

<i>Nereis. Revista Iberoamericana Interdisciplinar de Métodos, Modelización y Simulación</i>	9	49-62	Universidad Católica de Valencia San Vicente Mártir	Valencia (España)	ISSN 1888-8550
--	---	-------	---	-------------------	----------------

Aplicación de la Topología Molecular en la búsqueda de nuevos compuestos basados en Azaauronas derivados de las Auronas naturales como potenciales antimaláricos

Application of Molecular Topology in the search for new composites based on Azaaurones derived from natural Aurones as potential antimalarial

Fecha de recepción y aceptación: 28 de enero de 2017 y 1 de marzo de 2017

Jenny Soraya Carrillo^{1,3}, Cleide Rizza^{1,4}, Blanca Esthela Álvarez^{1,5}, Diana Hernández^{1,6}, Jorge Gálvez¹ y Ramón García-Domenech^{1,2*}

¹ Máster en Enfermedades Parasitarias Tropicales, Departamento de Farmacia y Tecnología Farmacéutica y Parasitología, Universitat de València.

² Departamento de Química Física, Facultad de Farmacia, Universitat de València.

³ Facultad de Medicina Veterinaria y Zootecnia, Universidad Central del Ecuador.

⁴ Facultad de Ciencias de la Salud, Universidad de Carabobo, Venezuela.

⁵ Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero, México.

⁶ Facultad de Ciencias de la Salud, Universidad de Montemorelos, México.

* Correspondencia: Universitat de València. Facultad de Farmacia. Departamento de Química Física. Avenida V. A. Estellés, s/n. 46100 (Burjassot), Valencia. España. E-mail: ramon.garcia@uv.es



RESUMEN

La malaria es una enfermedad parasitaria transmitida por vectores del género *Anopheles*, que tiene un efecto devastador en la salud pública a nivel mundial. En 2015, se reportaron 438.000 muertes por malaria, siendo el continente africano el más afectado. Una de las causas de esta preocupante tasa de mortalidad es el surgimiento de farmacoresistencias en las diferentes especies de *Plasmodium*. Por esto es importante el desarrollo de nuevos agentes antimaláricos. En el presente estudio se utilizó la topología molecular para el desarrollo de un modelo QSAR capaz de predecir la actividad antimalárica de un grupo de compuestos derivados de las auronas naturales. Utilizando el análisis lineal discriminante se encontró un modelo capaz de clasificar correctamente la actividad de 33 de los 35 compuestos estudiados (93 %). Con el fin de predecir la potencia antimalárica de los compuestos, se realizó un análisis de regresión multineal, que fue capaz de explicar el 83 % de la varianza. El modelo fue validado mediante una validación cruzada y un test de aleatoriedad. Finalmente, se aplicó a la búsqueda de nuevos compuestos con potencial actividad antimalárica.

PALABRAS CLAVE: *Plasmodium*, malaria, auronas, *Anopheles*, análisis QSAR, topología molecular.

ABSTRACT

Malaria is a parasitic disease transmitted by vectors of the genus *Anopheles*, which has a devastating effect on public health globally. In 2015, were reported 438.000 Malaria deaths, with the African continent being the most affected. One of the causes of this concerning rate of deaths is the appearance of pharmacoresistences in the different species of *Plasmodium*. For this reason is important the developing of new antimalarial agents. In the present study we used molecular topology to develop QSAR models capable of predicting the antimalarial activity of a group of compounds derived from natural Aurones. Using the linear discriminant analysis, a model was found capable of classify correctly the activity of 33 out of the 35 compounds studied (93%). In order to predict the compounds' antimalarial potency, we performed a multilinear regression analysis, which was capable of explaining 83% of the variance. The model was validated by a cross-validation and a randomness test. Finally, the selected model was applied in the search for new compounds with potential antimalarial activity.

KEYWORDS: *Plasmodium*, Malaria, Aurones, *Anopheles*, QSAR Analysis, Molecular Topology



INTRODUCCIÓN

La malaria es una enfermedad parasitaria transmitida por protozoos del género *Plasmodium*. En humanos la malaria es causada por cuatro especies de *Plasmodium* (*P. falciparum*, *P. vivax*, *P. malariae* y *P. ovale*), las cuales son transmitidas por mosquitos hembra del género *Anopheles*. En 2015 se estimaron 214 millones de casos de malaria en el mundo y 438.000 muertes, siendo la especie *Plasmodium falciparum* la responsable del mayor número de muertes por malaria y la más prevalente en el continente africano (WHO, 2015).

Hoy en día la malaria sigue siendo un problema de gran importancia a nivel mundial. Sin embargo, dentro de las comunidades más pobres se han logrado, durante los últimos 4-5 años, avances sustanciales en la instrucción y ampliación de los programas para proporcionar prevención y tratamiento a las personas más afectadas por esta enfermedad (WHO, 2016; White, 1999).

Entre las causas de la alta mortandad de la malaria se encuentran la resistencia del parásito a los fármacos antimaláricos de uso común (Souard *et al.*, 2010), el deterioro de servicios primarios de salud en muchas áreas y la resistencia a insecticidas utilizados para el control de vectores (WHO, 2016; Achan *et al.*, 2011).

La artemisinina y sus derivados hemisintéticos, artemeter y artesunato, que comparten una parte de endoperóxido, están siendo usados en terapias combinadas basadas en artemisinina (TCA) como tratamiento de primera línea para malaria no complicada de *P. falciparum*, lo que lleva a los organismos de salud oficiales a tener como prioridad global a nivel de salud pública la protección de la eficacia de las terapias combinadas (Ashley *et al.*, 2014; Muraleedharan y Avery, 2009). No obstante, en la actualidad ya se ha reportado resistencia de *P. falciparum* a la artemisinina en cinco países del sureste de Asia de la subregión del Gran Mekong (Camboya, Laos, Birmania o Myanmar, Tailandia y Vietnam), habiéndose encontrado a lo largo de la frontera entre Camboya y Tailandia resistencia de *P. falciparum* a casi todos los medicamentos disponibles (WHO, 2016).

Una de las investigaciones recientes se refiere a un grupo de 35 compuestos análogos de auronas naturales que presentan actividad antimalárica a nivel submicromolar (Souard *et al.*, 2010). Las auronas presentan un efecto inhibitorio *in vitro* de *Plasmodium falciparum* en la fase eritrocítica y de la bomba de flujo implicados en la resistencia del parásito, sin embargo también se han encontrado como blancos validados con expresión en la vacuola digestiva del parásito (Carrasco *et al.*, 2014; Muraleedharan y Avery, 2009). Según lo mencionado anteriormente, se plantea la necesidad de encontrar nuevas terapias que funcionen también contra el estadio hepático de la enfermedad, destacándose las azaauronas como una plataforma prometedora para los antimaláricos de doble etapa (Carrasco *et al.*, 2016).

Debido a la amplia propagación de la malaria y la resistencia del parásito a los fármacos de primera línea, se plantea desarrollar fármacos más activos y económicos para combatir la enfermedad (Mishra *et al.*, 2016).

Uno de los requisitos de la investigación sobre nuevos fármacos antimaláricos es que sean económicamente sostenibles, por ello se emplean técnicas *in silico* (Kier y Hall, 1986), basadas en los estudios de relación cuantitativa estructura-actividad o *quantitative structure-activity relationship* (QSAR). En efecto, esos métodos pueden reducir los costosos fallos en la selección de compuestos que son candidatos a fármacos en los ensayos clínicos, mediante el filtrado de bibliotecas combinatorias (Galvez *et al.*, 2012b). También son especialmente funcionales en ausencia de estructuras tridimensionales de dianas de fármacos específicos (Wang, Wu, Lin y Yang, 2015).

Dentro de los métodos QSAR destacan los que emplean los índices topológicos, ya que han llegado a predecir de forma rápida y precisa diversas propiedades fisicoquímicas y biológicas en grupos de compuestos con una amplia diversidad estructural. Además se han encontrado nuevos candidatos que tienen actividad antiinflamatoria (Gálvez-Llompарт *et al.*, 2011), antiparasitaria (García-Domenech *et al.*, 2008; Mahmoudi *et al.*, 2008; Martínez *et al.*, 2015), antimicrobiana, antiviral, antiherpética (de Julián-Ortiz *et al.*, 1999), antiepiléptica (Bruno-Blanch *et al.*, 2003), antifúngicos (Pastor *et al.*, 1998), antihistamínicos (Duart *et al.*, 2001), actividad de repelente frente a mosquitos (Martínez *et al.*, 2014) y antineoplasias específicamente en melanoma (Kar y Roy, 2012; Lee *et al.*, 2012).

El objetivo del presente trabajo es la elaboración de un modelo topológico-matemático de predicción QSAR usando la topología molecular, el análisis lineal discriminante y el análisis de regresión multilíneal. Las moléculas estudiadas fueron las azaauronas, que son derivados de las auronas descritos en investigaciones previas, siendo el propósito el de obtener nuevos compuestos con potencial actividad antimalárica.



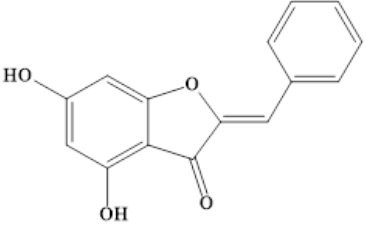
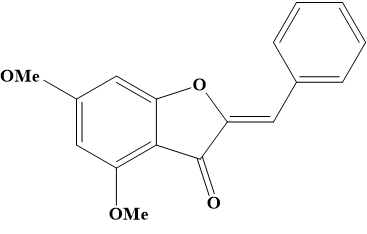
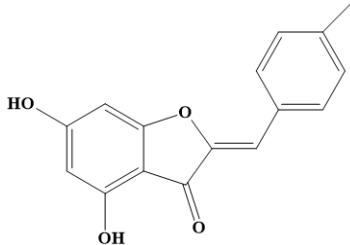
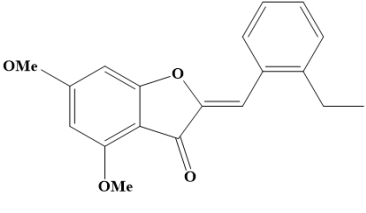
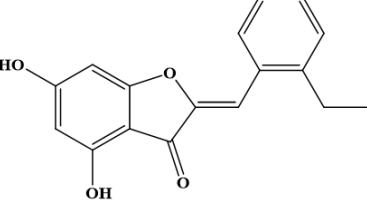
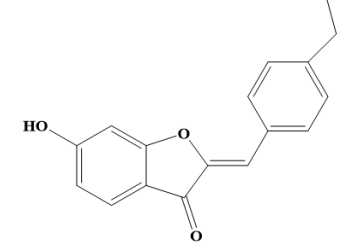
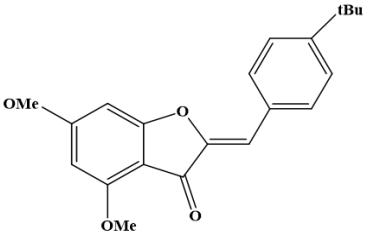
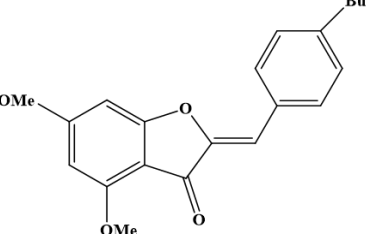
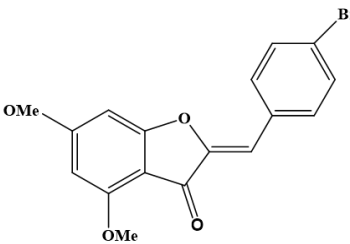
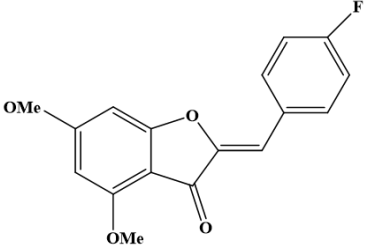
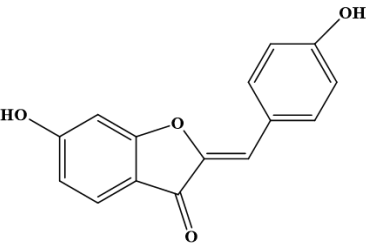
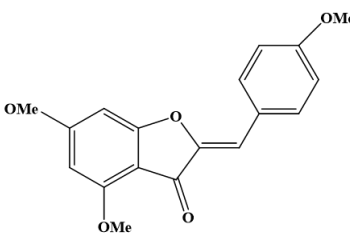
MATERIAL Y MÉTODOS

Compuestos analizados y ensayos realizados

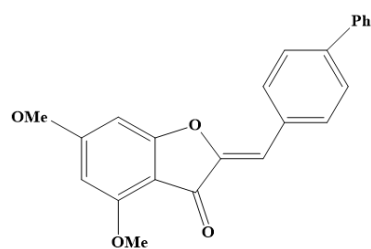
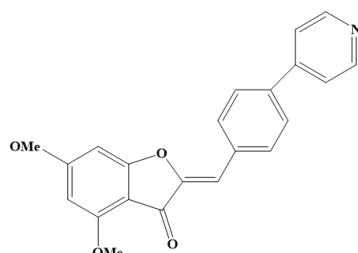
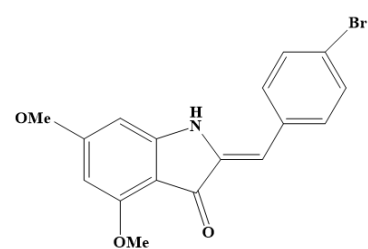
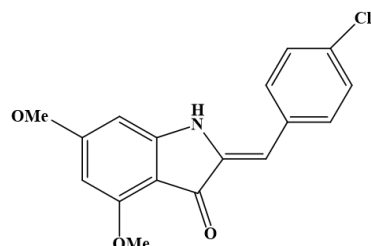
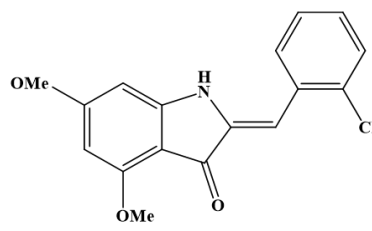
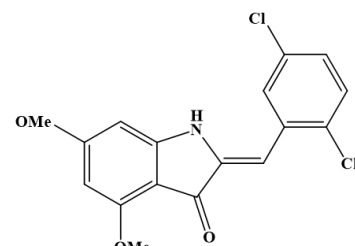
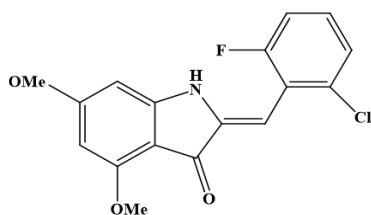
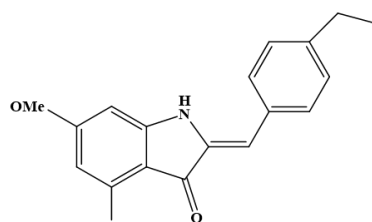
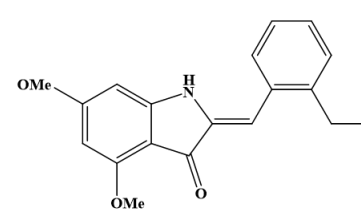
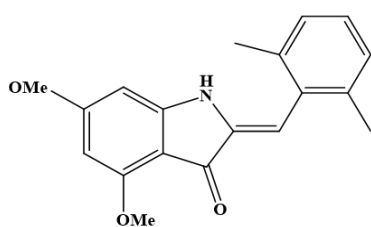
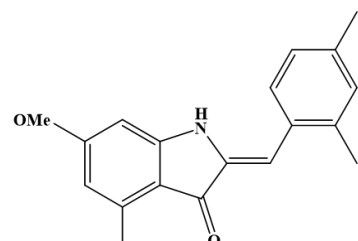
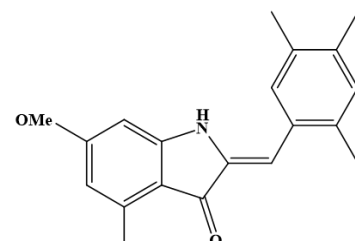
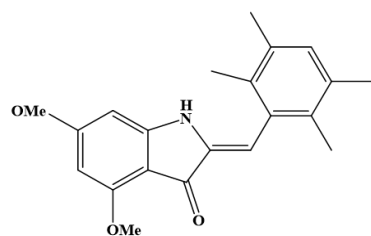
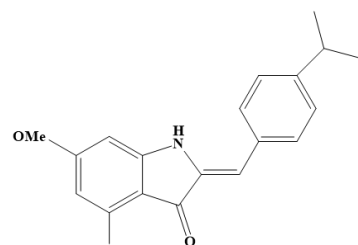
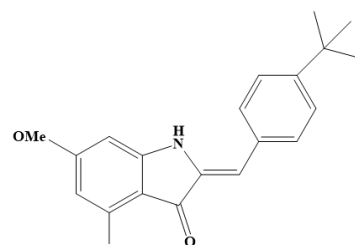
Para este estudio se seleccionaron 35 compuestos (14 auronas y 21 azaauronas) diseñados como análogos de las auronas naturales (2-bencilidenbenzofuran-3 (2H) -onas); de estos compuestos varios presentaron propiedades antimaláricas, las cuales han sido comprobadas experimentalmente en estudios previos, y por ende se dispone de los valores de concentración IC_{50} (μM), así como de su estructura química (Souard *et al.*, 2010) (tabla1).

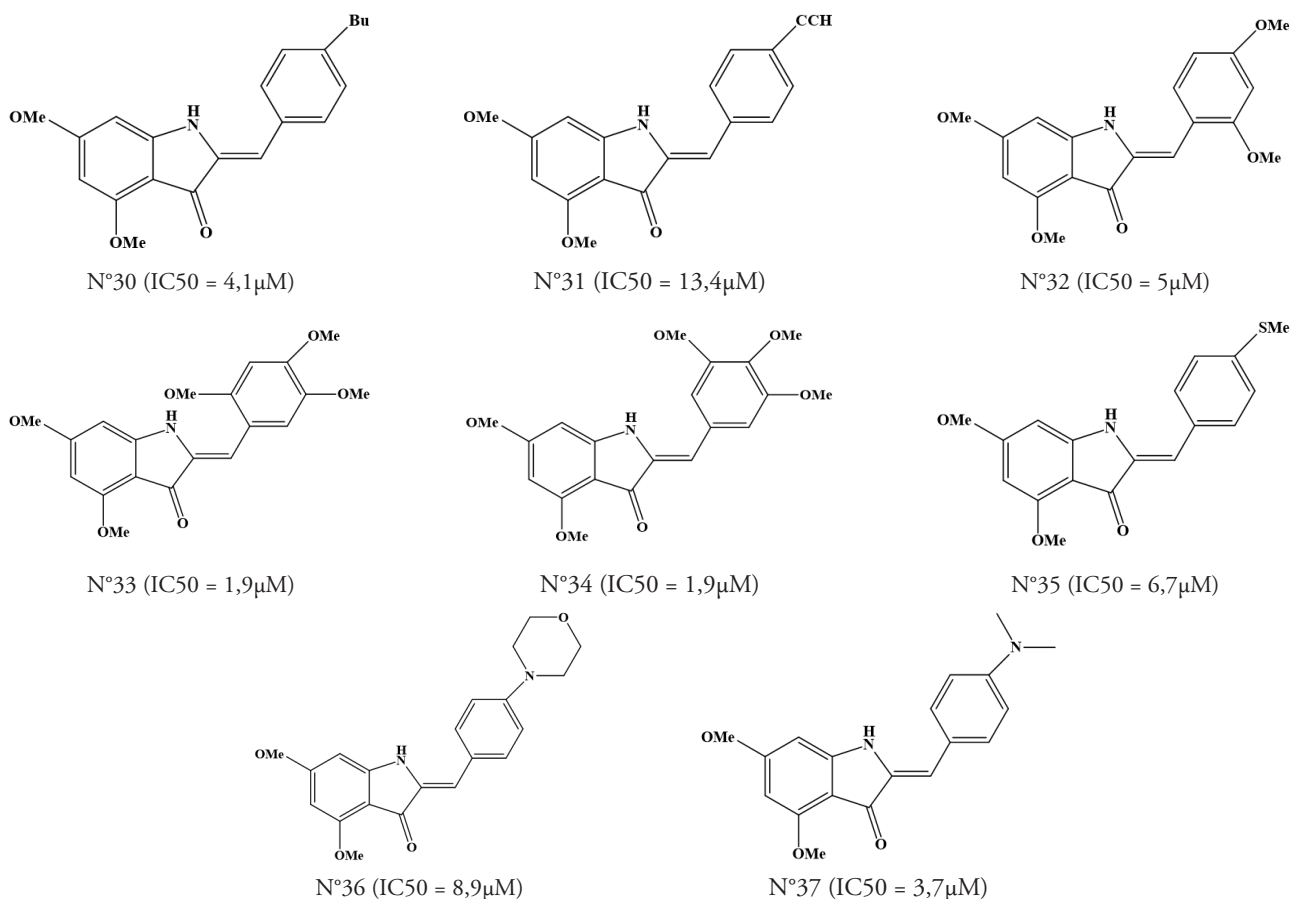
Cabe señalar que se ha mantenido la misma numeración de cada molécula con respecto al trabajo original de donde fueron tomados los compuestos que aparecen en la tabla 1 (Souard *et al.*, 2010).

Tabla 1. Estructura química de los compuestos estudiados y actividad antimalárica

 N°3 (IC_{50} = 94,5μM)	 N°4 (IC_{50} = 60,3μM)	 N°5 (IC_{50} = 63,4μM)
 N°6 (IC_{50} = 21μM)	 N°7 (IC_{50} = 113,5μM)	 N°8 (IC_{50} = 28μM)
 N°9 (IC_{50} = 13,3μM)	 N°10 (IC_{50} = 11,8μM)	 N°11 (IC_{50} = 49,8μM)
 N°12 (IC_{50} = 86,7μM)	 N°13 (IC_{50} = 130μM)	 N°14 (IC_{50} = 11μM)



N°15 (IC₅₀ = 234 μM)N°16 (IC₅₀ = 85 μM)N°17 (IC₅₀ = 49,8 μM)N°18 (IC₅₀ = 17 μM)N°19 (IC₅₀ = 9,9 μM)N°20 (IC₅₀ = 8,4 μM)N°21 (IC₅₀ = 9 μM)N°22 (IC₅₀ = 1 μM)N°23 (IC₅₀ = 12,8 μM)N°24 (IC₅₀ = 9,1 μM)N°25 (IC₅₀ = 3,6 μM)N°26 (IC₅₀ = 5,6 μM)N°27 (IC₅₀ = 8,9 μM)N°28 (IC₅₀ = 4,4 μM)N°29 (IC₅₀ = 7,2 μM)



Descriptores moleculares

Para dibujar la estructura química de las moléculas se utilizó el software ChemDraw® Professional 15.1. Posteriormente se importaron las moléculas al software DesMol2 (DesMol2), y se calcularon 62 descriptores topológicos, tales como: índices de conectividad Randić, Kier y Hall de hasta cuarto orden (${}^m\chi_t$, ${}^m\chi_v$) (L. Kier y Hall, 1976; L. B. Kier y Hall, 1986b); índices topológicos de carga (ITC) de hasta quinto orden (J_m , G_m , f_m^v , G_m^v) (Gálvez *et al.*, 1994); cocientes y diferencias entre índices de conectividad de valencia y no valencia (${}^mG_t = {}^m\chi_t / {}^m\chi_v$ y ${}^mD_t = {}^m\chi_t - {}^m\chi_v$); índice de Wiener (W) (Wiener, 1947), y descriptores constitucionales sencillos como el número de vértices (N), longitud (diámetro) del grafo (L), número de ramificaciones (R), pares de ramificaciones que están separados a distancias 1, 2 y 3 (PRi), etc.

Modelo QSAR

Para los análisis se utilizó el software STATISTICA, en su versión 8.0 (Statistics Statsoft, 2008).



Análisis lineal discriminante

El análisis lineal discriminante (ALD) es un algoritmo que nos permite distinguir entre dos o más categorías u objetos por medio de una función lineal. En nuestro caso se trata de diferenciar o discriminar entre compuestos activos e inactivos según los valores de los descriptores de sus moléculas (García-Doménech *et al.*, 2008; Izenman, 2013).

En este estudio, el grupo activo lo conforman compuestos con probada actividad antimalárica y con $IC_{50} < 10\mu M$, y el grupo inactivo, los compuestos con un $IC_{50} > 10\mu M$. La capacidad discriminante se determinó en función del porcentaje de compuestos clasificados correctamente dentro de cada grupo.

La selección de los descriptores se basó en el parámetro de distribución Fisher-Snedecor (F), y el criterio de clasificación empleado fue la distancia de Mahalanobis, clasificándose como activo o inactivo cualquier compuesto en función de su distancia más corta al centroide correspondiente (Gálvez *et al.*, 2012a; Izenman, 2013). El programa estadístico selecciona las variables para el cálculo de la función discriminante (FD) de manera sucesiva (*stepwise*), es decir, revisa todas las variables y la que contribuya más a la discriminación de los grupos será incluida en el modelo, mientras que la variable que menos contribuye a la predicción es eliminada. Para evaluar la calidad de la función discriminante se empleó el parámetro λ de Wilks (también llamado U-estadístico), utilizando la prueba de igualdad de las medias de grupo para evaluar si las variables introducidas tienen poder discriminante (Gálvez *et al.*, 2012b; Gramatica y Sangion, 2016). Asimismo se evaluó de manera sencilla la calidad de la FD aplicándola a un grupo externo; en este caso el grupo externo fueron los compuestos que no fueron incluidos en el análisis.

Una vez seleccionada la función discriminante (FD), se realiza el correspondiente diagrama de distribución de la actividad farmacológica (DDAF). Estos diagramas son gráficos de tipo histograma, útiles para establecer los intervalos de la función FD en donde la expectativa (o expectancia, E) o probabilidad de encontrar compuestos activos es máxima. En los DDAF, E se ubica en el eje de las ordenadas y los valores de la función discriminante (FD), en el eje de las abscisas. Para cada intervalo arbitrario de cualquier función, se define la expectancia de actividad como: $E_a = a/(i+1)$, donde “a” es el cociente entre el número de compuestos activos en el intervalo dividido por el número total de compuestos activos, e “i” es el número de compuestos inactivos en el intervalo dividido por el número total de compuestos inactivos. La expectativa de inactividad también se obtiene de la misma manera, como $E_i = i/(a+1)$. Esta representación nos permite ver las zonas en las que la superposición o solapamiento es mínima, así como determinar los intervalos de la función FD donde la probabilidad de encontrar nuevos compuestos activos es máxima en relación con la elección de un falso activo (Gálvez *et al.*, 1996).

Análisis de regresión multineal

El análisis de regresión multineal (ARML) busca una ecuación que correlacione un grupo de descriptores (variables independientes o explicativas) con la propiedad de interés o variable dependiente mediante un análisis de tipo cuantitativo (Cuadras, 2007). Para obtener la ecuación de regresión multineal se utilizaron los índices topológicos como variables independientes y la actividad antimalárica, IC_{50} , en su transformación logarítmica, $pIC_{50} = -\log IC_{50}$, como variable dependiente. El criterio para la selección de variables consistió en usar la agrupación con menor número de variables para evitar un sobreajuste, que el valor del coeficiente de correlación múltiple, R^2 , sea alto ($R^2 > 0,8$) y que el error estándar de estimación fuera mínimo.

Para validar la función de predicción se realizó un test de validación interna y un test de aleatoriedad. La validación interna se basa en una validación cruzada tipo *leave-one-out* (determinación del coeficiente de predicción Q^2), en la cual cada compuesto es eliminado del modelo y el valor de la actividad, pIC_{50} , se recalcula con los demás compuestos y descriptores de la ecuación seleccionada. El proceso se repite tantas veces como compuestos se estudian. Con los valores predichos se determina el valor del coeficiente de predicción, Q^2 , y se compara con el valor del R^2 . Valores de $Q^2 > 0,7$ nos indican que la función obtenida es predictiva y, por ende, que el modelo seleccionado es de buena calidad. El test de aleatoriedad se realizó para identificar una posible regresión de tipo fortuito, por ello los valores de la propiedad fueron intercambiados aleatoriamente entre los compuestos, recalculándose los valores de R^2 y Q^2 para cada uno. El modelo se considera estable si los valores de R^2 y Q^2 son muy bajos con respecto al modelo de regresión seleccionado.



RESULTADOS Y DISCUSIÓN

Un total de 35 compuestos fueron utilizados para construir un modelo topológico-matemático que contribuyera a predecir la actividad antimalárica. La elaboración del modelo fue abordada en dos etapas: la primera etapa se centró en la búsqueda de una función discriminante capaz de distinguir entre compuestos activos e inactivos, y la segunda obtuvo una función topológica capaz de predecir el valor de la actividad antimalárica en términos pIC_{50} . Estas dos funciones son la base del modelo matemático que va a contribuir a la búsqueda y selección de nuevos compuestos con potencial actividad antimalárica.

La función discriminante obtenida por medio del ALD fue aplicada a un total de 35 compuestos. Los compuestos pertenecientes al grupo activo tienen una $IC_{50} < 10\mu M$, mientras tanto los compuestos del grupo inactivo presentan una $IC_{50} > 10\mu M$.

La FD seleccionada fue:

$$FD = 171,72 - 3,21 G_2^v + 12,72 G_3 - 954,8 J_5 + 22,73 {}^2D - 121,11 {}^4C \quad \text{Ec. 1}$$

$$N = 35 \quad F = 5,26 \quad \lambda \text{ (Lambda de Wilks)} = 0,305 \quad p = < 0,0000$$

Los índices topológicos que aparecen en la Ec. 1 son de carga, G_3 , G_2^v y J_5 , que tienen en cuenta la transferencia de carga intramolecular a diferentes distancias topológicas, y de conectividad molecular, 2D y 4C_p , que informan sobre el ensamblaje topológico de la molécula.

A partir de la función discriminante, un determinado compuesto será seleccionado como un potencial antimalárico si la función $FD > 0$, e inactivo si la función $FD < 0$. Los parámetros estadísticos que se presentan en la Ec. 1 nos señalan una alta significación estadística del modelo $p < 0,00001$, así como los demás parámetros que lo conforman.

La matriz de clasificación clasificó correctamente en el grupo activo al 94,1 % de los compuestos (16 de 17) y al 93,3 % (14 de 15) en el inactivo. En total, el modelo clasificó correctamente la actividad del 93,7 % de compuestos.

La clasificación de cada compuesto, junto con el valor numérico de la función FD y la probabilidad de actividad, se ilustran en la tabla 2.

Tabla 2. Resultados de la clasificación obtenida del análisis lineal discriminante y regresión multilínea para cada compuesto analizado

Comp. N.º	$IC_{50} \text{ exp}^a$ (μM)	$pIC_{50} \text{ Exp.}$	Clas. Exp.	FD	Prob. (Activo)	Clas. ^b Cal.	$pIC_{50} \text{ Cal.}$	$pIC_{50} \text{ Cal(vc)}$
3	94,5	4,0	I	-7,79	0,00	I	3,9	3,8
4	60,3	4,2	I	-10,62	0,00	I	4,4	4,4
5	63,4	4,2	I	-6,95	0,00	I	4,0	4,0
6	21,0	4,7	I	-3,43	0,03	I	4,8	4,8
8	28,0	4,6	I	-1,34	0,20	I	4,3	4,3
9	13,3	4,9	I	-1,03	0,25	I	4,8	4,8
10	11,8	4,9	I	-2,37	0,08	I	5,3	5,3
11	49,8	4,3	I	-8,13	0,00	I	4,4	4,4
12	86,7	4,1	I	-5,16	0,00	I	4,6	4,7
14	11,0	5,0	I	-4,66	0,00	I	4,9	4,9



Comp. N.º	IC ₅₀ exp ^a (µM)	pIC ₅₀ Exp.	Clas. Exp.	FD	Prob. (Activo)	Clas. ^b Cal.	pIC ₅₀ Cal.	pIC ₅₀ Cal(vc)
16	85,0	4,1	I	-3,75	0,02	I	3,8	3,7
17	49,8	4,3	I	-3,79	0,02	I	4,6	4,7
18	17,0	4,8	I	-1,29	0,21	I	4,8	4,8
19	9,9	5,0	A	2,89	0,94	A	4,8	4,8
20	8,4	5,1	A	6,13	0,99	A	4,7	4,7
21	9,0	5,0	A	8,42	1,00	A	5,0	5,0
22	1,0	6,0	A	-0,33	0,41	I	5,3	5,3
23	12,8	4,9	I	0,63	0,64	NC	5,0	5,0
24	9,1	5,0	A	5,99	0,99	A	5,1	5,1
25	3,6	5,4	A	3,49	0,96	A	5,2	5,1
26	5,6	5,3	A	2,92	0,94	A	5,1	5,1
27	8,9	5,1	A	8,23	1,00	A	5,3	5,4
28	4,4	5,4	A	2,47	0,92	A	5,4	5,4
29	7,2	5,1	A	2,80	0,94	A	5,0	5,0
30	4,1	5,4	A	1,22	0,76	A	5,5	5,5
31	13,4	4,9	I	-2,08	0,10	I	5,1	5,1
32	5,0	5,3	A	4,36	0,98	A	5,3	5,3
33	1,9	5,7	A	9,29	1,00	A	5,7	5,7
34	1,9	5,7	A	5,99	0,99	A	5,7	5,7
35	6,7	5,2	A	3,50	0,97	A	5,5	5,6
36	8,9	5,1	A	1,57	0,82	A	4,9	4,8
37	3,7	5,4	A	6,29	0,99	A	5,2	5,2
7	113,5	3,95	I	-0,69	0,32	I	4,34	4,39
13	130,0	3,89	I	-1,08	0,25	I	3,81	3,79
15	234,0	3,63	I	0,97	0,720	NC	3,92	4,08

^a Valor de IC₅₀ (µM) obtenido de Souard *et al.*^b Obtenido de la Ec. 1

A: activo; I: inactivo; NC: compuesto no clasificado



En la figura 1 se observa el diagrama de distribución de la actividad antimalárica obtenida con la función FD (las columnas negras y blancas representan al grupo activo e inactivo, respectivamente). El dominio de aplicabilidad de la función FD para seleccionar compuestos activos con teórica actividad antimalárica cubre el intervalo de la función FD entre 1 y 10, por lo tanto la expectancia o probabilidad de actividad antimalárica estará en este intervalo. Asimismo, un compuesto se considera no-clasificable si la función $FD < -12$, si $FD > 10$ o si FD está en el intervalo entre 0 y 1.

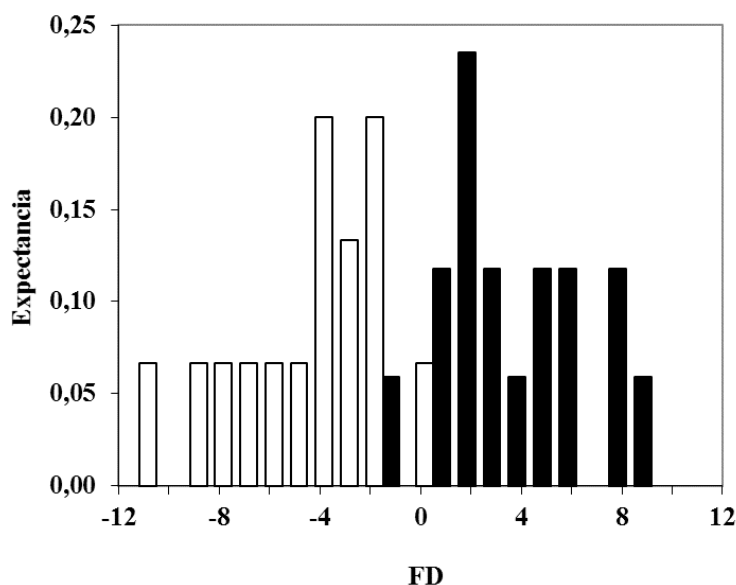


Fig. 1. Diagrama de distribución de la actividad antimalárica obtenida al graficar la expectancia (E) frente a la función discriminante FD (Ec.1). [Las barras de color negro representa el grupo activo y las de color blanco el grupo inactivo].

Además, una forma sencilla de evaluar la calidad de cualquier función discriminante es aplicarla a un grupo test externo. En nuestro caso utilizamos 3 compuestos (n.º 7, 13 y 15) que no fueron incluidos en el análisis lineal discriminante, con valores expresados en $IC_{50} > 100\mu M$. La tabla 2 nos muestra los resultados del análisis para cada compuesto. Como puede observarse, dos de los compuestos fueron clasificados correctamente como inactivos ($FD < 0$), mientras que el compuesto n.º 15 no pudo clasificarse como activo ya que FD no cubre el intervalo entre 1 y 10. Esto demuestra que nuestro modelo es muy específico, es decir, es eficiente para detectar compuestos inactivos, lo que minimiza el riesgo de falsa actividad.

Los parámetros estadísticos y la predicción de la actividad de cada compuesto se obtuvieron del análisis de regresión multilíneal (ARML), donde la variable dependiente está expresada en términos pIC_{50} y las variables independientes son los índices topológicos.

La función de regresión multilíneal seleccionada fue:

$$pIC_{50} = 10,747 - 0,587 {}^0\chi^v + 1,558 G_3 - 4,605 {}^4C_p + 0,357 L - 0,354 W \quad \text{Ec. 2}$$

$$N = 35 \quad R^2 = 0,830 \quad F = 5,26 \quad p = < 0,00000 \quad EEE = 0,259 \quad Q^2 = 0,755$$

En este caso, además de índices topológicos de carga y de conectividad, G_3 , ${}^0\chi^v$ y 4C_p , aparece la longitud de la molécula, L, y el índice de Wiener.



Los valores de R^2 están por encima de 0,8 y valores de Q^2 por encima de 0,7, y los valores del error estándar de estimación EEE son bajos. Todo ello certifica la calidad del modelo predictivo.

La tabla 2 (columna 8) y la figura 2 muestran los valores de pIC_{50} predichos para cada compuesto analizado. Los valores del pIC_{50} predicho o calculado con la Ec. 2 son muy similares al pIC_{50} experimental, lo cual nos indica que la predicción es buena y que la ecuación de regresión nos permite diferenciar correctamente los compuestos más activos de los menos activos, seleccionando como más activos aquellos que tienen un valor de $pIC_{50} > 5$ ($IC_{50} < 10\mu M$). La predicción de pIC_{50} también es aceptable para los tres compuestos inactivos del grupo test externo.

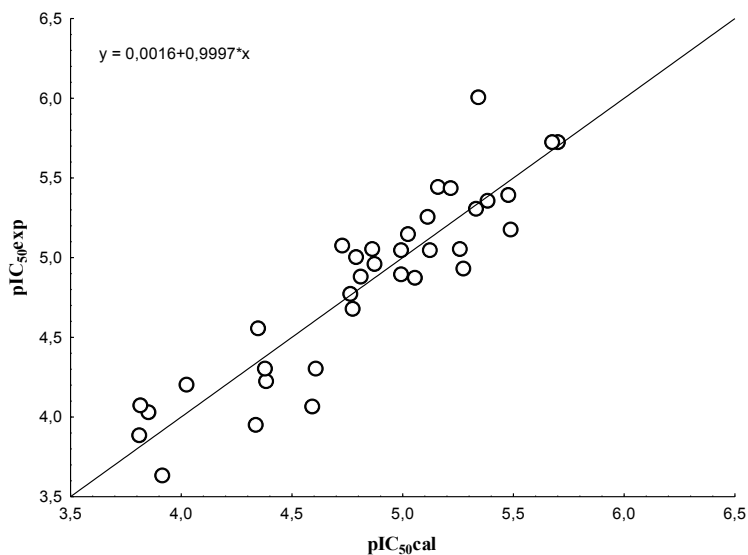


Fig. 2. Representación gráfica de los valores de experimental vs. el calculado de cada compuesto con la Ec. 2.

Como se mencionó anteriormente, el poder predictivo del modelo fue evaluado, como se deduce de los tests de validación interna (validación cruzada) y de aleatoriedad. En cuanto a la técnica de validación cruzada (*leave one out*), se obtuvo un coeficiente de predicción $Q^2 = 0,755$ para pIC_{50} , como se ilustra en la tabla 2 (columna 9). Los valores de Q^2 son solo ligeramente inferiores a R^2 , lo cual sugiere una aceptable capacidad predictiva de la Ec.2. La figura 3 nos muestra los resultados de los residuales obtenidos frente a los residuales de la validación cruzada, en la cual todos los compuestos siguen una dependencia lineal y no se observa la presencia de ningún *outlier*.

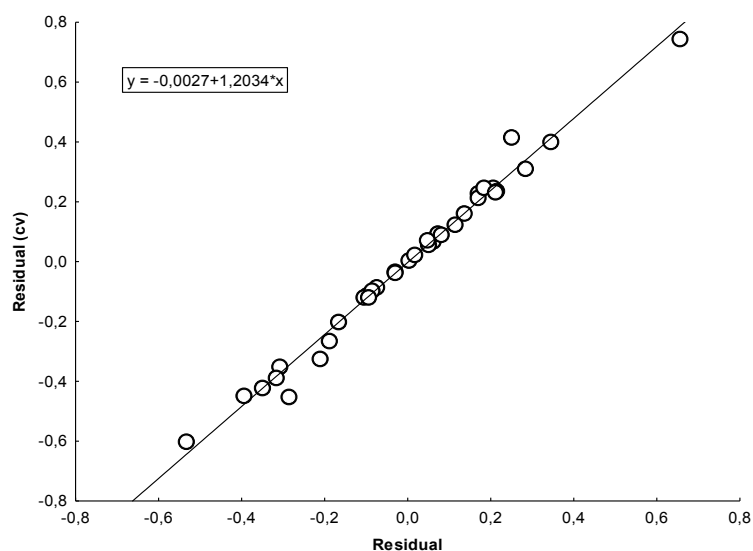


Fig. 3. Representación gráfica de los residuales obtenidos con cada compuesto *vs.* los residuales de la validación cruzada.

En el test de aleatoriedad se asignaron valores aleatorios a la variable dependiente y se calcularon los valores de R^2 y Q^2 . Los resultados se ilustran en la figura 4, donde se observa que las regresiones obtenidas con las propiedades asignadas aleatoriamente a los compuestos nos dieron valores de R^2 y Q^2 muy bajos, siendo solamente el modelo seleccionado (representado en la figura por un punto negro) el que obtiene valores de predicción óptimos.

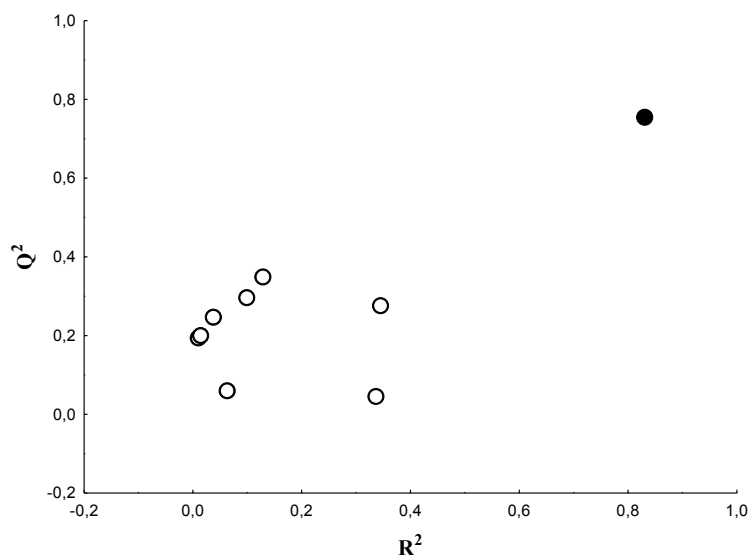


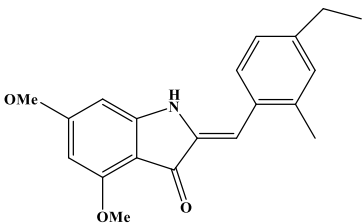
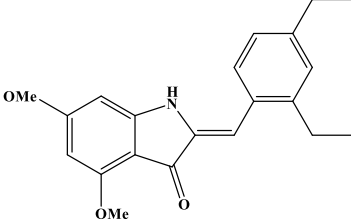
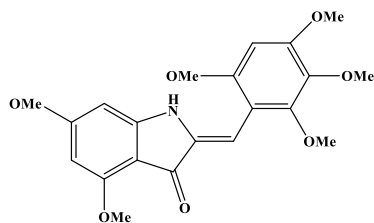
Fig. 4. Representación gráfica del coeficiente de predicción, Q^2 , *vs.* el coeficiente de correlación, R^2 , obtenidos del estudio de aleatoriedad del modelo de predicción del pIC_{50} . [Punto negro seleccionado].



Una vez obtenidos los modelos de predicción mediante el análisis lineal discriminante y la regresión multilíneal (Ec.1 y Ec.2), se plantea como objetivo aplicar estos modelos a la búsqueda de nuevos compuestos realizando pequeñas modificaciones sobre la estructura de partida, con el fin de predecir su valor de actividad teórica. El criterio de selección se basó en el diagrama de distribución farmacológica (DDF), seleccionando los compuestos con FD que se encuentren en el intervalo de 1 a 10 y cuyo valor calculado de pIC_{50} sea > 5 ($IC_{50} < 10\mu M$).

La tabla 3 muestra las estructuras químicas de tres moléculas y los valores predichos pIC_{50} . Las modificaciones se hicieron partiendo de la estructura base de los compuestos 26, 27 y 33, en los cuales se observó una mejora al sustituir grupos metilo por etilo. En la estructura 33 se adicionaron grupos metoxi.

Tabla 3. Resultados del cribado molecular con el modelo de predicción antimalárica

Estructura química		
		
FD = 3,1 pIC_{50} cal = 5,7	FD = 5 pIC_{50} cal = 5,8	FD = 8,4 pIC_{50} cal = 5,8

CONCLUSIONES

Los resultados obtenidos del análisis de 35 compuestos análogos de las auronas naturales (14 auronas y 21 azaauronas), todos derivados de 2-bencilidenbenzofuran-3 (2H) -onas, muestran que el análisis QSAR basado en la topología molecular es muy eficaz en la búsqueda de nuevos derivados antimaláricos. Dicho análisis se basó en dos modelos matemáticos, uno derivado del análisis lineal discriminante a nivel cualitativo (Ec. 1) y otro basado en la regresión multilíneal a nivel cuantitativo (Ec. 2). Por otra parte, el cribado virtual de las estructuras análogas demostró que pueden seleccionarse nuevos compuestos potencialmente activos con un nivel de eficacia similar o superior al obtenido con los compuestos de partida. Ello permite esperar que los modelos reseñados aquí podrán usarse en el futuro para el rastreo de bases moleculares que permitan localizar nuevos fármacos antimaláricos.

AGRADECIMIENTOS

Al Máster en Enfermedades Parasitarias Tropicales, del Departamento de Farmacia y Tecnología Farmacéutica y Parasitología de la Universitat de València, y a los profesores responsables de la cátedra de Química Física de dicha universidad, quienes nos transmitieron sus conocimientos de topología molecular y nos brindaron su apoyo para la elaboración del trabajo.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Achan, J., Talisuna, A. O., Erhart, A., Yeka, A., Tibenderana, J. K., Baliraine, F. N., D'Alessandro, U. (2011). Quinine, an old anti-malarial drug in a modern world: Role in the treatment of malaria, *Malaria Journal*, **10**, 144-2875-10-144. DOI:1475-2875-10-144 [pii].



- [2] Ashley, E. A., Dhorda, M., Fairhurst, R. M., Amaratunga, C., Lim, P., Suon, S. Tracking Resistance to Artemisinin Collaboration (TRAC). (2014). Spread of artemisinin resistance in plasmodium falciparum malaria, *The New England Journal of Medicine*, **371**(5) 411-423. DOI:10.1056/NEJMoa1314981.
- [3] Statistica 8.0 (2008) Statsoft Inc Retrieved from <http://www.statsoft.com/Textbook>.
- [4] Bruno-Blanch, L., Galvez, J., García-Doménech, R. (2003). Topological virtual screening: A way to find new anticonvulsant drugs from chemical diversity, *Bioorganic y Medicinal Chemistry Letters*, **13**(16) 2749-2754.
- [5] Carrasco, M. P., Machado, M., Goncalves, L., Sharma, M., Gut, J., Lukens, A. K., Moreira, R. (2016). Probing the azaaurone scaffold against the hepatic and erythrocytic stages of malaria parasites, *ChemMedChem*, **11**(19), 2194-2204. DOI:10.1002/cmdc.201600327.
- [6] Carrasco, M. P., Newton, A. S., Gonçalves, L., Góis, A., Machado, M., Gut, J., Moreira, R. (2014). Probing the aurone scaffold against plasmodium falciparum: Design, synthesis and antimalarial activity, *European Journal of Medicinal Chemistry*, **80**, 523-534. DOI: <http://dx.DOI.org/10.1016/j.ejmech.2014.04.076>.
- [7] Cuadras, C. M. (2007). *Nuevos métodos de análisis multivariante*, CMC Editions.
- [8] DE JULIÁN-ORTIZ, J. V., GÁLVEZ, J., MUÑOZ-COLLADO, C., GARCÍA-DOMENECH, R., GIMENO-CARDONA, C. (1999). Virtual combinatorial syntheses and computational screening of new potential anti-herpes compounds 1, *Journal of Medicinal Chemistry*, **42**(17), 3308-3314.
- [9] DesMol2 Software. Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular. Facultad de Farmacia, Universitat de València, Spain. Retrieved from <http://ahila2017.uv.es/desmol2/>.
- [10] Duart, M., Garcia-Domenech, R., Anton-Fos, G., Galvez, J. (2001). Optimization of a mathematical topological pattern for the prediction of antihistaminic activity, *Journal of Computer-Aided Molecular Design*, **15**(6), 561-572.
- [11] Gálvez, J., García-Domenech, R., de Gregorio Alapont, C., de Julián-Ortiz, J., Popa, L. (1996). Pharmacological distribution diagrams: A tool for de novo drug design, *Journal of Molecular Graphics*, **14**(5), 272-276.
- [12] Galvez, J., Galvez-Llompарт, M., Garcia-Domenech, R. (2012a). Introduction to molecular topology: Basic concepts and application to drug design, *Current Computer-Aided Drug Design*, **8**(3), 196-223. DOI:CCADD-EPUB-20120625-2 [pii]
- [13] Galvez, J., Galvez-Llompарт, M., García-Doménech, R. (2012b). Molecular topology as a novel approach for drug discovery, *Expert Opinion on Drug Discovery*, **7**(2), 133-153. DOI:10.1517/17460441.2012.652083.
- [14] Galvez, J., Garcia, R., Salabert, M. T., Soler, R. (1994). Charge indexes. new topological descriptors, *Journal of Chemical Information and Computer Sciences*, **34**(3), 520-525. DOI:10.1021/ci00019a008.
- [15] Gálvez-Llompарт, M., Recio, M. C., García-Doménech, R. (2011). Topological virtual screening: A way to find new compounds active in ulcerative colitis by inhibiting NF- κ B, *Molecular Diversity*, **15**(4), 917-926.
- [16] García-Doménech, R., López-Peña, W., Sanchez-Perdomo, Y., Sanders, J. R., Sierra-Araujo, M. M., Zapata, C., Gálvez, J. (2008). Application of molecular topology to the prediction of the antimalarial activity of a group of uracil-based acyclic and deoxyuridine compounds, *International Journal of Pharmaceutics*, **363**(1), 78-84.
- [17] García-Doménech, R., Lopez-Pena, W., Sanchez-Perdomo, Y., Sanders, J. R., Sierra-Araujo, M. M., Zapata, C., Galvez, J. (2008). Application of molecular topology to the prediction of the antimalarial activity of a group of uracil-based acyclic and deoxyuridine compounds, *International Journal of Pharmaceutics*, **363**(1-2), 78-84. DOI:10.1016/j.ijpharm.2008.07.003.
- [18] GRAMATICA, P., SANGION, A. (2016). A historical excursus on the statistical validation parameters for QSAR models: A clarification concerning metrics and terminology, *Journal of Chemical Information and Modeling*, **56**(6), 1.127-1.131. DOI:10.1021/acs.jcim.6b00088.
- [19] Izenman, A. J. (2013). Linear discriminant analysis, *Modern multivariate statistical techniques*, Springer 237-280.
- [20] Kar, S., Roy, K. (2012). QSAR of phytochemicals for the design of better drugs, *Expert Opinion on Drug Discovery*, **7**(10), 877-902. DOI:10.1517/17460441.2012.716420.
- [21] Kier, L. B., Hall, L. H. (1986). *Molecular connectivity in structure-activity analysis*, Research Studies.
- [22] Lee, K., Jung, S., Naik, R., Cho, A. E. (2012). 3D-QSAR study of melanin inhibiting (S)-(-)-decursin and its analogues by pharmacophore mapping, *Bulletin of the Korean Chemical Society*, **33**(1), 149-152.
- [23] Mahmoudi, N., García-Doménech, R., Galvez, J., Farhati, K., Franetich, J. F., Sauerwein, R., Mazier, D. (2008). New active drugs against liver stages of plasmodium predicted by molecular topology, *Antimicrobial Agents and Chemotherapy*, **52**(4), 1215-1220. DOI:10.1128/AAC.01043-07.



- [24] Martínez, J. R. B., Ovando, C. D. T., Ovand, N. T., Llompart, M. G., Álvarez, J. G. (2015). Aplicación de topología molecular a la predicción de la actividad frente a trypanosoma brucei rhodesiense de compuestos bicíclicos derivados de la 4-metilpiperazina, *Nereis: Revista Iberoamericana Interdisciplinar De Métodos, Modelización Y Simulación*, (7), 27-38.
- [25] Martínez, M. Á R., López-Peñalver, R. J. S., Tomás, G. A., Álvarez, J. G., Llompart, M. G., Doménech, R. G. (2014). Aplicación de la topología molecular a la predicción y optimización de la actividad repelente de mosquitos de derivados de la N-acypiperidina, *Nereis: Revista Iberoamericana Interdisciplinar De Métodos, Modelización Y Simulación*, (6), 19-26.
- [26] Mishra, M., Mishra, V. K., Kashaw, V., Iyer, A. K., Kashaw, S. K. (2016). Comprehensive review on various strategies for antimalarial drug discovery, *European Journal of Medicinal Chemistry*, **125**, 1300-1320. DOI:S0223-5234(16)30963-1 [pii].
- [27] Muraliedharan, K. M., Avery, M. A. (2009). Progress in the development of peroxide-based anti-parasitic agents, *Drug Discovery Today*, **14**(15-16), 793-803. DOI:10.1016/j.drudis.2009.05.008.
- [28] Pastor, L., Garcia-Domenech, R., Galvez, J., Wolski, S., García, M. (1998). New antifungals selected by molecular topology, *Bioorganic y Medicinal Chemistry Letters*, **8**(18), 2577-2582.
- [29] Souard, F., Okombi, S., Beney, C., Chevalley, S., Valentin, A., Boumendjel, A. (2010). 1-azaurones derived from the naturally occurring aurones as potential antimalarial drugs, *Bioorganic y Medicinal Chemistry*, **18**(15), 5724-5731. DOI:10.1016/j.bmc.2010.06.008.
- [30] Wang, T., Wu, M. B., Lin, J. P., Yang, L. R. (2015). Quantitative structure-activity relationship: Promising advances in drug discovery platforms, *Expert Opinion on Drug Discovery*, **10**(12), 1283-1300. DOI:10.1517/17460441.2015.1083006.
- [31] White, N. (1999). Antimalarial drug resistance and combination chemotherapy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **354**(1.384), 739-749. DOI:10.1098/rstb.1999.0426. WHO | global technical strategy for malaria 2016–2030. Retrieved from <<http://www.who.int/malaria/publications/atoz/9789241564991/en/>>.
- [32] WHO | update on artemisinin and ACT resistance. Retrieved from <<http://www.who.int/malaria/publications/atoz/update-artemisinin-resistance-october2016/en/>>.
- [33] WHO | world malaria report 2015. Retrieved from <<http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/>>.
- [34] Wiener, H. (1947). Structural determination of paraffin boiling points, *Journal of the American Chemical Society*, **69**(1), 17-20. DOI:10.1021/ja01193a005.

