



# Predicting anti-*Trypanosoma brucei rhodesiense* activity of a group of Dicationic 3,5-diphenylisoxazols by Molecular Topology

**Title in Spanish:** *Predicción de la actividad anti-Trypanosoma brucei rhodesiense de un grupo de 3,5-Difenilisoxazoles dicationicos por medio de la Topología Molecular*

Karla Vasco-Aguas<sup>1</sup>, Elena Barrio-Miguel<sup>1</sup>, Carmen Ferrando-Hernández<sup>1</sup>, María Álvarez-Izquierdo<sup>1</sup>, Jorge Gálvez<sup>2</sup>, Ramón García-Domenech<sup>2\*</sup>

<sup>1</sup>Departamento de Farmacia y Tecnología Farmacéutica y Parasitología, Facultad de Farmacia, Universitat de Valencia. Avd. V.A Estellés. s/n, 46100-Burjassot. Valencia, España. <sup>2</sup>Departamento de Química Física, Facultad de Farmacia. Universitat de Valencia. Avd. V.A Estellés. s/n, 46100-Burjassot. Valencia, España.

**ABSTRACT:** Currently, there are more than 65 million people at risk of contracting sleeping sickness caused by *Trypanosoma brucei*. Available treatments for this parasitic disease are limited and have side effects. There is a necessity to investigate novel, effective compounds that also yield low in toxicity. Quantitative Structure-Activity Relationship models (QSAR) have been used in this study for predicting anti-*Trypanosoma brucei* activity and potency in a group of dicationic 3,5-Diphenylisoxazoles. Linear Discriminant Analysis (LDA) correctly classified anti-*Trypanosoma brucei* activity in 100 % of cases (sensitivity) and inactivity in 86.7 % (specificity). Theoretical bioactivity (pIC<sub>50</sub>) of the studied compounds was predicted using Multilinear Regression Analysis (MLRA), demonstrating a high correlation between *in vitro* and *in silico* results ( $r^2=0.82$ ). Topologic models obtained with LDA and MLRA were applied to the screening of a group of dicationic analogs which presented a high *in silico* activity. These compounds when selected for *in vitro* analysis could reduce both time and costs in future drug research.

**RESUMEN:** Actualmente más de 65 millones de personas se encuentran en riesgo de contraer la enfermedad del sueño producida por *Trypanosoma brucei*. Los tratamientos actuales para esta parasitosis son limitados y poseen efectos colaterales, por lo que es necesario buscar nuevos compuestos activos que sean potentes y de baja toxicidad. En este estudio se utilizaron los métodos de relación cuantitativa estructura-actividad (QSAR) para predecir potencia y actividad anti-*Trypanosoma brucei rhodesiense* de un grupo de 3,5-Difenilisoxazoles Dicationicos. Mediante el Análisis Lineal Discriminante (ALD) se clasificó correctamente la actividad anti-*Trypanosoma brucei* en el 100 % de los casos (sensibilidad) y la inactividad en el 86,7 % (especificidad). Adicionalmente, a través del análisis de regresión multilíneal (MLRA) se predijo la bioactividad teórica (pIC<sub>50</sub>) de los compuestos estudiados, mostrando una elevada correlación de los resultados obtenidos *in vitro* e *in silico* ( $r^2=0,82$ ). Finalmente, los modelos topológicos obtenidos en el ALD y MLRA fueron aplicados a un grupo de análogos dicationicos para cribado molecular que presentaron una teórica elevada actividad y podrían ser seleccionados para su futuro análisis *in vitro* ahorrando tiempo y costes en la búsqueda de futuros tratamientos.

\*Corresponding Author: ramon.garcia@uv.es

Received: April 12, 2017 Accepted: May 7, 2017

An Real Acad Farm Vol. 83, N° 2 (2017), pp. 241-250

Language of Manuscript: Spanish

## 1. INTRODUCCIÓN

La tripanosomiasis africana humana (TAH) o enfermedad del sueño, es una parasitosis producida por *Trypanosoma brucei* (*T. brucei gambiense* en el 98 % de los casos notificados y *T. brucei rhodesiense* en el porcentaje restante (1). Este protozoario es transmitido por las moscas del género *Glossina*, comúnmente denominadas moscas tse-tse, que se encuentran distribuidas en el continente africano (2–4). Treinta y seis países subsaharianos son considerados endémicos de una u otra forma de la enfermedad (3). En el 2014, la Organización Mundial de la Salud (OMS) reportó 3796

casos de la enfermedad y se estima que existen aproximadamente 65 millones de personas en riesgo (1).

Esta parasitosis tiene dos fases: la aguda donde el parásito se desarrolla en sangre y linfa, y la crónica en la que atraviesa la barrera hematoencefálica e invade el sistema nervioso central pudiendo causar la muerte (1).

A día de hoy, se cree posible llegar a la eliminación de la enfermedad del sueño debido a la disminución de un 73,4 % de los casos en el periodo que va desde 2001 a 2010 (2,4). Actualmente, el protocolo seguido para el tratamiento de la THA consiste en la administración de pentamidina o suramina en la primera etapa de la

enfermedad y de melarsoprol o eflornitina en la segunda (1). Todos ellos, especialmente los dos últimos, producen efectos colaterales no deseables que llevan a la búsqueda de nuevos principios activos contra *Trypanosoma brucei* (1). El tratamiento ideal debe ser de administración oral, efectivo frente a las dos formas de la enfermedad, de bajo costo para la población y de aplicación en ausencia de personal cualificado (2,3).

En los últimos años, varios candidatos a fármacos para el tratamiento de la THA han sido estudiados en fases clínicas. Uno de estos compuestos es el maleato de furamidina (pafuramidina), profármaco de la furamidina (análogo y alternativo a la pentamidina) (5,6). La pafuramidina es mucho menos activa *in vitro* que la furamidina pero resulta activa por administración oral. Debido a los buenos resultados obtenidos, este profármaco llegó a fase clínica III pero fue abandonado por producir hepatotoxicidad e insuficiencia renal (7).

La presente investigación se centra en el estudio de los análogos del 3,5-difenilisoaxazol, activos *in vitro* contra la TAH, donde el anillo central de la furamidina es sustituido por un isoxazol (8). El 3,5-Bis(4-amidinofenil)isoxazol es el compuesto principal de la serie ( $n^{\circ}$  3 en Tabla 1) y el más parecido estructuralmente a la furamidina. Otras tres diamidinas ( $n^{\circ}$  13,  $n^{\circ}$  22 y  $n^{\circ}$  32 en Tabla 1) son regioisómeros de  $n^{\circ}$  3 y presentan diferente orientación de los grupos amidina, lo cual se ha comprobado que es definitivo para la actividad anti-*Trypanosoma* (8).

La selección de moléculas potentes y activas frente a la TAH puede simplificarse a través de la aplicación de la topología molecular (TM), que se basa en la generación de modelos matemáticos mediante estudios de relación cuantitativa estructura-actividad (QSAR) (9,10). Dicho de otra forma, los estudios QSAR se basan en la suposición de que la estructura de la molécula define sus propiedades. Este método simboliza al compuesto como un grafo de hidrógenos suprimidos, definido por un conjunto de vértices (átomos) y aristas (enlaces) (11), que posteriormente es representado como una matriz que servirá para el cálculo de índices topológicos (IT). Los IT

empleados con mayor frecuencia son: el índice de Wiener, el índice de conectividad de Randic, y los índices de carga (11,12). El cálculo de estos IT tiene como principal ventaja, que sus valores son independientes de la conformación de la molécula en un determinado momento (13). A partir de los IT obtenidos, se realiza un análisis de regresión multilínea y un análisis lineal discriminante para obtener el nivel de bioactividad y una función discriminante capaz de clasificar una molécula dada como “buena” (activa) o “mala” (inactiva) respecto a cierta propiedad farmacológica. Una vez que se obtiene la ecuación matemática basada en estos IT para la predicción de una determinada propiedad, se lleva a cabo un cribado virtual o screening para rastrear e identificar nuevos compuestos activos (10). De hecho, la topología molecular se ha empleado en varios estudios relacionados con la identificación de nuevos compuestos anti *T. brucei* (13,14), anti-protozoáricos (15,16), anti-maláricos (17) y anti-microbianos (18).

Este estudio tiene como objetivo la predicción de la actividad anti-*Trypanosoma brucei rhodesiense* de un grupo de 3,5-Difenilisoxazoles dicationicos previamente testados *in vitro* por medio de los métodos QSAR utilizando el análisis lineal discriminante y el análisis de regresión multilínea. Por otro lado, se aplicó el modelo QSAR para testear un grupo de moléculas diseñadas *in silico* para seleccionar aquellos compuestos con alta bioactividad teórica.

## 2. MATERIAL Y MÉTODOS

### 2.1. Compuestos analizados

Se analizó un grupo de entrenamiento compuesto por 43 moléculas análogas del 3,5-difenilisoaxazol (Tabla 1) cuya actividad antiprotozoaria,  $IC_{50}$  (nM), fue ensayada previamente *in vitro* con *Trypanosoma brucei rhodesiense* (*T. brucei rhodesiense*) STIB900 (8). Las moléculas dicationicas derivadas del isoxazol, representadas en la Tabla 1, se diseñaron *in silico* con el programa ChemDraw® Professional 15.1 (PerkinElmer Inc.) para el posterior cálculo de índices topológicos.

**Tabla 1. Estructuras de los Derivados Cationicos del Difenilisoaxazol y su actividad anti-*T. brucei rhodesiense*  $IC_{50}$  (nM).**

Compuesto	X <sub>1</sub>	R <sub>1</sub>	R <sub>2</sub>	X <sub>2</sub>	$IC_{50}$ (nM) <i>T. brucei</i> <i>rhodesiense</i>
-----------	----------------	----------------	----------------	----------------	--

1	<i>p</i> -Am	H	H	H	1719
2	H	H	H	<i>p</i> -Am	1121
3	<i>p</i> -Am	H	H	<i>p</i> -Am	5
4	<i>p</i> -iPrAm	H	H	<i>p</i> -iPrAm	196
5	<i>p</i> -Im	H	H	<i>p</i> -Im	87
6	<i>p</i> -Am	NO <sub>2</sub>	H	<i>p</i> -Am	14
7	<i>p</i> -Am	Cl	H	<i>p</i> -Am	16
8	<i>p</i> -Am	OMe	H	<i>p</i> -Am	11
9	<i>p</i> -Am	H	NO <sub>2</sub>	<i>p</i> -Am	32
10	<i>p</i> -Am	H	Cl	<i>p</i> -Am	23
11	<i>p</i> -Am	H	OMe	<i>p</i> -Am	16
12	<i>p</i> -Am	OMe	OMe	<i>p</i> -Am	13
13	<i>p</i> -Am	H	H	<i>m</i> -Am	6
14	<i>p</i> -iPrAm	H	H	<i>m</i> -iPrAm	488
15	<i>p</i> -Im	H	H	<i>m</i> -Im	1554
16	<i>p</i> -Am	NO <sub>2</sub>	H	<i>m</i> -Am	9
17	<i>p</i> -Am	Cl	H	<i>m</i> -Am	6
18	<i>p</i> -Am	OMe	H	<i>m</i> -Am	7
19	<i>p</i> -Am	H	NO <sub>2</sub>	<i>m</i> -Am	51
20	<i>p</i> -Am	H	Cl	<i>m</i> -Am	21
21	<i>p</i> -Am	H	OMe	<i>m</i> -Am	12
22	<i>m</i> -Am	H	H	<i>p</i> -Am	4
23	<i>m</i> -iPrAm	H	H	<i>p</i> -iPrAm	316
24	<i>m</i> -Im	H	H	<i>p</i> -Im	974
25	<i>m</i> -Am	Cl	H	<i>p</i> -Am	21
26	<i>m</i> -Am	OMe	H	<i>p</i> -Am	4
27	<i>m</i> -Am	H	NO <sub>2</sub>	<i>p</i> -Am	19
28	<i>m</i> -Am	H	Cl	<i>p</i> -Am	12
29	<i>m</i> -Am	H	OMe	<i>p</i> -Am	6
30	<i>m</i> -Am	OMe	NO <sub>2</sub>	<i>p</i> -Am	34
31	<i>m</i> -Am	OMe	OMe	<i>p</i> -Am	6
32	<i>m</i> -Am	H	H	<i>m</i> -Am	29
33	<i>m</i> -iPrAm	H	H	<i>m</i> -iPrAm	4983
34	<i>m</i> -Im	H	H	<i>m</i> -Im	18454
35	<i>m</i> -Am	NO <sub>2</sub>	H	<i>m</i> -Am	1628
36	<i>m</i> -Am	Cl	H	<i>m</i> -Am	25
37	<i>m</i> -Am	OMe	H	<i>m</i> -Am	7
38	<i>m</i> -Am	H	NO <sub>2</sub>	<i>m</i> -Am	85
39	<i>m</i> -Am	H	Cl	<i>m</i> -Am	46
40	<i>m</i> -Am	H	OMe	<i>m</i> -Am	27
41	<i>m</i> -Am	OMe	OMe	<i>m</i> -Am	4
42	<i>m</i> -iPrAm	OMe	OMe	<i>m</i> -iPrAm	374
43	<i>m</i> -Im	OMe	OMe	<i>m</i> -Im	45

## 2.2. Descripciones moleculares

Se calcularon 62 descriptores topológicos con el programa Desmol2 disponible en línea (<http://ahila2017.uv.es/desmol2/index.xhtml>). Se obtuvo los índices de Randić–Kier–Hall hasta el cuarto orden ( ${}^m\chi_t$ ,

${}^m\chi_t^v$ )(19,20), índices topológicos de carga hasta el quinto orden ( $J_m$ ,  $G_m$ ,  $J_m^v$ ,  $G_m^v$ ) (12,21) y los cocientes y diferencias entre los índices de conectividad de valencia y no valencia ( ${}^mC_t = {}^m\chi_t / {}^m\chi_t^v$  y  ${}^mD_t = {}^m\chi_t - {}^m\chi_t^v$ ) descritos en la Tabla 2.

**Tabla 2. Descriptores utilizados en este estudio.**

Símbolo	Nombre	Definición	Ref.
${}^k\chi_t$ , $k=0-4$ $t=p,c,pc$	Índices derivados de índice de Randić de orden $k$ y tipo path (p), cluster (c) y path-cluster (pc)	${}^k\chi_t = \sum_{j=1}^{k n_t} \left( \prod_{i \in S_j} \delta_i \right)^{-1/2}$ Donde $\delta_i$ es el número de enlaces, $\sigma$ or $\pi$ , del átomo $i$ para átomos no hidrogenados. $S_j$ es la sub-estructura $j$ th de orden $k$ y tipo $t$ .	(20)
${}^k\chi_t^V$ , $k=0-4$ $t=p,c,pc$	Índices de Kier-Hall de orden $k$ y tipo path (p), cluster (c) y path-cluster (pc)	${}^k\chi_t^V = \sum_{j=1}^{k n_t} \left( \prod_{i \in S_j} \delta_i^V \right)^{-1/2}$ Donde $\delta_i^V$ es la valencia de Kier-Hall de los átomos. $S_j$ es la sub-estructura $j$ th de orden $k$ y tipo $t$ .	(20)
$G_k$ , $k=1-5$	Índices topológicos de carga de orden $k$	$G_k = \sum_{i=1}^{N-1} \sum_{j=i+1}^N  M_{ij} - M_{ji}  \Leftrightarrow \delta(k, D_{ij})$ Donde $M=AQ$ es el producto de la matriz de adyacencia y del inverso del cuadrado de la matriz de distancia para el grafo molecular de hidrógenos suprimidos. $D$ es la matriz de distancia. $\delta$ es el Kronecker delta	(12)
$G_k^V$ , $k=1-5$	Índices de valencia topológica de carga de orden $k$	$G_k^V = \sum_{i=1}^{N-1} \sum_{j=i+1}^N  M_{ij}^V - M_{ji}^V  \delta(k, D_{ij})$ Donde $M^V=A^VQ$ es el producto de la matriz de adyacencia de electronegatividad modificada y del inverso del cuadrado de la matriz de distancia para el grafo molecular de hidrógenos suprimidos. $D$ es la matriz de distancia. $\delta$ es el Kronecker delta	(12)
$J_k$ , $k=1-5$	Índices topológicos de carga normalizados de orden $k$	$J_k = \frac{G_k}{N-1}$	(12)
$J_k^V$ , $k=1-5$	Índices topológicos de carga de valencia normalizada de orden $k$	$J_k^V = \frac{G_k^V}{N-1}$	(12)
${}^kD_t$ , $k=0-4$ $t=p,c,pc$	Diferencias de conectividad de orden $k$ tipo path (p), cluster (c) y path-cluster (pc)	${}^kD_t = {}^k\chi_t - {}^k\chi_t^V$	(22)
${}^kC_t$ , $k=0-4$ $t=p,c,pc$	Cocientes de conectividad de orden $k$ y tipo path (p), cluster (c) y path-cluster (pc)	${}^kC_t = \frac{{}^k\chi_t}{{}^k\chi_t^V}$	(22)

### 2.3. Algoritmos QSAR

#### 2.3.a. Análisis lineal discriminante

Con el objetivo de identificar correctamente la actividad/inactividad anti-*T. brucei* (IC<sub>50</sub>) del subgrupo de 3,5-difenilisoxazoles dicatiónicos ( $n^\circ$  1-43) se realizó el análisis lineal discriminante (ALD) con el programa STATISTICA 8 (Statsoft Inc, Tulsa, Okla) (23). Se clasificó a los compuestos en dos grupos según su actividad anti-*T. brucei*: activos *in vitro* (IC<sub>50</sub><50nM) e inactivos *in vitro* (IC<sub>50</sub>>50nM). Se seleccionó un modelo de ALD estadísticamente significativo con capacidad discriminante medida con el coeficiente Lambda de Wilks. El diagrama de distribución de la actividad anti-*T. brucei* (PDD) se realizó con la función discriminante de cada compuesto, calculada con la ecuación de discriminación y la expectativa de actividad/inactividad. La expectativa de actividad se calculó con la ecuación ( $E_a = a/(i+1)$ ) y la de inactividad con ( $E_i = i/(a+1)$ ); donde "a" es la fracción de compuestos activos en el intervalo entre el número total de compuestos activos, e "i" es la fracción de compuestos inactivos en el intervalo entre el número total de compuestos inactivos.

#### 2.3.b. Análisis de regresión multilínea

Se realizó el análisis de regresión multilínea (MLRA) para predecir los valores de IC<sub>50</sub> *in silico*. Los valores experimentales de IC<sub>50</sub> se transformaron a su logaritmo negativo (pIC<sub>50</sub>) expresado en concentración molar. La ecuación de predicción se obtuvo correlacionando los valores experimentales (pIC<sub>50</sub> *in vitro*) con los índices topológicos en el programa STATISTICA 8 (Statsoft Inc, Tulsa, Okla)(23). La selección de la mejor ecuación se basó en el valor de regresión con el menor número de descriptores (24), significancia estadística ( $p \leq 0,05$ ) y error estándar de estimación inferior al 10 % de variabilidad de la actividad estudiada. Para validar el modelo de MLRA escogido se realizó un test interno y uno de aleatoriedad. El test de validación interna o cros-validación se basó en el algoritmo *leave-one-out* que consiste en eliminar un caso y realizar el análisis de regresión con los casos restantes y sus descriptores, para posteriormente predecir el valor de la propiedad del caso retirado. El procedimiento se repitió 43 veces (igual al número de casos) y se calculó el coeficiente de predicción que informa la calidad de la función seleccionada (25). El test de aleatoriedad se realizó intercambiando al azar los valores de pIC<sub>50</sub> entre los compuestos y realizando la regresión con los mismos descriptores del modelo. Este proceso se llevó a cabo 10 veces y se calculó los valores de  $r^2$  y  $q^2$  de cada uno para

luego compararlos con el seleccionado.

### 3. RESULTADOS Y DISCUSIÓN

En este estudio se predijo con elevada exactitud la actividad anti-*T. brucei* de 43 análogos del 3,5-difenilisoxazol, empleando métodos computacionales basados en relaciones cuantitativas de estructura-actividad (QSAR). En primer lugar, se seleccionó un modelo de ALD estadísticamente significativo para identificar la actividad/inactividad anti-*T. brucei* (IC<sub>50</sub>) del subgrupo de 43 compuestos análogos del isoxazol. Este modelo incluyó dos variables: IT de Kier de orden 4, tipo path (p) (<sup>4</sup>χ<sub>p</sub>) e IT de carga de orden 2 (G<sub>2</sub>) cuya ecuación discriminante para análogos del 3,5-difenilisoxazol se describe a continuación.

$$FD = -18,35 - 15,05 \chi_p + 5,10 G_2$$

$$n=43; \lambda = 0,274; F(40,2) = 52,96; p < 0,0001$$

Ecuación 1

Una vez que el modelo de ALD fue aplicado a los 43

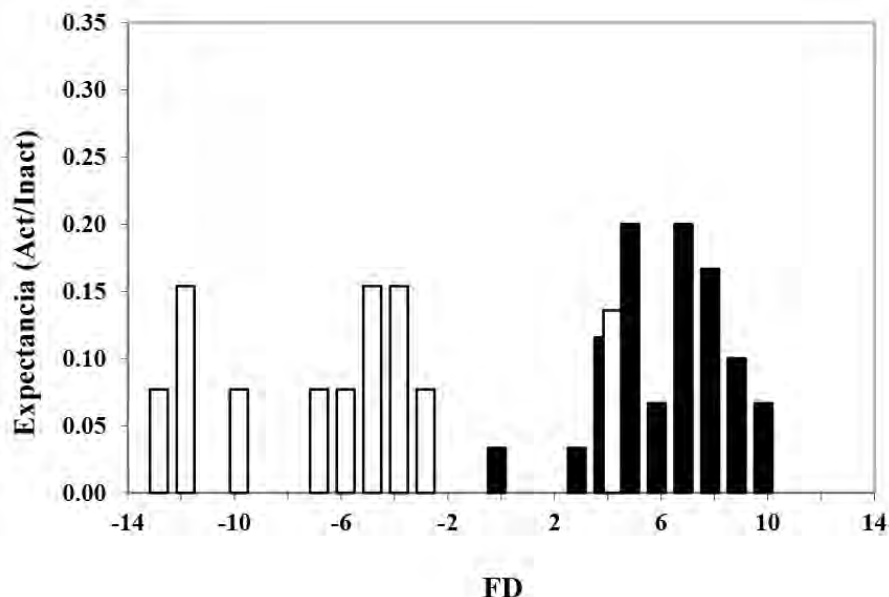
**Tabla 3. Resultados de la predicción obtenida por regresión multilínea y análisis lineal discriminante de la IC<sub>50</sub> de *T. brucei* en el grupo de entrenamiento.**

Compuesto	IC <sub>50</sub> (nM) <sup>a</sup> exp.	Clasif. exp. <sup>b</sup>	Probabilidad de Actividad	FD <sup>c</sup>	Clasif. calc. <sup>d</sup>	pIC <sub>50</sub> exp.	pIC <sub>50</sub> calc.	pIC <sub>50</sub> calc. <sup>e</sup>
1	1719	i	0,02	-4,0	i	5,77	5,61	5,54
2	1121	i	0,02	-4,0	i	5,95	5,61	5,46
3	5	A	1,00	5,9	A	8,29	8,32	8,32
4	196	i	0,00	-9,8	i	6,71	6,72	6,72
5	87	i	0,08	-2,5	i	7,06	6,70	6,60
6	14	A	1,00	5,7	A	7,86	7,69	7,68
7	16	A	1,00	9,3	A	7,81	8,15	8,18
8	11	A	1,00	7,2	A	7,98	8,03	8,03
9	32	A	1,00	5,9	A	7,49	7,77	7,79
10	23	A	1,00	9,5	A	7,64	8,32	8,41
11	16	A	1,00	7,3	A	7,79	8,13	8,15
12	13	A	1,00	8,6	A	7,89	8,04	8,05
13	6	A	0,99	4,6	A	8,20	8,08	8,07
14	488	i	0,00	-11,3	i	6,31	6,00	5,92
15	1554	i	0,02	-3,9	i	5,81	6,14	6,21
16	9	A	0,99	4,4	A	8,05	7,38	7,31
17	6	A	1,00	7,9	A	8,20	7,88	7,86
18	7	A	1,00	5,8	A	8,19	7,74	7,72
19	51	A	1,00	5,8	A	7,29	7,50	7,52
20	21	A	1,00	10,3	A	7,67	7,89	7,91
21	12	A	1,00	7,5	A	7,93	8,02	8,03
22	4	A	0,99	4,6	A	8,46	8,09	8,06
23	316	i	0,00	-11,3	i	6,50	6,01	5,90
24	974	i	0,02	-3,9	i	6,01	6,16	6,19
25	21	A	1,00	10,1	A	7,68	7,73	7,74
26	4	A	1,00	7,4	A	8,37	7,94	7,92
27	19	A	0,99	4,5	A	7,71	7,47	7,45
28	12	A	1,00	8,1	A	7,94	8,07	8,08
29	6	A	1,00	6,0	A	8,23	7,85	7,84
30	34	A	1,00	7,3	A	7,47	7,62	7,64
31	6	A	1,00	8,8	A	8,24	7,90	7,88
32	29	A	0,96	3,2	A	7,54	7,85	7,87
33	4983	i	0,00	-12,7	i	5,30	5,30	5,29

compuestos estudiados se observó una sensibilidad del 100 %, especificidad del 86,67 % y una exactitud del 95,56 %. De entre los 43 compuestos 30 fueron activos tanto *in vitro* como *in silico*; mientras que de los 13 compuestos restantes que fueron inactivos *in vitro*, 11 fueron inactivos *in silico* y dos erróneamente clasificados como activos (falsos activos). Asimismo, el diagrama de distribución de la actividad anti-*T. brucei* mostró una capacidad discriminante del modelo aceptable, donde los compuestos inactivos se agrupan entre valores de -2 y -13 y los activos entre 2 y 11 (Tabla 3). Dos de las moléculas inactivas fueron clasificadas en el rango de activos (n° 35 FD=4,27 y n° 38 FD=4,39) y una de las activas en un rango intermedio (n° 43 FD=0,46) (Tabla 3; Figura 1). La función discriminante clasifica a los compuestos con valores superiores a 0 como activos (A) y menores a 0 como inactivos (i), por lo que los resultados obtenidos en su mayoría concuerdan con la teoría.

34	18454	i	0,01	-5,4	i	4,73	5,60	5,79
35	1628	i	0,99	4,3	A	5,79	7,12	7,32
36	25	A	1,00	8,7	A	7,60	7,47	7,46
37	7	A	1,00	6,0	A	8,13	7,65	7,63
38	85	i	0,99	4,4	A	7,07	7,21	7,22
39	46	A	1,00	8,9	A	7,33	7,64	7,66
40	27	A	1,00	6,2	A	7,57	7,75	7,76
41	4	A	1,00	9,0	A	8,38	7,76	7,72
42	374	i	0,00	-6,9	i	6,43	6,80	6,96

<sup>a</sup> Valores de IC<sub>50</sub> obtenidos en estudio previo(8) <sup>b</sup>A=activos (IC<sub>50</sub><50nM); i=inactivos (IC<sub>50</sub>>50nM) <sup>c</sup>Función discriminante obtenida con la ecuación 1 <sup>d</sup>A=activos (FD>0); i=inactivos (FD<0) <sup>e</sup>Calculado con el test interno de cross-validación.



**Figura 1. Diagrama de la distribución farmacológica de la actividad anti-*T. brucei* mediante la relación entre expectativa (E) y función discriminante (las barras negras representan los compuestos con IC<sub>50</sub> <50 nM y las barras blancas los compuestos con IC<sub>50</sub> >50 nM).**

En segundo lugar, mediante el análisis de regresión multilíneal se seleccionó un modelo de predicción de pIC<sub>50</sub> estadísticamente significativo. Este modelo incluyó cuatro descriptores topológicos: índice de Kier-Hall de orden 3 y tipo cluster (<sup>3</sup>χ<sub>c</sub><sup>v</sup>), el IT de carga normalizado de orden 1 (J<sub>1</sub>), diferencias de conectividad de orden 4 tipo path (D<sup>4</sup>χ<sub>p</sub>) y el índice de Wiener (W), descritos en la Ecuación 2 de predicción del MLRA para análogos del 3,5-difenilisoxazol.

$$pIC_{50} = -1,09 - 16,66 {}^3\chi_c^v + 53,52 J_1 - 8,60 D^4\chi_p + 0,008 W$$

$$n=43; R^2=0,816; F(38,4)=42,2; p=0; EEE=0,42; Q^2=0,759$$

Ecuación 2

El modelo de predicción del MLRA obtuvo una correlación alta-positiva, entre los valores de pIC<sub>50</sub> calculados y experimentales, y la cross-validación arrojó un valor de q<sup>2</sup>=0,759, que muestra la alta capacidad predictiva de los valores de pIC<sub>50</sub> observados en la Tabla 3 y en la Figura 2.

Por otro lado, el test de aleatoriedad confirma que el modelo responde a un patrón topológico y no es de tipo estocástico, debido a los bajos valores de r<sup>2</sup> y q<sup>2</sup> que se obtienen al modificar los valores de pIC<sub>50</sub> al azar representados en la Figura 3.

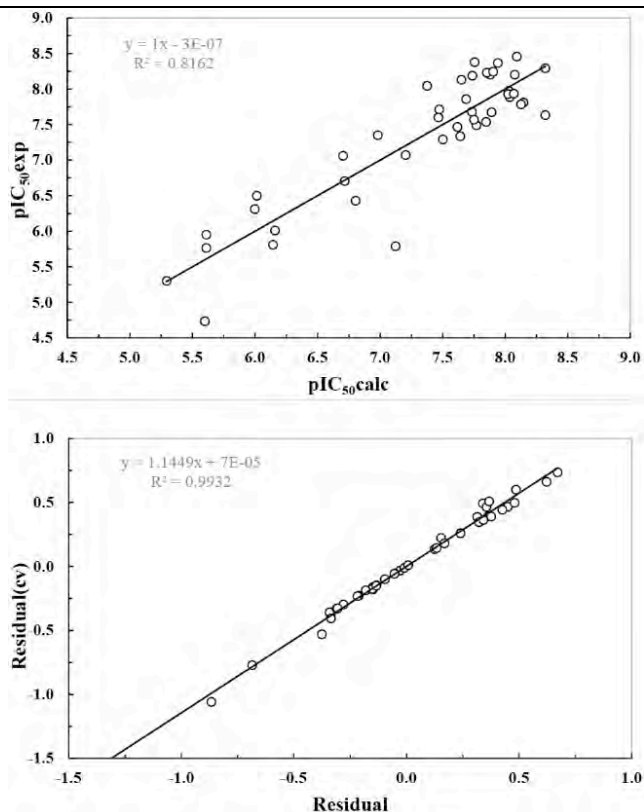


Figura 2. Predicción de inhibición del crecimiento de *T. brucei rhodesiense* pIC<sub>50</sub> para el grupo de derivados 3,5-Difenilisoxazoles Dicationicos. Arriba: representación gráfica de pIC<sub>50</sub> experimental frente al pIC<sub>50</sub> calculado a partir de la ecuación 2. Abajo: representación gráfica de los residuales obtenidos en la validación cruzada frente a los obtenidos en el grupo de entrenamiento.

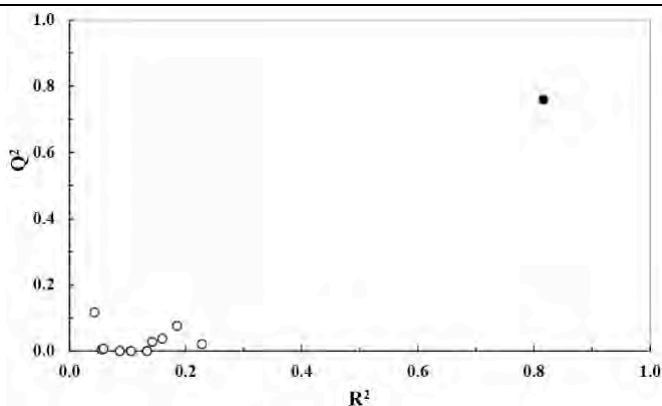
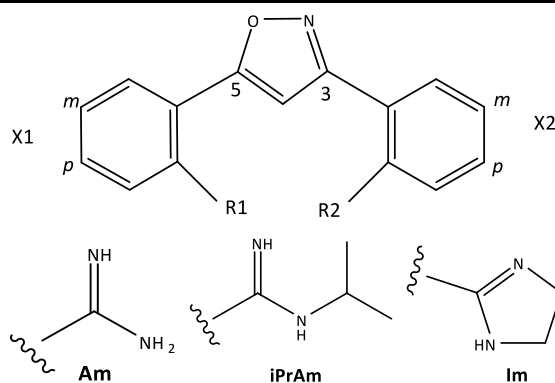


Figura 3. Test de aleatoriedad del grupo de entrenamiento manipulando los valores de pIC<sub>50</sub> y correlacionando los valores de  $r^2$  y  $q^2$ . El punto negro indica los valores de  $r^2$  y  $q^2$  obtenidos para el modelo seleccionado.

En consecuencia, los métodos QSAR para este grupo de compuestos dicationicos (grupo de entrenamiento) permiten discriminar de forma altamente fiable su actividad anti-*T. brucei* y su potencia en términos de pIC<sub>50</sub> ( $r^2=0,816$ ).

Finalmente, a partir de las moléculas con valores de IC<sub>50</sub> más bajos se diseñaron 28 análogos en los que se aplicaron los modelos QSAR adicionando dos radicales nuevos: Bromo (Br) y grupo etoxilo (OEt). El rango de la función discriminante de este grupo de compuestos para

cribado molecular comprendió valores entre 3 y 14, clasificándose todos ellos como activos, entre los cuales se destacan compuestos con elevada potencia teórica (Tabla 4). En cuanto al análisis de regresión multilíneal, este grupo de compuestos arrojó valores de entre 7,4 y 9,3 (Tabla 4), los cuales están en un rango similar al del grupo de entrenamiento donde los compuestos activos presentan valores de entre 7,3 y 8,5 y los inactivos entre 4,7 y 7,1 (Tabla 3).

**Tabla 4. Resultados de la predicción obtenida por regresión multilíneal y análisis lineal discriminante de la IC<sub>50</sub> de *T. brucei* en el grupo testado.****Estructuras base de los Derivados Catiónicos del Difenilisoxazol.**

Compuesto	X <sub>1</sub>	R <sub>1</sub>	R <sub>2</sub>	X <sub>2</sub>	FD	pIC <sub>50</sub> calc
1t	<i>m</i> -Am	Br	H	<i>p</i> -Am	10,1	7,78
2t	<i>m</i> -Am	OEt	H	<i>p</i> -Am	4,7	8,12
3t	<i>m</i> -Am	Br	OMe	<i>p</i> -Am	11,5	7,58
4t	<i>m</i> -Am	OEt	OMe	<i>p</i> -Am	6,1	8,28
5t	<i>m</i> -Am	Br	Cl	<i>p</i> -Am	13,6	7,73
6t	<i>m</i> -Am	OEt	Cl	<i>p</i> -Am	8,2	8,18
7t*	<i>m</i> -Am	H	Br	<i>p</i> -Am	9,1	9,31
8t	<i>m</i> -Am	H	OEt	<i>p</i> -Am	3,2	7,97
9t*	<i>m</i> -Am	OMe	Br	<i>p</i> -Am	10,9	8,59
10t	<i>m</i> -Am	OMe	OEt	<i>p</i> -Am	6,0	8,22
11t	<i>m</i> -Am	Cl	Br	<i>p</i> -Am	13,6	8,32
12t	<i>m</i> -Am	Cl	OEt	<i>p</i> -Am	8,7	7,69
13t	<i>m</i> -Am	OMe	OEt	<i>m</i> -Am	6,3	8,11
14t	<i>m</i> -Am	OMe	Br	<i>m</i> -Am	11,8	7,55
15t	<i>m</i> -Am	OEt	OMe	<i>m</i> -Am	6,3	8,11
16t	<i>m</i> -Am	Br	OMe	<i>m</i> -Am	11,7	7,45
17t*	<i>p</i> -Am	Br	H	<i>p</i> -Am	9,3	8,76
18t	<i>p</i> -Am	OEt	H	<i>p</i> -Am	4,4	8,16
19t*	<i>p</i> -Am	Br	OMe	<i>p</i> -Am	10,6	8,59
20t	<i>p</i> -Am	OEt	OMe	<i>p</i> -Am	5,8	8,37
21t*	<i>p</i> -Am	Br	Cl	<i>p</i> -Am	12,8	8,73
22t	<i>p</i> -Am	OEt	Cl	<i>p</i> -Am	7,9	8,25
23t*	<i>p</i> -Am	H	Br	<i>p</i> -Am	9,5	8,96
24t	<i>p</i> -Am	H	OEt	<i>p</i> -Am	4,6	8,26
25t*	<i>p</i> -Am	OMe	Br	<i>p</i> -Am	10,7	8,70
26t	<i>p</i> -Am	OMe	OEt	<i>p</i> -Am	5,8	8,37
27t*	<i>p</i> -Am	Cl	Br	<i>p</i> -Am	12,8	8,76
28t	<i>p</i> -Am	Cl	OEt	<i>p</i> -Am	7,9	8,18

\*Compuestos destacados por tener valores de pIC<sub>50</sub> ≥ 8,5, que representan valores de IC<sub>50</sub> ≤ 3nM (actividad elevada).

#### 4. CONCLUSIONES

Los modelos QSAR basados en la topología molecular empleados en este trabajo, han permitido predecir con elevada exactitud (95,6 %) la actividad anti-*Trypanosoma brucei rhodesiense* del grupo de 3,5-Difenilisoxazoles dicationicos analizado. El análisis discriminante predijo correctamente la actividad anti-*Trypanosoma brucei* en el 100 % de los casos y la inactividad en el 86,7 %. Adicionalmente, el análisis de regresión multilíneal mostró

una alta correlación entre los resultados de pIC<sub>50</sub> predichos y los obtenidos *in vitro*. Por último, la aplicación de los modelos permitió identificar un grupo de moléculas con alta bioactividad teórica, comprobada con el análisis discriminante y el de regresión multilíneal.

Este estudio muestra la utilidad de la topología molecular para predecir la actividad antiprotozoaria de nuevos compuestos químicos ahorrando tiempo y costes en la búsqueda de futuros tratamientos, ante la aparición de



resistencias y la necesidad de un espectro más amplio de fármacos que sean de fácil administración y baja toxicidad.

## 5. CONFLICTO DE INTERESES

Los autores declaran no tener ningún conflicto de intereses.

## 6. AGRADECIMIENTOS

Se agradece al Master de Enfermedades Parasitarias Tropicales de la Universitat de València por su apoyo en la realización de este trabajo.

## 7. REFERENCIAS

1. WHO: World Health Organization. La tripanosomiasis africana (enfermedad del sueño). 2016 [acceso 14 de diciembre de 2016]. Disponible en: <http://www.who.int/mediacentre/factsheets/fs259/es/>
2. Simarro PP, Jannin J, Cattand P. Eliminating human African trypanosomiasis: Where do we stand and what comes next? *PLoS Med.* 2008;5(2):0174–80.
3. Simarro PP, Diarra A, Postigo JAR, Franco JR, Jannin JG. The human african trypanosomiasis control and surveillance programme of the world health organization 2000-2009: The way forward. *PLoS Negl Trop Dis.* 2011;5(2).
4. Organización Mundial de la Salud. Relevé épidémiologique hebdomadaire. Weekly epidemiological record [revista en internet]. 2006 febrero. [acceso 14 de diciembre 2016];81(8). Disponible en:<http://www.who.int/wer/2006/wer8108.pdf>.
5. Silva MRS. Avaliação fenotípica in vitro de novas amidinas aromáticas sobre *Trypanosoma cruzi* [tesis doctoral]. Rio de Janeiro: Instituto Oswaldo Cruz, Fiocruz;2015.
6. da Silva CF. Atividade de amidinas aromáticas sobre *Trypanosoma Cruzi*: estudos in vitro e in vivo [tesis doctoral]. Rio de Janeiro: Instituto Oswaldo Cruz, Fiocruz; 2011.
7. Ríos-Martínez C. Síntesis y estudio de nuevos agentes antiprotozoarios derivados de compuestos dicationicos para el tratamiento de la tripanosomiasis africana humana [tesis doctoral]. Universidad Complutense de Madrid; 2014. [acceso 01 de abril de 2017] Available from: <http://eprints.ucm.es/25956/1/T35418.pdf>
8. Patrick DA, Bakunov SA, Bakunova SM, Kumar EVKS, Lombardy RJ, Jones SK, et al. Synthesis and in vitro antiprotozoal activities of dicationic 3,5-diphenylisoxazoles. *J Med Chem.* 2007;50(10):2468–85.
9. Amigó JM, Gálvez J, Villar VM. A review on molecular topology: applying graph theory to drug discovery and design. *Naturwissenschaften.* Springer; 2009;96(7):749–61.
10. Gálvez J, Gálvez-Llompарт M, García-Domenech R. Molecular topology as a novel approach for drug discovery. *Expert Opin Drug Discov.* Taylor & Francis; 2012;7(2):133–53.
11. Garcia-Domenech R, Galvez J, de Julian-Ortiz JV, Pogliani L. Some new trends in chemical graph theory. *Chem Rev* 2008 Mar;108(3):1127-1169
12. Galvez J, Garcia R, Salabert MT, Soler R. Charge indexes. New topological descriptors. *J Chem Inf Comput Sci.* ACS Publications; 1994; 34(3):520–5.
13. Bretón JR, Torrico C, Torrico N, Gálvez-Llompарт M, Gálvez J, García-Domenech R. Aplicación de topología molecular a la predicción de la actividad frente a *Trypanosoma brucei rhodesiense* de compuestos bicíclicos derivados de la 4-metilpiperazina. *Nereis*, 2015; (7):27–38.
14. Fernández BG, Zanni R, Pellicer M, Llompарт MG, Domenech RG. Aplicación de la topología molecular en la búsqueda de nuevos compuestos derivados del 4-nitro-imidazol activos frente al *Trypanosoma brucei*. *An. R. Acad. Nac. Farm* 2012; 78(4).
15. López AD, Gómez II, Moreira LR, Álvarez JG, Domenech RG. Aplicación de la topología molecular en la predicción de la actividad antiprotozoaria de derivados del benzimidazol. *Nereis Rev Iberoam Interdiscip métodos, Model y simulación.* 2013;(5):9–18.
16. Jiménez Mateo A, Santano García R, Martín Sarmiento A, García Guidet L, Gálvez J, García-Domenech R. Application of molecular topology for predicting the leishmanicidal activity of a group of compounds derived from pyrrolo [1,2-  $\alpha$  ] quinoxaline. *An. Real Acad. Farm.* 2016;82(3):317-23.
17. Mahmoudi N, Garcia-Domenech R, Galvez J, Farhati K, Franetich JF, Sauerwein R, et al. New active drugs against liver stages of *Plasmodium* predicted by molecular topology. *Antimicrob Agents Chemother.* 2008;52(4):1215–20.
18. García-García Á, Gálvez J, de Julián-Ortiz J-V, García-Domenech R, Muñoz C, Guna R, et al. New agents active against *Mycobacterium avium* complex selected by molecular topology: a virtual screening method. *J Antimicrob Chemother.* 2004 Jan 1;53(1):65–73.
19. Kier LB, Murray WJ, Randić M, Hall LH. Molecular connectivity V: connectivity series concept applied to density. *J Pharm Sci.* Wiley Online Library; 1976;65(8):1226–30.
20. Kier LB, Hall LH. General definition of valence delta values for molecular connectivity. *J Pharm Sci.* Wiley Online Library; 1983;72(10):1170–3.
21. Galvez J, Garcia-Domenech R, De Julian-Ortiz V, Soler R. Topological approach to analgesia. *J Chem Inf Comput Sci.* ACS Publications; 1994;34(5):1198–203.
22. Galvez J, Garcia-Domenech R, de Julian-Ortiz J V, Soler R. Topological approach to drug design. *J Chem Inf Comput Sci.* United States; 1995;35(2):272–84.
23. StatSoft I. *Statistica* (data analysis software system). Version 8.0. 2007. 2010.

24. Hocking RR. Criteria for selection of a subset regression: which one should be used? *Technometrics*. Taylor & Francis; 1972;14(4):967–76.
25. Besalú E. Fast computation of cross-validated properties in full linear leave-many-out procedures. *J Math Chem*. Springer; 2001;29(3):191–204.