *36*, 62–77.

(32) Drouin, E.; Charpentier, F.; Gauthier, C.; Laurent, K.; Le Marec, H. Electrophysiologic characteristics of cells spanning the left ventricular wall of human heart: evidence for presence of M cells. *J. Am. Coll. Cardiol.* **1995**, *26*, 185–192.

(33) Yan, G. X.; Shimizu, W.; Antzelevitch, C. Characteristics and distribution of M cells in arterially perfused canine left ventricular wedge preparations. *Circulation* **1998**, *98*, 1921–1927.

(34) Clayton, R. H.; Holden, A. V. Propagation of normal beats and re-entry in a computational model of ventricular cardiac tissue with regional differences in action potential shape and duration. *Prog. Biophys. Mol. Biol.* **2004**, *85*, 473–499.

(35) Snyders, D. J.; Chaudhary, A. High affinity open channel block by dofetilide of HERG expressed in a human cell line. *Mol. Pharmacol.* **1996**, *49*, 949–955.

(36) Jurkiewicz, N. K.; Sanguinetti, M. C. Rate-dependent prolongation of cardiac action potentials by a methanesulfonanilide class III antiarrhythmic agent. Specific block of rapidly activating delayed rectifier K+ current by dofetilide. *Circ. Res.* **1993**, *72*, 75–83.

(37) Gögelein, H.; Brüggermann, A.; Gerlach, U.; Brendel, J.; Busch, A. E. Inhibition of IKs channels by HMR 1556. *Naunyn Schmiedebergs Arch. Pharmacol.* **2000**, *362*, 480–488.

(38) Gerlach, U.; Brendel, J.; Lang, H. J.; Paulus, E. F.; Weidmann, K.; Brüggermann, A.; Busch, A. E.; Suessbrich, H.; Bleich, M.; Greger, R. Synthesis and activity of novel and selective I(Ks)-channel blockers. *J. Med. Chem.* **2001**, *44*, 3831–3837.

(39) 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*, 233–244.

(40) Rico, S.; Antonijoan, R. M.; Barbanoj, M. J. Ebastine in the light of CONGA recommendations for the development of third-generation antihistamines. *J. Asthma Allergy* **2009**, *2*, 73–92.

(41) Polak, S.; Wiśniowska, B.; Brandys, J. Collation, assessment and analysis of literature in vitro data on hERG receptor blocking potency for subsequent modeling of drugs' cardiotoxic properties. *J. Appl. Toxicol.* **2009**, *29*, 183–206.

(42) Kohl, P.; Noble, D. Systems biology and the virtual physiological human. *Mol. Syst. Biol.* **2009**, *5*, 292.