

2020 CASERB

LIBRO DE ACTAS

XXXVIII Congreso
Anual de la
Sociedad Española
de Ingeniería Biomédica

ORGANIZADORES Y COLABORADORES:



GRUPO DE INGENIERÍA BIOMÉDICA
UNIVERSIDAD DE VALLADOLID



BioRes
BIOENGINEERING
AND RESUSCITATION

Sociedad Española de
INGENIERÍA BIOMÉDICA



Dispositivo laparoscópico mecánico articulado: Análisis de la curva de aprendizaje	129
Modelos Predictivos en el Tratamiento de Pacientes con Alcoholismo	133
Clasificación de bloqueadores de Ikr basada en protocolos de voltage clamp y técnicas de machine learning	137
Arrhythmogenicity quantification of two genetic defects affecting IKr channel in AF patients	141
Feature Selection and Machine Learning for Predicting Multi-drug Resistance just After ICU Admission	145
Construcción de pupilómetro binocular para estudio de lentes intraoculares	149

Clasificación de Bloqueadores de I_{Kr} Basada en Protocolos de Voltage Clamp y Técnicas de Machine Learning.

Fernando Escobar¹, Julio Gomis-Tena¹, Javier Saiz¹, Lucía Romero¹

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

Resumen

La evaluación de la cardiotoxicidad es clave en el desarrollo de nuevos compuestos y se suele realizar analizando el grado de bloqueo que provocan en los canales de potasio relacionados con el gen humano ether-à-go-go (hERG). El objetivo de este trabajo es desarrollar un clasificador que determine la preferencia de un fármaco para unirse a un estado determinado del canal.

Se ha creado un conjunto de 2600, divididos en 13 clases, bloqueadores virtuales con diferentes afinidades y cinéticas por los estados conformacionales del canal. Se realizaron simulaciones utilizando tres protocolos de estimulación que aumentan la probabilidad del canal de ocupar ciertos estados. Para cada simulación se han tomado tres medidas: IC_{50} , la constante de recuperación de la corriente de potasio I_{Kr} y una estimación del tiempo necesario para alcanzar el estado estacionario. Por lo tanto, se obtuvieron 9 variables por cada fármaco estudiado. Se desarrolló, entrenó y evaluó un clasificador en dos pasos. En primer lugar, se usaron máquinas de soporte vectorial sobre las IC_{50} para separar las 13 clases en tres grupos con 4, 5 y 4 clases respectivamente. En segundo lugar, se usaron redes neuronales en cada grupo para terminar de clasificar los bloqueadores.

Los tres clasificadores obtuvieron una precisión global en los grupos de test del 90.83, 88.66 y 89.16% respectivamente.

1. Introducción

La evaluación preclínica de la cardiotoxicidad de nuevos fármacos se ha convertido en una prioridad para las compañías farmacéuticas debido a potenciales efectos adversos como arritmias ventriculares [1]. Hoy en día, la forma más común de evaluar la seguridad de un fármaco se basa en medir el bloqueo in vitro que producen en el gen humano ether-à-go-go (hERG) y la prolongación in vivo del intervalo QT. Ambos fenómenos han sido relacionados con la aparición de Torsades de Pointes (TdP).

Aunque estos test de seguridad han tenido éxito en evitar que fármacos potencialmente peligrosos llegasen al mercado, también han evitado el desarrollo de compuestos con potencial de ser útiles. De hecho, existen fármacos como el verapamil, reconocido como un potente bloqueador de la corriente I_{Kr} , que no está asociado al desarrollo de TdP.

Estudios recientes han demostrado que la evaluación de la seguridad de fármacos mejora cuando se tienen en cuenta las dinámicas y cinéticas de los mismos [2]. Las

diferencias en la unión de un compuesto al canal dependiendo del estado conformacional en el que se encuentre el canal pueden alterar significativamente los valores de la IC_{50} , que es la concentración de fármaco a la cual se produce la mitad del bloqueo de una corriente iónica. A su vez, dichos valores también pueden variar dependiendo del protocolo de estimulación utilizado [3]. Por lo tanto, considerar el comportamiento del fármaco y estandarizar los protocolos empleados es necesario para mejorar la evaluación de los mismos. Un estudio reciente ha propuesto un conjunto de tres protocolos que maximizan la probabilidad del canal I_{Kr} de encontrarse en cada uno de los estados posibles [4]. En este trabajo se utilizarán dichos protocolos con el objetivo de determinar el estado de unión preferente haciendo uso de técnicas de *machine learning*.

2. Materiales y métodos

2.1. Modelos

Se han utilizado seis variantes del modelo de Markov de I_{Kr} de ventrículo humano de Fink et al. [5] para simular la interacción fármaco-canal, modelos mostrados en la figura 1. El modelo sin fármaco original tiene cinco estados conformacionales, tres cerrados (C_1 , C_2 , y C_3), uno abierto (O) y uno inactivado (I). Se incorporaron otros cinco estados para representar los estados del canal a los que el fármaco está unido (C_{1d} , C_{2d} , C_{3d} , O_d and I_d). Las transiciones entre los estados con y sin fármaco están reguladas por una constante denominada $k_{(C, O, I)}$ que se multiplica por la concentración de fármaco para obtener la velocidad de transición. La liberación de un compuesto desde un estado está marcada por la velocidad de transición denominada $r_{(C, O, I)}$.

2.2. Protocolos

Tres protocolos de *voltage clamp* (mostrados en la figura 2) previamente desarrollados se utilizaron para simular los efectos de los fármacos [4]. Dichos protocolos, denominados P40, P0 y P-80, maximizan el tiempo que el canal permanece a 40, 0 y -80 mV

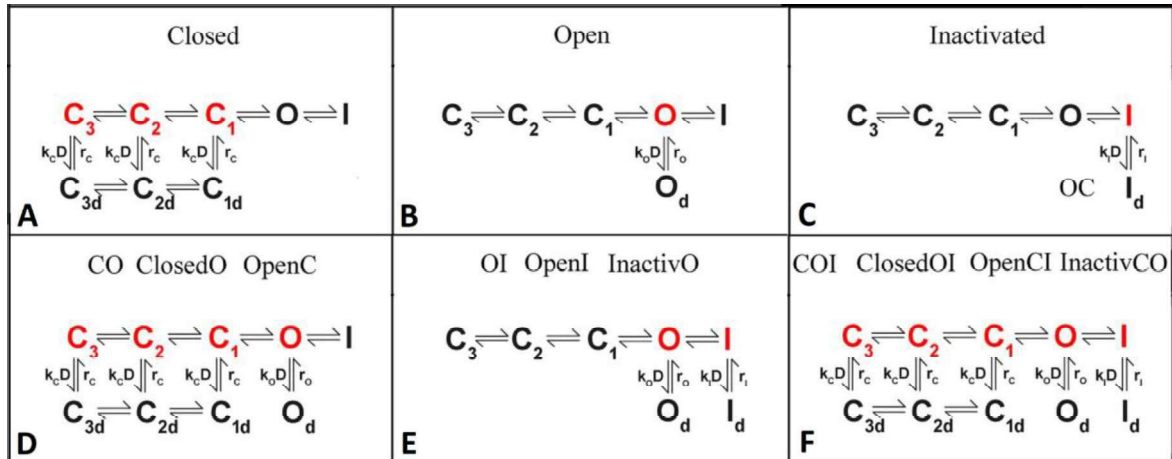


Figura 1: Modelos de Markov con estados libres (A, B y C) y con fármaco (D, E y F). D es la concentración de fármaco, k_c , k_o y k_i se multiplican por D para calcular la velocidad de asociación desde los estados cerrados, abiertos e inactivados libres a los ocupados respectivamente. r_c , r_o y r_i son las velocidades de disociación del compuesto. Las letras rojas indican estados a los que se une el fármaco. Los nombres de las clases parecen con el modelo de Markov correspondiente.

respectivamente. P40 consiste en un paso de acondicionamiento de 5 s a 40 mV seguido de un paso de test de 0.2 s a -60 mV y 0.2 s de recuperación a -80 mV. P0 es similar a P40 pero el paso de acondicionamiento se produce a 0 mV. P-80 consiste en un test de 0.5 s a -50 mV y un paso de acondicionamiento de 4.5 s a -80 mV. Los protocolos aumentan la posibilidad del canal de encontrarse en el estado inactivado, abierto y cerrado, respectivamente [6].

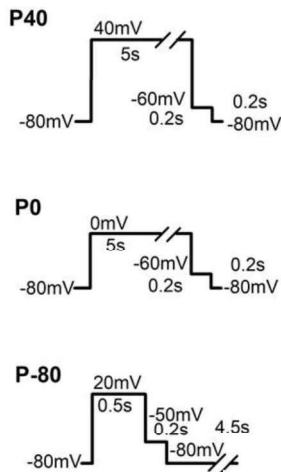


Figura 2: Protocolos de voltage clamp utilizados para simular los efectos de fármacos.

2.3. Fármacos

Se creó un conjunto de 2600 fármacos virtuales con diferentes dinámicas y cinéticas por lo estados conformacionales del canal. Los nombres de las diferentes clases de compuesto aparecen con el modelo de Markov correspondiente en la figura 1 y se definen como: Fármacos que se unen exclusivamente al estado cerrado (Closed), abierto (Open) e inactivado (Inactivated). Fármacos que se unen simultáneamente al estado cerrado y al abierto, con igual afinidad (CO), con preferencia por el cerrado (ClosedO) o por el abierto (OpenC). Compuestos que se unen con igual afinidad a los estados abierto e inactivado (OI), con preferencia por el abierto (OpenI) y con preferencia por el inactivado (InactivO). Finalmente, fármacos que se unen a los tres estados a la vez, de igual manera (COI), con mayor afinidad por el cerrado (ClosedOI), con preferencia por el abierto (OpenCI) y con preferencia por el inactivado (InactivCO). En todos los casos, la velocidad de asociación es igual para todos los estados bloqueados. La mayor afinidad por un determinado estado se simula multiplicando por 100 la constante de disociación del estado correspondiente. Se simularon 200 fármacos de cada clase, cuyas constantes k y r fueron generadas aleatoriamente dentro de un rango de valores fisiológicos usando Matlab (Mathworks Inc.).

3. Resultados

Los efectos de los fármacos virtuales se simularon usando los tres protocolos mencionados anteriormente. Para cada simulación se obtuvieron tres variables por cada protocolo, 1) IC_{50} , el parámetro más utilizado para evaluar la seguridad cardíaca, 2) un estimación basada en derivadas del número de pulsos requeridos para alcanzar el

estado estacionario y 3) la constante de recuperación de la evolución de la inhibición de la corriente de potasio a la concentración de la IC_{50} estimada mediante un ajuste exponencial. Por lo tanto, se obtuvieron 9 variables por compuesto.

Tras obtener todas las variables se desarrolló un clasificador en dos pasos para intentar clasificar los fármacos en las trece clases objetivo (mostradas en la figura 1). Todo el análisis de datos se realizó con Matlab.

3.1. Máquinas de soporte vectorial

Se realizó un análisis exploratorio de los datos para visualizar las variables obtenidas. Se construyeron gráficos 3D donde se mostraron los valores de cada variable asignando cada protocolo a un eje. La representación 3D de la IC_{50} se muestra en la figura 3 (gráficos del resto de variables no mostrados).

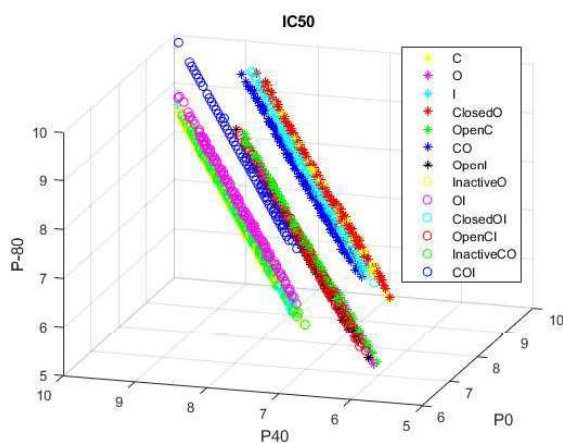


Figura 3: representación 3D de los valores de IC_{50} de los 2600 bloqueadores virtuales. Los ejes representan el $-\log([IC_{50}])$ obtenido con los tres protocolos (P40, P0 y P-80).

Como se observa en la figura 3, existen tres grupos claramente diferenciados entre los fármacos, separables mediante un hiperplano. Por este motivo se decidió utilizar máquinas de soporte vectorial (SVM) para crear las separaciones entre grupos. Por lo tanto, se entrenaron dos SVMs utilizando un enfoque de uno contra todos. El primer SVM separaba el grupo superior, considerando todos los compuestos presentes como una única clase (Figura 4, panel superior). El segundo SVM separaba el grupo inferior del resto, de nuevo considerando como una todas las clases presentes en el mismo (Figura 4, panel inferior). Todos aquellos compuestos que no pertenecieran a ninguno de los dos grupos anteriores se asignaron al grupo central.

Este primer paso separaba los fármacos en tres grupos utilizando exclusivamente los valores de IC_{50} . El primer grupo se corresponde con fármacos con mayor afinidad por el estado cerrado y los que bloqueaban de igual manera el

cerrado y el abierto (Closed, ClosedO, ClosedOI, CO). El grupo central contenía los compuestos con preferencia por el estado abierto y con igual afinidad por los tres estados a la vez (Open, OpenC, OpenI, OpenCI, COI). Finalmente, el grupo inferior agrupaba los compuestos con mayor afinidad por el estado inactivado y los que bloqueaban simultáneamente y de igual manera el abierto e inactivado (Inactivated, InactivO, InactivCO, OI).

Por lo tanto, haciendo uso solamente de la IC_{50} obtenida para cada protocolo se pueden separar los compuestos en base a su estado preferente.

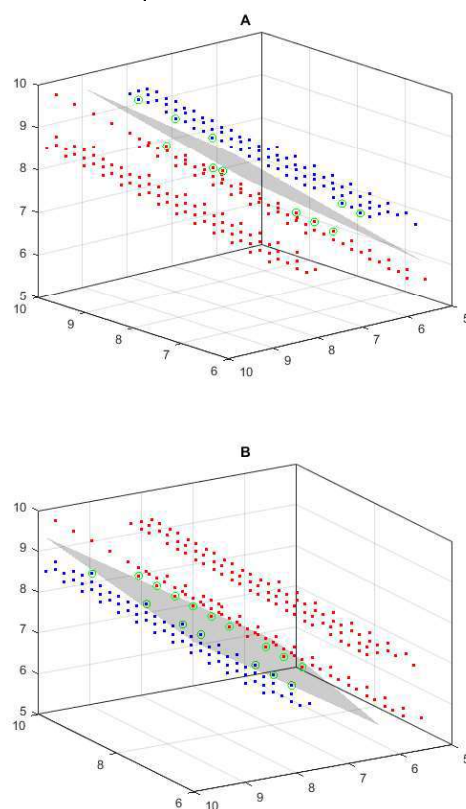


Figura 4: Hiperplanos creados mediante los SVMs para separar los compuestos en tres grupos por los valores de IC_{50} obtenidos con cada protocolo. (A) SVM que separa el grupo superior (preferencia por el estado cerrado). (B) SVM para separar el grupo inferior (preferencia por el estado inactivado). Los puntos azules representan fármacos que pertenecen al grupo separado y los rojos aquellos que no.

3.2. Redes Neuronales

Tras el primer paso del proceso de clasificación, se entrenó y evaluó un conjunto de tres redes neuronales utilizando la *toolbox* de Matlab. Las redes se usaron para clasificar los fármacos de cada uno de los grupos en las trece clases objetivo. Los grupos de entrenamiento y test fueron seleccionados automáticamente por la *toolbox*. 120,

126 y 120 fármacos fueron seleccionados para cada uno de los grupos de test. Para este paso se consideraron las nueve variables obtenidas. Las redes tienen una capa de entrada con 9 nodos (tantos como variables usadas), una capa oculta y una capa de salida con 3, 4 y 3 nodos respectivamente. El número de neuronas en la capa oculta se seleccionó evaluando un rango de 1 a 18 (el doble de las neuronas de la capa de entrada) y seleccionando el mejor resultado [7]. Para cada grupo, el mejor resultado se obtenía con 11, 7 y 9 neuronas respectivamente. Los resultados para los grupos de test se muestran en la figura 5. La figura muestra una precisión total de 90.83, 88.66 y 89.16% para los grupos de test.

Figura 5: Matrices de confusión de los grupos de test para las redes neuronales. El número superior en las celdas representa el número absoluto de fármacos y el inferior el porcentaje relativo. (A) grupo de preferencia por el estado cerrado. (B) grupo de preferencia por el estado abierto. (C) grupo de preferencia por el estado inactivado. Las filas se corresponden con las clases predichas y las columnas con las clases reales. Celdas azules muestran los fármacos clasificados correctamente y las celdas naranjas los incorrectamente clasificados. La columna de la derecha muestra el valor predictivo positivo (verde) y ratio de falsos (rojo). La fila de abajo muestra el ratio de verdaderos positivos (verde) y el ratio de falsos negativos (rojo). La celda gris oscura muestra la precisión total.

		A Closed state preference group				
Output Class	C	32 26.7%	4 3.3%	0 0.0%	1 0.8%	86.5% 13.5%
	ClosedO	1 0.8%	19 15.8%	0 0.0%	4 3.3%	79.2% 20.8%
	CO	0 0.0%	0 0.0%	30 25.0%	0 0.0%	100% 0.0%
	ClosedOI	0 0.0%	0 0.0%	1 0.8%	28 23.3%	96.6% 3.4%
		97.0% 3.0%	82.6% 17.4%	96.8% 3.2%	84.8% 15.2%	90.8% 9.2%
	C	ClosedO	CO	ClosedOI		
		Target Class				

		B Open state preference group					
Output Class	O	26 17.3%	1 0.7%	1 0.7%	1 0.7%	0 0.0%	89.7% 10.3%
	OpenC	2 1.3%	30 20.0%	1 0.7%	1 0.7%	0 0.0%	88.2% 11.8%
	OpenI	5 3.3%	0 0.0%	21 14.0%	0 0.0%	0 0.0%	80.8% 19.2%
	OpenCI	1 0.7%	3 2.0%	1 0.7%	20 13.5%	0 0.0%	80.0% 20.0%
	COI	0 0.0%	0 0.0%	0 0.0%	0 0.0%	36 24.0%	100% 0.0%
	76.5% 23.5%	88.2% 11.8%	87.5% 12.5%	90.9% 9.1%	100% 0.0%	88.7% 11.3%	
	O	OpenC	OpenI	OpenCI	COI		
		Target Class					

		C Inactivated state preference group				
Output Class	I	26 21.7%	5 4.2%	0 0.0%	1 0.8%	81.3% 18.8%
	InactiveO	2 1.7%	18 15.0%	0 0.0%	1 0.8%	85.7% 14.3%
	OI	0 0.0%	0 0.0%	35 29.2%	2 1.7%	94.6% 5.4%
	InactiveCO	0 0.0%	2 1.7%	0 0.0%	28 23.3%	93.3% 6.7%
		92.9% 7.1%	72.0% 28.0%	100% 0.0%	87.5% 12.5%	89.2% 10.8%
	I	InactiveO	OI	InactiveCO		
		Target Class				

4. Conclusión

Hemos desarrollado una herramienta útil para la clasificación de fármacos de acuerdo a sus afinidades y basándonos en medidas obtenidas de protocolos de *voltage clamp* y con elevada precisión

5. Referencias

[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(4): 741–750.

[2] 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

[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>