

Medicent Electrón. 2017 ene.-mar.;21(1)

UNIVERSIDAD CENTRAL DE LAS VILLAS  
«MARTA ABREU»  
SANTA CLARA, VILLA CLARA

## ARTÍCULO ORIGINAL

**Identificación de compuestos líderes con potencial actividad antileishmaniásica: una contribución a la farmacéutica médica**

**Identification of lead compounds with potential antileishmanial activity: a contribution to pharmaceutical medicine**

Naiví Flores Balmaseda<sup>1</sup>, Yovani Marrero Ponce<sup>2</sup>, Judith Louvina Sayer<sup>1</sup>, Dalton Sebastián Tocto Merchán<sup>3</sup>

1. Universidad Central de las Villas. Facultad de Química y Farmacia. Santa Clara, Villa Clara. Cuba. Correo electrónico: [nflores@uclv.edu.cu](mailto:nflores@uclv.edu.cu)
2. Universidad Tecnológica de Bolívar. Colombia.
3. Universidad de Ciencias Médicas Dr. Serafín Ruiz de Zárate Ruiz. Santa Clara, Villa Clara. Cuba.

### RESUMEN

**Introducción:** la leishmaniasis es una enfermedad provocada por protozoos del género *Leishmania*, cuyas repercusiones pueden ser fatales o provocar discapacidad. En el presente trabajo se identifican nuevos compuestos con potencial actividad antileishmaniásica empleando estudios *in silico*.

**Objetivo:** confeccionar una nueva base de datos de compuestos evaluados contra este parásito e identificar nuevos compuestos líderes con potencial actividad antileishmaniásica a través de estudios *in silico*.

**Métodos:** los compuestos incluidos se recopilaron de la literatura y bases de datos internacionales, tomando la concentración inhibitoria media como medida de clasificación en activos e inactivos. Se emplearon los softwares DRAGON y STATISTICA 8.0 para calcular los descriptores moleculares; el análisis de clúster y la distancia euclidiana se emplearon en la confección de las series de entrenamiento y predicción. Los modelos fueron validados, y se realizó una búsqueda de nuevos compuestos líderes a través del tamizaje virtual de bases de datos.

**Resultados:** se confeccionó una base de datos de 2 145 compuestos, que fue dividida aleatoriamente. Tras el cribado virtual, fueron identificados 316 compuestos como potencialmente antileishmaniásicos, entre estos, conocidos fármacos con otras propiedades farmacológicas y varios productos naturales con diversos usos.

**Conclusiones:** se confeccionó una base de datos de compuestos a los que se les ha evaluado experimentalmente su actividad antileishmaniásica, y se obtuvo un modelo que permite predecir la actividad de nuevos compuestos a través del cribado virtual.

*DeCS:* leishmaniasis, base de datos, preparaciones farmacéuticas.

## ABSTRACT

**Introduction:** leishmaniasis is a disease caused by protozoan parasites of the genus *Leishmania*; its repercussions can be fatal and cause disability. In the present work, some new compounds with potential antileishmanial activity are identified using in silico studies.

**Objective:** to create a new database of the evaluated compounds against this parasite and to identify new lead compounds with potential antileishmanial activity through in silico studies.

**Methods:** the included compounds were compiled from literature and international databases taking the minimum inhibitory concentration to classify them as active and inactive. Molecular descriptors were calculated using STATISTICA 8.0 and DRAGON software; cluster analysis and Euclidean distance were used for training series and time - series prediction. Models were validated and a search for new lead compounds was carried out through virtual screening of databases.

**Results:** a database of 2145 compounds randomly divided was made up. There were identified 316 potentially antileishmanial compounds after carrying out the virtual screening, among which, known drugs with other pharmacological properties and some natural products with diverse uses were found.

**Conclusions:** there was created a database of compounds in which the antileishmanial activity was evaluated experimentally, as well as, a model was obtained allowing us to predict the activity of new compounds through virtual screening.

*DeCS:* leishmaniasis, database, pharmaceutical preparations.

## INTRODUCCIÓN

La leishmaniasis es una enfermedad zoonótica desatendida, provocada por protozoos del género *Leishmania* y transmitida por la picadura de insectos flebótomos. Sus manifestaciones clínicas pueden ser objeto de estigma social, discapacidad y muerte prematura; es endémica en 88 países. Se estima que el número de personas en riesgo de infección asciende a aproximadamente a 350 millones y que al año aparecen cerca de 500 000 casos nuevos de leishmaniasis visceral y 1,5 millones de la cutáneo mucosa.<sup>1</sup> La Organización Mundial de la Salud (OMS) ha precisado que la incidencia del leishmaniasis ha aumentado desde los años 80, y que ha ganado una posición relevante en el mundo entre las causas de muerte por enfermedades infecciosas.<sup>1</sup>

Los antimoniales, que son los medicamentos más frecuentemente empleados contra la leishmaniasis, han estado en uso desde los años 20 y tienen una ventana terapéutica estrecha debido a su toxicidad.<sup>1</sup> Actualmente no existe ningún método profiláctico eficaz para prevenir esta enfermedad, y el arsenal farmacológico disponible para tratarla es limitado. Se ha desarrollado

resistencia a las distintas quimioterapias existentes que, además, presentan un alto índice de reacciones adversas graves, poseen un elevado costo y requieren, en su mayoría, de largos períodos de tratamiento. Dada la diversidad de especies del género *Leishmania*, las formas en que se presenta el parásito (promastigote y amastigote), su complejo ciclo de vida y la variedad de manifestaciones clínicas que presenta la enfermedad, se hace muy engorroso encontrar nuevos compuestos líderes efectivos.<sup>2,12</sup>

La consecuencia final es la necesidad de desarrollar nuevos fármacos con alto índice terapéutico, posibilidad de producción a bajo costo, baja inducción de resistencia, que exploten nuevas dianas, e idealmente con actividad adicional sobre otros parásitos, aparte de *Leishmania*.<sup>1</sup> Hasta el momento, no se han encontrado compuestos con la potencia, seguridad y selectividad deseadas; por ello, las investigaciones en esta rama de la ciencia continúan siendo un reto para la industria farmacéutica,<sup>1</sup> la cual desarrolla sus investigaciones, fundamentalmente, a través de los métodos tradicionales de «prueba-error». Por todo lo anteriormente mencionado, es necesario descubrir y caracterizar nuevas entidades moleculares con actividad antileishmaniásica para reducir al mínimo los efectos adversos presentes en los fármacos ya existentes.

El diseño y descubrimiento de fármacos asistido por computadoras ofrece una alternativa más económica, rápida y eficaz en este sentido. Los métodos *in silico* tienen potencial para mejorar significativamente el descubrimiento y posterior desarrollo de fármacos. Asimismo, las técnicas *in silico* (químio- y bioinformáticas) ahorran un promedio de 130 millones de USD por cada fármaco, y aceleran el proceso de selección de nuevos compuestos líderes bioactivos. Por este medio, varias técnicas computacionales son usadas para seleccionar un reducido número de compuestos potencialmente activos de una gran biblioteca química o combinatoria que esté disponible, con un consecuente ahorro de recursos de todo tipo, lo que aporta racionalidad al proceso de identificación de nuevos compuestos líderes.<sup>3</sup>

Actualmente, la Facultad de Química y Farmacia de la Universidad Central «Marta Abreu» de las Villas, realiza investigaciones dirigidas al desarrollo de nuevos principios activos para la industria farmacéutica en las ramas de la medicina veterinaria, humana y la sanidad vegetal. Dentro de ella, el Grupo de Diseño Molecular, Asistido por Computadora e Investigaciones Bioinformáticas (CAMD-BIR Unit, por sus siglas en inglés) ha presentado resultados relevantes en este campo empleando métodos computacionales, cuya aplicación en estudios de relación cuantitativa de estructura actividad (QSAR) ha permitido el desarrollo de nuevos métodos grafo-teóricos para el diseño molecular de compuestos orgánicos potencialmente bioactivos, entre los que se incluyen los antileishmaniásicos.<sup>1</sup>

Estos métodos llamados *in silico* engloban todas las técnicas asistidas por ordenadores usadas en el diseño, descubrimiento y optimización de compuestos con propiedades deseadas y han desempeñado un papel fundamental en el desarrollo de un número de fármacos que actualmente se encuentran en el mercado.<sup>4</sup>

Estos presupuestos confirman la necesidad de desarrollar el presente trabajo, con el objetivo de confeccionar una nueva base de datos de compuestos evaluados experimentalmente contra la leishmaniasis y obtener un modelo QSAR, que permita predecir la actividad antileishmaniásica de nuevos compuestos.

## MÉTODOS

La investigación fue desarrollada en el Grupo de Diseño Molecular Asistido por Computadora e Investigaciones Bioinformáticas (CAMD-BIR Unit), perteneciente al Departamento de Farmacia, Facultad de Química y Farmacia, de la Universidad Central de las Villas, durante el período comprendido entre enero del 2012 y mayo del 2014.

En primer lugar, se realizó una extensa revisión bibliográfica que permitió incluir en el presente estudio compuestos recopilados de la literatura internacional actualizada y de las bases de datos Martin Negwer Handbook<sup>5</sup> y Merk Index,<sup>6</sup> así como de los que aparecen en BioAssays del PubChem,<sup>7</sup> consistentes en varios ensayos primarios de la actividad antileishmaniásica; quedaron, finalmente, 2 145 compuestos evaluados experimentalmente para determinar su actividad antileishmaniásica. Los datos fueron publicados en 365 artículos científicos de investigación relacionados con la evaluación farmacológica de nuevos compuestos y varios ensayos antiparasitarios. Los cálculos se efectuaron en un ordenador (PC) Centrino Centrino Duo Core 1.66 GHz, y los datos fueron almacenados en matrices, utilizando el programa Microsoft Excel. El análisis estadístico se efectuó con el paquete de procesamiento *STATISTICA* 8.0 para Windows.<sup>8</sup> Todas las estructuras (2 145 compuestos en total) fueron representadas usando el módulo ChemDraw Ultra del software ChemOffice 2005 versión 9.0,<sup>9</sup> exportadas como ficheros .mol del tipo MDL y llevadas al Software DRAGON 5.5 Professional<sup>10</sup> donde se calcularon las diferentes familias de descriptores moleculares (DM) (0-2)D implementadas en él, que totalizaron un conjunto de más de 2000 DM.

En el presente trabajo se emplearon dos técnicas de reconocimiento de patrones: el análisis de conglomerados (AC) y el análisis de componentes principales (ACP). Con el fin de conformar las series de entrenamiento (SE) y de predicción (SP), les fue aplicado un AC del tipo k-MCA (por sus siglas en inglés, *k-Means Cluster Analysis* o Análisis de Conglomerados en *k-Medias*), con la distancia euclidiana como medida de disimilitud y agente disgregante/aglomerante, los cuales se distribuyeron, por separado, en varios conglomerados, para garantizar una buena representatividad química en cada caso (SE y SP). Los compuestos incluidos en la SP nunca se usaron en el desarrollo de los modelos cuantitativos. La SE quedó conformada por el 63,2 % del total de los compuestos y la SP por el 36,8 %.

En el cálculo de las variables a emplear, se utilizó el paquete de procesamiento *STATISTICA* 8.0 para Windows y la técnica *forward stepwise* o pasos hacia delante, cuyo algoritmo consiste en construir un modelo paso a paso, revisando todas las variables y evaluando cuál puede contribuir más a la discriminación entre grupos. Esta variable podrá ser incluida en el modelo. El algoritmo finaliza cuando entre las variables no seleccionadas, ninguna discrimina de forma significativa.

La ley de parsimonia (*lex parsimoniae*) o el rasurador de Occam: «*Numquam ponenda est pluralitas sine necessitate*» o «La pluralidad nunca se debe postular sin necesidad (de ello)»,<sup>3</sup> se tomó en cuenta a la hora de seleccionar el modelo. En este caso, se seleccionó el modelo con la significación estadística más alta y con la menor cantidad de parámetros o variables posible. La calidad estadística del modelo se determinó empleando como estadístico determinante el análisis de la Lambda de Wilks ( $\lambda$ ), que permite evaluar la hipótesis de que dos o más grupos provienen de poblaciones con las mismas medias para un conjunto de variables. La  $\lambda$  de Wilks toma valores en el rango de 0 (discriminación perfecta) a 1 (no discrimina). También se calculó el cuadrado de la distancia de Mahalanobis ( $D^2$ ) como criterio para determinar la existencia de dos poblaciones, la razón de Fisher ( $F_{exp} > F_{tab}$ ,  $\alpha = 0,05$ ) y el número de variables en cada ecuación. Para evaluar la habilidad (poder discriminante) de la función discriminante (FD) obtenida, se tuvieron en consideración, además, los porcentajes de buena clasificación en cada uno de los grupos y del modelo en general en la serie de entrenamiento (SE), mientras que la clasificación de los casos es desarrollada por medio de la probabilidad de clasificación posterior. Usando la  $D^2$  para la clasificación, se pueden obtener probabilidades; la probabilidad de que un caso pertenezca a un grupo particular es básicamente 'proporcional' a la  $D^2$  del caso al centroide del grupo. En resumen, la probabilidad posterior es la probabilidad basada en el conocimiento de los valores de otras variables, de que el respectivo caso o compuesto pertenece a un grupo en particular, dígame: es activo o inactivo contra el parásito.

Para acceder al poder predictivo del modelo QSAR de Análisis Discriminante Lineal (ADL) obtenido, se desarrollaron procedimientos de validación externa. La calidad del modelo también fue evaluada chequeando los principales parámetros estadísticos propuestos en la literatura para los modelos de clasificación: coeficiente de correlación de Matthews (C), exactitud (Qtotal), sensibilidad, especificidad y la proporción de falsos activos.

Finalmente, el modelo es empleado en el cribado virtual de bases de datos, con el propósito de identificar nuevos compuestos líderes, los cuales son sometidos a análisis. Se empleó la base de datos internacional Spectrum collection,<sup>11</sup> que consta de más de 2 000 compuestos biológicamente activos y diversamente estructurales, tanto de origen sintético como natural, así como las funciones obtenidas, utilizando para ello los softwares DRAGON Professional versión 5.5 -2007 y STATISTICA 8.0.

## RESULTADOS

Se construyó una nueva base de datos de compuestos evaluados experimentalmente contra *Leishmania*, que constó de 2 145 compuestos de composición muy heterogénea, de ellos 537 activos y 1 618 inactivos, tomando como criterio de actividad la concentración inhibitoria media (IC<sub>50</sub>), lo cual constituye algo novedoso, pues no se encontró en la literatura científica la confección de una base de datos similar, a partir de la evaluación de esta actividad. Esta base de datos posee un alto grado de variabilidad estructural, e incluye análogos de purinas, diamidinas aromáticas, péptidos, fluoroquinolonas, arilquinuclidinas, quinolinas, flavonas, dicationes aromáticos, saponinas, benzodiazepinas, chalconas, pirimidinas sustituidas, imidazo-compuestos, pronas, derivados de la THTT (tetrahydro tiadiazin-thiona), porfirinas catiónicas, derivados de quinazolina, diterpenos, taninos, entre otros; asimismo, podrá ser empleada para la modelación ser empleada para la modelación molecular y el diseño computacional de modelos capaces de evaluar compuestos y determinar su posible actividad contra parásitos de este género.

Con el ánimo de modelar la actividad antileishmaniásica, se empleó la técnica estadística ADL y los índices 0-2D del software DRAGON como descriptores moleculares. La mejor función de clasificación obtenida con el procedimiento 'forward y backward stepwise' fue la ecuación (Ec) 3.1:

$$\begin{aligned} \text{Class} = & 0.238 \text{ nO} + 0.909 \text{ nR10} - 0.998 \text{ MAXDP} + 19.884 \text{ PW3} - 20.354 \text{ X0Av} \\ & + 43.954 \text{ X3Av} - 3.7619 \text{ MATS2m} + 2.038 \text{ MATS6e} - 0.668 \text{ GATS8p} + 95.431 \text{ JGI4} \\ & - 1.263 \text{ nCONN} + 0.0067 \text{ ALOGP2} - 1.057 \text{ B02[O-O]} - 1.538 \text{ B04[O-S]} + 1.67 \quad (3.1) \end{aligned}$$

que presentó los siguientes parámetros estadísticos empleando 14 variables:

**N = 1357**  $\lambda = 0.60$   $D^2 = 3.32$  **F = 62.7** **Rcan = 0.63**  $\chi^2 = 678.45$  **p < 0.0001** donde, N es el número de compuestos,  $\lambda$  es el estadístico de Wilks,  $D^2$  es el cuadrado de la distancia de Mahalanobis, F es el estadístico de Fischer, Rcan: Raíz canónica,  $\chi^2$ : estadígrafo, y p: nivel de significación estadística de la técnica estadística ADL.

Una de las aplicaciones más importantes de un modelo QSAR es su uso en la predicción de la actividad de nuevas entidades químicas que, incluso, no tienen que poseer existencia física aún. En este sentido, uno de los criterios para la aceptación o no de un modelo QSAR, como los que se han desarrollado en este trabajo, están basados en los estadísticos de la serie de predicción externa. En esta serie, la Ec. 3.1 clasificó correctamente el 76,02 % y el 76,50 % de los compuestos como activos e inactivos, respectivamente, para un porcentaje de buena clasificación global de 76,40 %. Como puede observarse en esta serie, el modelo ajustado describió adecuadamente la actividad biológica de estos compuestos.

En la tabla 1 se resumen los resultados de la clasificación, obtenidos en el ajuste del modelo, mediante la serie de validación externa. El modelo obtenido describe, de una forma adecuada, la actividad, y se evidencia un adecuado porcentaje de falsos activos ( $Q = 82,76\%$ ).

**Tabla 1.** Resultados de la clasificación de los compuestos incluidos en las series de entrenamiento y de predicción.

Serie	Exactitud 'Q <sub>Total</sub> ' (%)	Sensibilidad (%)	Especificidad (%)	Razón de falsos positivos (%)
SE	82,76	63,98	82,51	17,15
SP	76,40	47,27	76,02	23,50

## DISCUSIÓN

El cribado virtual (*screening*) ha emergido como una interesante alternativa para la evaluación masiva de compuestos químicos.<sup>12</sup> Este proceso consiste en utilizar los modelos obtenidos en el cribado de grandes bases de datos, en virtud de encontrar un conjunto reducido de compuestos predichos con la actividad biológica deseada.

Utilizando la función obtenida, se desarrolló una búsqueda virtual de nuevos fármacos antileishmaniásicos. En este sentido, puede plantearse que existen dos enfoques diferentes para encontrar un nuevo compuesto con actividad leishmanicida mediante cribado virtual:

- 1) Diseño de nuevos compuestos para ser sintetizados en un laboratorio químico y, posteriormente, evaluarlos de forma experimental.
- 2) Encontrar en bases de datos químicas, compuestos conocidos con otras actividades o usos (fundamentalmente fármacos que son usados en la clínica), y probar su actividad antileishmaniásica experimentalmente.

Ambos enfoques son importantes y muy utilizados en la práctica farmacéutica: el primero permite diseñar nuevos compuestos líderes con la actividad deseada, pero ellos necesitan ser sintetizados en primer lugar, ser evaluados desde el punto de vista farmacológico y, finalmente, pasar a través de rigurosos ensayos toxicológicos y farmacodinámicos. Sin embargo, los compuestos seleccionados con el segundo enfoque ya tienen métodos de síntesis bien establecidos y, en muchos casos, su comportamiento toxicológico y farmacodinámico es bien conocido, sobre todo para el caso de compuestos comercializados como fármacos. Por esta razón, se seleccionó en este estudio, el segundo método de búsqueda para encontrar nuevos compuestos antileishmaniásicos.

En este artículo solo se han mostrado, a modo de ejemplo, los resultados obtenidos en el tamizaje virtual de la colección *Spectrum*, de la cual se pueden obtener cualesquiera de los 2 000 compuestos que incluye, para una futura evaluación experimental. El modelo clasificó como candidatos promisorios a un reducido grupo de 316 estructuras como posibles compuestos anti-*Leishmania*. Entre estos, se pueden encontrar conocidos fármacos con otras propiedades farmacológicas y varios productos naturales con diversos usos.

Estos resultados son unos de los más importantes criterios de validación del modelo desarrollado en este trabajo, pues estos son capaces de detectar una serie de compuestos como activos, a partir de miles de compuestos incluidos en las bases de datos.

## CONCLUSIONES

- Se confeccionó una base de datos que comprendió 2 145 compuestos a los que se les ha evaluado experimentalmente su actividad antileishmaniásica, de los cuales 537 son activos.
- Se obtuvo un modelo QSAR utilizando descriptores moleculares y análisis discriminante lineal, que permite predecir la actividad antileishmaniásica a través del cribado virtual de bases de datos.
- Se identificaron 316 estructuras como posibles compuestos antileishmaniásicos tras el cribado virtual de una base de datos internacional, entre los que se pueden encontrar conocidos fármacos con otras propiedades farmacológicas y varios productos naturales con diversos usos.

## REFERENCIAS BIBLIOGRÁFICAS

1. Informe final Segunda Reunión de Gerentes de Programas de leishmaniasis de Las Américas, 8-10 oct. 2013. Ciudad de Panamá: PANAFTOSA - OPS/OMS; 2014 jun.
2. Guía de Atención Integral de Leishmaniasis 2010 (Documento actualizado de Versión final Convenio 256/09) Convenio Cooperación Técnica No. 637 de 2009 Ministerio De La Protección Social. Washington, DC: Organización Panamericana de la Salud; jul. 2010.
3. Fina BL, Lombarte M, Rigalli A. Investigación de un fenómeno natural: ¿estudios *in vivo*, *in vitro* o *in silico*? Actual Osteol. 2013;9(3):294-9.
4. Negwer M. Organic-Chemical Drugs and their Synonyms. Berlin: Akademie 9th, extensively enl. ed. - Weinheim: Wiley-VCH; 2007.
5. Budavari S, O'Neill M, Smith A, Heckelman P, Kinneary J (Merck & Co., Inc.) The Merck Index: An Encyclopedia of Chemicals, Drugs & Biologicals. 12th. ed. New York: Chapman & Hall; 1997.
6. National Center for Biotechnology Information. Antileishmanial activity against antimony-resistant *Leishmania infantum* MCAN/BR/2002/BH400 promastigote infected in Swiss mouse peritoneal macrophages after 72 hrs. PubChem OPEN Chemistry Database. Bethesda, USA: U.S National Library of Medicine; 2014  
<https://pubchem.ncbi.nlm.nih.gov/bioassay/744771#section=Top>
7. StatSoft, Inc. (2007). STATISTICA (data analysis software system), version 8.0. USA: SAS Institute Inc.; 2008.
8. Perkin E. Software: ChemOffice v, 9.0. Reino Unido: Universidad de Cambridge Soft Corporation; 2003.
9. Todeschini R. Talete srl, DRAGON for Windows (Software for Molecular Descriptor Calculations) v, 5.5 [internet]. Milano (Italy): University of Milano-Bicocca P.za della Scienza; 2007 [citado 10 feb. 2016]. Disponible en: <http://www.talete.mi.it>
10. The Spectrum Collection [internet]. United States: Micro Source Discovery Systems, Inc.; 2012 [citado 21 mayo 2014]. Disponible en: <http://www.msdiscovery.com/spectrum.html>
11. Tropsha A. Best Practices for QSAR Model Development, Validation, and Exploitation. Mol Inf. 2010;29:476- 88.
12. World Health Organization. Sustaining the drive to overcome the global impact of neglected tropical diseases: second WHO report on neglected tropical diseases: summary. Geneva: WHO; 2013

Recibido: 25 de enero de 2016

Aprobado: 27 de mayo de 2016

Naiví Flores Balmaseda. Universidad Central de las Villas. Facultad de Química y Farmacia. Santa Clara, Villa Clara. Cuba. Correo electrónico: [nflores@uclv.edu.cu](mailto:nflores@uclv.edu.cu)