

UCLV
Universidad Central
"Marta Abreu" de Las Villas



FQF
Facultad de
Química y Farmacia

Departamento de Ciencias Farmacéuticas

TRABAJO DE DIPLOMA

Título: Empleo de estudios *"in silico"* en la identificación de nuevos compuestos con potencial actividad contra *Leishmania major*

Autor: José Ernesto González Sorí

Tutores: MSc. Naiví Flores Balmaseda

Dr. Juan Alberto Castillo Garit

Santa Clara Julio, 2019
Copyright©UCLV

UCLV
Universidad Central
"Marta Abreu" de Las Villas



FQF
Facultad de
Química y Farmacia

Pharmaceutical Sciences Department

DIPLOMA THESIS

Title: Use of "*in silico*" studies in identification of new compounds with potential activity against *Leishmania major*

Author: José Ernesto González Sorí

Thesis Director: MSc. Naiví Flores Balmaseda

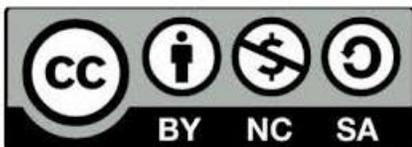
Dr. Juan Alberto Castillo Garit

Santa Clara July, 2019
Copyright©UCLV

Este documento es Propiedad Patrimonial de la Universidad Central “Marta Abreu” de Las Villas, y se encuentra depositado en los fondos de la Biblioteca Universitaria “Chiqui Gómez Lubian” subordinada a la Dirección de Información Científico Técnica de la mencionada casa de altos estudios.

Se autoriza su utilización bajo la licencia siguiente:

Atribución- No Comercial- Compartir Igual



Para cualquier información contacte con:

Dirección de Información Científico Técnica. Universidad Central “Marta Abreu” de Las Villas. Carretera a Camajuaní. Km 5½. Santa Clara. Villa Clara. Cuba. CP. 54 830

Teléfonos.: +53 01 42281503-1419

La ciencia no conoce país, porque el conocimiento pertenece a la humanidad, y es la antorcha que ilumina el mundo.

- Louis Pasteur

La tenacidad es, para muchas personas, una de las cualidades más valoradas de todas a la hora de lograr el éxito.

- Louis Pasteur



Dedicatoria

A mi familia y a todas las personas que de una manera u otra contribuyeron a mi formación profesional y a superarme cada día de este mundo. Siempre los recordaré...

Agradecimientos

Quiero darles un reconocimiento especial a mis padres por apoyarme en todo momento y estar ahí día a día.

Gracias a mis amigos, los de verdad, los que siempre estuvieron ahí, sin fallos en las fiestas y en los estudios.

A mis tutores por dedicar largas horas de atención y poder terminar a tiempo este trabajo, especialmente a Naivi, que tuvo que caerme atrás desde el primer día porque de otra manera no hubiera terminado nunca.

Gracias a mis compañeros de aula que compartimos cada seminario, el estudio para cada examen, las fiestas, las comidas entre amigos y todo lo que hicimos cada día a lo largo de estos 5 años. Especial agradecimiento a Yi y Lau, ustedes fueron con las que más recuerdo me quedo. Siempre estarán conmigo, imposible olvidarlas.

Gracias a los compañeros del cuarto por la convivencia pacífica que logramos tener y aguantar los desvelos con fines recreativos o de estudios.

Gracias a los profesores que me brindaron sus conocimientos y permitieron llegar hasta este momento.

A todos los que ya mencioné y a todos los que faltan: MIL GRACIAS

Resumen

La leishmaniasis agrupa un conjunto de enfermedades infecciosas causada por parásitos protozoarios del género *Leishmania*. Se han descrito alrededor de 30 especies, la mayoría patógenas para el hombre. La *Leishmania major* es la causa fundamental de leishmaniasis cutánea en regiones del norte de África, Este Medio, India Oriental y otros países del Medio Oriente siendo además la única especie detectada en Mali. Los antimonios pentavalentes constituyen el tratamiento más eficaz contra la leishmaniasis aunque poseen cierta toxicidad. El actual incremento en la resistencia a tratamientos convencionales e incidencia de casos hace necesario el descubrimiento de nuevos medicamentos más eficientes, efectivos y seguros. Los estudios *in silico* constituyen herramientas alternativas para dar solución a esta problemática. En el presente trabajo se emplean técnicas de Inteligencia Artificial para la obtención de modelos matemáticos capaces de predecir e identificar nuevos compuestos potencialmente antileishmaniásicos. Se confeccionó una base de datos de 120 compuestos con actividad comprobada contra amastigotes de *Leishmania major*. Se obtuvieron a través del análisis de conglomerados *k-medias*, 3 series: entrenamiento, predicción y validación externa. Se confeccionaron cinco modelos de clasificación empleando las técnicas IBk, MLP, J48 y SMO con porcentajes de clasificación entre 93.59% y 83.33%, cuyo poder predictivo se verificó a través de procedimientos de validación externa e interna. Se efectuó el cribado virtual del Cuadro Básico De Medicamentos de Cuba, compuestos de síntesis y origen natural a partir de un ensamblado de los modelos obtenidos identificándose de forma óptima 114 compuestos potencialmente activos contra amastigotes de *Leishmania major*.

Abstract

Leishmaniasis is a group of infectious diseases caused by protozoan parasites of the genus *Leishmania*. About 30 species have been described, most of them pathogenic to humans. *Leishmania major* is the fundamental cause of cutaneous leishmaniasis in regions of North Africa, Middle East, Eastern India and other Middle Eastern countries. It is also the only specie detected in Mali. Pentavalent antimonies are the most effective treatment against leishmaniasis although they have some adverse reactions. The current increase in resistance to conventional treatments and incidence of cases makes necessary the discovery of new drugs more efficient, effective and safe. *In silico* studies are alternative tools to solve this problem. In the present work, Artificial Intelligence techniques are used to obtain mathematical models capable of predicting and identifying new potentially antileishmaniasis compounds. A database of 120 compounds was created with proven activity against *Leishmania major* amastigotes. Three series were obtained through *k-medium* cluster analysis: training, prediction and external validation. Five classification models were made using IBk, MLP, J48 and SMO techniques with classification percentages between 93.59% and 83.33%, whose predictive power was verified through external and internal validation procedures. Virtual screening of the Basic Table of Drugs of Cuba, compounds of synthesis and natural origin was carried out from an assembly of the models obtained, identifying in an optimal way 114 compounds potentially active against amastigotes of *Leishmania major*.

Glosario

A	Número de átomos
AC	Análisis de conglomerados
ADL	Análisis Discriminante Lineal
AmB	Anfotericina B
ANN	Redes Neuronales Artificiales
C	Coeficiente de Correlación de Matthews
CAMD-BIR	Grupo de Diseño Molecular Asistido por Computadora e Investigaciones Bioinformáticas
DMS	Descriptores Moleculares
DMS-0D	Descriptores Constitucionales
DMS-1D	Descriptores Unidimensionales
DMS-2D	Descriptores Bidimensionales, Invariantes de Grafos
DMS-3D	Descriptores Tridimensionales
DMS-4D	Descriptores Tetradimensionales
FDA	(Por sus siglas en inglés <i>Food and Drug Administration</i>)
FP_r	Razón de Falsos Positivos
IA	Inteligencia Artificial
IBk	Técnica de <i>k</i> -vecino más cercano
IC₅₀	Concentración Inhibitoria media
IM	Indices Moleculares
J48	Técnica de Árbol de clasificación

Glosario

k-MCA	Análisis de conglomerados <i>k-medias</i>
KNN	<i>K-Nearest-Neighbours</i> (algoritmo del vecino más cercano)
L-AmB	Anfotericina B liposómica
LC	Leishmaniasis Cutánea
LCD	Leishmaniasis Cutáneo Difusa
LCL	Leishmaniasis Cutáneo Localizada
LMC	Leishmaniasis Mucocutánea
LV	Leishmaniasis Visceral
mg/kg	Miligramos/kilogramo
MLP	Técnica de Perceptrón Multicapa
OMS	Organización Mundial de la Salud
PKDL	Leishmaniasis Dérmica <i>Post-kala-azar</i>
Q	Exactitud o Porcentaje global de buena clasificación
QSAR	Relaciones Cuantitativas Estructura-Actividad
QSPR	Relaciones Cuantitativas Estructura-Propiedad
RN	Red Neuronal
ROC	Área Bajo la Curva
SbV	Antimoniales Pentavalentes
SE	Serie de Entrenamiento
SMO	Técnica de Optimización Secuencial Mínima
SP	Serie de Predicción
SVM	Técnica de Máquina vectorial de soporte

Glosario

VC	Validación Cruzada
VE	Validación Externa
VIH	Virus de Inmunodeficiencia Humana
VN	Verdaderos Negativos
VP	Verdaderos Positivos
WEKA	Software de entorno para análisis del conocimiento, de la Universidad de Waikato
µg /mL	Microgramos/mililitro

Contenido

Introducción.....	1
1 CAPÍTULO I. REVISIÓN BIBLIOGRÁFICA.....	4
1.1 Leishmaniasis	4
1.1.1 Generalidades	4
Taxonomía de la <i>Leishmania</i>	5
1.1.2 Ciclo biológico	6
1.1.3 Formas de leishmaniasis. Manifestaciones clínicas	7
1.1.4 Distribución geográfica	10
1.1.5 Co-infección <i>Leishmania</i> /VIH	10
1.1.6 Tratamiento y resistencia	13
1.1.7 Vacunas para tratar la leishmaniasis.....	17
1.2 <i>Leishmania major</i> . Generalidades	19
1.3 Enfoques para la búsqueda de medicamentos destinados a la leishmaniasis 22	
1.4 Estudios <i>in silico</i>	23
1.5 Estudios QSAR.....	24
1.6 Descriptores moleculares.....	26
1.7 Softwares empleados en los estudios QSAR.....	27
1.7.1 DRAGON.....	27
1.7.2 WEKA.....	28
1.7.3 ChemOffice	29
1.7.4 ChemAxon. Principales características y aplicaciones.....	30
1.7.5 I S I D A / Q S P R.....	32
1.8 Inteligencia Artificial	32

Tabla de Contenido

1.8.1	Aprendizaje o adquisición de conocimiento en la Inteligencia Artificial	33
1.9	Clasificación. Algoritmos de aprendizaje automático para clasificación...	34
1.9.1	kNN	35
1.9.2	Árboles de decisión	36
1.9.3	Máquinas vectoriales de soporte	37
1.9.4	Redes neuronales	38
2	CAPÍTULO II: MATERIALES Y MÉTODOS	39
2.1	Gestión de la data	39
2.2	Estrategias computacionales	40
2.3	Parametrización de la estructura química	41
2.4	Análisis estadístico de los resultados.....	42
2.5	Obtención y validación de los modelos	42
2.6	Cribado virtual y sistema multclasificador de ensamblado basado en modelos QSAR.....	44
3	CAPÍTULO III: RESULTADOS Y DISCUSIÓN.....	46
3.1	Gestión de la data.....	46
3.2	Parametrización de la estructura química. Estrategias computacionales. Análisis estadístico de datos	47
3.3	Validación de los modelos obtenidos	52
3.4	Cribado virtual de base de datos	53
4	Conclusiones	55
	Recomendaciones	62
5	Anexos.....	¡Error! Marcador no definido.
5.1	Base de datos de compuestos activos biológicamente comprobados frente amastigotes de <i>L. major</i>	¡Error! Marcador no definido.

Tabla de Contenido

- 5.2 Análisis de Conglomerados del tipo k -MCA **¡Error! Marcador no definido.**
- 5.3 Manifestaciones clínicas de la leishmaniasis 68

Introducción

La leishmaniasis es una enfermedad causada por protozoos parásitos del género *Leishmania* y se caracteriza por presentar una variedad clínica y epidemiológica en relación con las especies involucradas en la infección, los vectores transmisores, así como factores ligados al hospedero. El parásito cuenta con más de 30 especies diferentes, de las cuales alrededor de 20 especies son patógenas para el hombre y se distribuyen en distintas regiones del Viejo y Nuevo Mundos.⁽¹⁾ Según la Organización Mundial de la Salud, la enfermedad es endémica en 98 países de 4 continentes, 34 con registro de coinfección *Leishmania*-VIH. Anualmente se producen entre 700 000 y un millón de nuevos casos y cerca de 65 000 defunciones; la OMS la considera como una de las seis enfermedades tropicales de mayor importancia a nivel mundial.⁽²⁾

Existen dos formas del parásito: promastigote, extracelular, flagelado, predominante en el vector y el amastigote, intracelular, no flagelado, en el hospedador vertebrado.⁽³⁾ El vector responsable de la transmisión natural de *Leishmania* es la hembra hematófaga de insectos dípteros de los géneros *Phlebotomus* en el Viejo Mundo (Europa, norte de África, Medio Este y Asia), y *Lutzomyia* en el Nuevo Mundo (sur de Estados Unidos hasta el norte de Argentina).⁽⁴⁾ En Cuba no se reconocen vectores transmisores de *Leishmania*. En 1920 se describió para Cuba y las Islas Caimán la presencia de 2 especies de *Lutzomyia*: *Lu. Cubensis* y *Lu. orestes*. En la década de los 80 se notificaron 2 especies y una nueva subespecie del género para Cuba: (*Lu. novoae* y *Lu. diazi*) y (*Lu. Cayennensis cruz*); solo *Lu. orestes*, entre todas, es la que manifiesta hábitos antropofílicos y hematófagos. Al mismo tiempo, se realizó la caracterización biológica y ecológica de estos vectores, y se notificó su distribución en todas las regiones de la geografía cubana, aunque no parecen transmitir la enfermedad.⁽¹⁾

Las diferentes formas clínicas de la leishmaniasis se manifiestan en dependencia de la especie de *Leishmania* de que se trate y de la respuesta inmunitaria del hospedero a la infección (leishmaniasis cutánea, cutánea localizada, cutánea difusa, mucocutánea y visceral). Esta última es la forma más grave y resulta fatal en

Introducción

casi todos los casos si no se trata.⁽⁵⁾ Entre las especies de *Leishmania* se encuentra la *L. major*, la cual es considerada patógena para el hombre en el Viejo Mundo y la causa fundamental de leishmaniasis cutánea en regiones del norte de África, semidesiertos del Este Medio, India Oriental y Sudán, Pakistán y Sub-Sahara siendo además la única especie detectada en Mali.⁽⁶⁾ La identificación de compuestos con actividad antileishmaniásica son objeto de estudio a nivel mundial, pero dada la diversidad de especies del género, el complejo ciclo de vida del parásito y la variedad de manifestaciones clínicas que presenta la enfermedad se hace muy complejo hallar nuevos compuestos líderes.⁽⁷⁾ Existen varios medicamentos en el tratamiento contra la leishmaniasis cutánea, pero son accesibles o efectivos sólo a un porcentaje de la población afectada. Los niños presentan una constelación de desafíos: se ha demostrado una mayor patogenicidad de la infección por *Leishmania* en comparación con los adultos, son más propensos a tener lesiones en la zona de la cara, dependen de los adultos para tener acceso a la atención médica, la información farmacológica en este grupo etáreo es limitada, siendo además probable que la farmacocinética y farmacodinamia de los medicamentos contra la *Leishmania* sea distinta.⁽⁸⁾

Los compuestos antimoniales se han considerado como la primera línea de tratamiento para todas las especies de *Leishmania* por más de 60 años y continúan siendo el tratamiento más eficaz con el que se logran resultados clínicos y microbiológicos satisfactorios, pero el desarrollo de resistencia parasitaria ha provocado una disminución progresiva de su eficacia.⁽⁹⁾ La toxicidad de los medicamentos disponibles, unido al alto costo, la necesidad de hospitalización (debido a los efectos adversos y requerimientos de administración sistémica), disminución de la eficacia y la resistencia a los fármacos disponibles unido a la ubicación geográfica de Cuba y el principio internacionalista de nuestro sistema social constituyen aspectos por lo cual se hace necesario el descubrimiento y caracterización de nuevas entidades moleculares con actividad antileishmaniásica. El descubrimiento de medicamentos asistido por computadoras o también llamados estudios *in silico*, basado en relaciones cuantitativas de estructura actividad (modelos QSAR), resalta entre las opciones para el descubrimiento de nuevas

Introducción

alternativas terapéuticas antileishmaniásicas que permite un diseño de candidatos a nuevos fármacos mucho más económico y rápido. Por todo lo planteado anteriormente el presente proyecto de investigación estuvo guiado por el siguiente

problema científico:

Los fármacos empleados en la terapéutica actual para el tratamiento de las formas de leishmaniasis causadas por *Leishmania major* presentan una alta frecuencia de efectos citotóxicos, reacciones adversas severas y resistencia, por lo cual es necesario el descubrimiento de nuevas entidades moleculares activas contra este parásito.

Como vía para darle solución se formula la siguiente **hipótesis:**

Es posible la identificación de nuevos compuestos con potencial actividad contra *Leishmania major* mediante el empleo de métodos QSAR y técnicas de Inteligencia Artificial.

Teniendo en cuenta lo anterior, se propone como **objetivo general:**

Identificar nuevos compuestos con potencial actividad en el tratamiento de la leishmaniasis provocada por *L. major*.

y como **objetivos específicos:**

1. Confeccionar una base de datos de compuestos a los que se les ha evaluado experimentalmente su actividad contra amastigotes de *Leishmania major*.
2. Obtener modelos QSAR (predictivos) empleando técnicas de Inteligencia Artificial.
3. Validar la calidad estadística y predictiva de los modelos obtenidos a través de procesos de validación interna y externa.
4. Identificar nuevos compuestos con potencial actividad contra amastigotes de *Leishmania major* mediante el cribado virtual de bases de datos y compuestos de síntesis.

1 CAPÍTULO I. REVISIÓN BIBLIOGRÁFICA

1.1 Leishmaniasis

1.1.1 Generalidades

Leishmaniasis es un término que define un conjunto de enfermedades de muy variada presentación clínica, como la leishmaniasis dérmica ⁽¹⁰⁾, leishmaniasis mucocutánea (MCL), y kala-azar (VL) cuya característica común es ser producida por la infección con parásitos pertenecientes al género *Leishmania*.^(11, 12) Según la Organización Mundial de la Salud la enfermedad afecta a las poblaciones más pobres del planeta y está asociada a la malnutrición, los desplazamientos de población, las malas condiciones de vivienda, la debilidad del sistema inmunitario y la falta de recursos. También se vincula a cambios ambientales, como la deforestación, la construcción de presas, los sistemas de riego, la migración y la urbanización. La propia organización estimó que cada año se producen entre 700 000 y un millón de nuevos casos y hasta 65 000 defunciones, pero solo una parte de las personas infectadas acaban padeciendo la enfermedad.⁽¹³⁾

La *Leishmania* es un protozoo flagelado, parásito intracelular obligado de mamíferos que habita en vectores.⁽¹⁴⁾ Morfológicamente tiene dos formas: intracelular (amastigote) y extracelular (promastigote). Los amastigotes viven dentro de vacuolas lisosomales de las células fagocíticas, carecen de flagelo y presentan una forma ovalada o redondeada con dimensiones que oscilan entre 2-5 μm de longitud y 1,5-2,5 μm de ancho, presentan un núcleo voluminoso, generalmente excéntrico y adyacente a este se encuentra el kinetoplasto. El citoplasma es granular y vacuolado. El pH y temperatura óptimos para su desarrollo son de 4-5 y 37 ° C respectivamente. Bioquímicamente su principal fuente de energía son los ácidos grasos de cadena larga, los cuales puede transformar en CO₂ y agua en mayor grado que los promastigotes.^(15, 16) La forma promastigote es móvil, fusiforme, extracelular y presenta un tamaño mayor que la forma intracelular. Su tamaño oscila entre 10-30 μm de largo y 1,5-3 μm de ancho. Presenta núcleo oval central y posee un flagelo adherido a un organelo similar a la mitocondria, llamado kinetoplasto.

Revisión Bibliográfica

Metabólicamente utilizan aminoácidos y azúcares como fuente principal en condiciones aeróbicas. El pH es neutro y la temperatura de desarrollo varía de 25 a 27°C, pareciéndose a las condiciones biológicas del intestino del insecto.

El perro es considerado el principal reservorio doméstico de estos parásitos.⁽¹⁷⁾ La transmisión en la naturaleza, generalmente se produce a través de la picadura del flebótomo díptero de la familia *Psychodidae* y subfamilia *Phlebotominae*, del género *Phlebotomus* en Europa y *Lutzomyia* en América.^(14, 18, 19) La transmisión es de tipo zoonótica, es decir, establecerse entre un animal enfermo (reservorio), el flebótomo (vector) y la persona sana, o antroponótica si se efectúa entre una persona enferma, el flebótomo y otra persona sana.^(4, 20) Los miembros del género *Leishmania* (*L.*) son parásitos de mamíferos. Los dos subgéneros: *Leishmania* y *Viannia*, están separados en base a su localización en el intestino del vector. Inicialmente, la clasificación de las especies se basó en distintos criterios extrínsecos, como características biológicas, geográficas y clínicas; por ejemplo: *L. guyanensis* (aislada en Guayana), *L. peruviana* (aislada en Perú), *L. infantum* (aislada de un niño en Túnez) y *L. gerbilis* (aislada de un gerbo).⁽²¹⁾ La epidemiología de la leishmaniasis depende de las características de la especie del parásito, de las características ecológicas de los lugares de transmisión, de la exposición actual y pasada de la población humana al parásito, y del comportamiento humano, que es muy variable.⁽²¹⁾

Taxonomía de la *Leishmania*

La leishmaniasis es causada por parásitos intracelulares del orden *Kinetoplastida*, familia *Trypanosomatidae* y del género *Leishmania*. Su taxonomía es compleja y ha evolucionado a lo largo de los años. En el pasado, eran clasificadas según la localización geográfica y el aspecto clínico que producía la enfermedad. En la actualidad, se han descrito alrededor de 30 especies de *Leishmania* de las cuales alrededor de 20 son patógenas al hombre. Hoy en día las agrupaciones taxonómicas se basan en características bioquímicas y moleculares, por lo que ya no coinciden con los patrones clínicos específicos de la enfermedad. De esta manera, el género *Leishmania* se divide en dos subgéneros: *Leishmania* y *Viannia*. Cada subgénero se divide en complejos, por ejemplo: *Viannia* abarca dos

complejos: *L. braziliensis* y *L. guyanensis*. En el primero se encuentran las siguientes especies: *L. braziliensis*, *L. peruviana*, *L. colombiensis*, *L. lainsoni*, *L. shawi* y *L. naiffi* y en el segundo aparecen: *L. guyanensis* y *L. panamensis*. Además, en el subgénero *Viannia* existen dos especies híbridas: *L. braziliensis/L. panamensis* y *L. braziliensis/L. guyanensis*. En el subgénero *Leishmania* se encuentran cinco complejos: *L. major*, *L. tropica*, *L. aethiopica*, *L. donovani* y *L. mexicana*, formados por una a tres especies.⁽²²⁾

1.1.2 Odobiológico

La leishmaniasis es transmitida mayoritariamente por insectos vectores. No obstante, se ha descrito con menos frecuencia la transmisión por vía sexual, transfusional, trasplante de órganos, vía transplacentaria, inoculación accidental en el laboratorio y por contacto directo lesión - lesión. Los insectos pueden transmitirla de una persona infectada a una sana (transmisión antroponótica) o de un animal enfermo al hombre (transmisión antropozoonótica). Cuando la leishmaniasis es transmitida por la picadura de un insecto flebótomo hembra infectado (Figura 1). El vector inyecta el estadio infectante (promastigote) desde su proboscis durante la ingestión de sangre. Los promastigotes introducidos en la piel son fagocitados por los macrófagos y otros tipos de células fagocíticas mononucleares. Dentro de estas células, los promastigotes se transforman en amastigotes, los cuales se multiplican por fisión binaria y luego infectan otras células fagocíticas mononucleares. Factores dependientes del parásito y del hospedero determinan si la infección se hace sintomática y la forma de presentación de la leishmaniasis. Los flebótomos se contagian cuando ingieren células infectadas al alimentarse de sangre. En el intestino de los insectos flebótomos los amastigotes se transforman en promastigotes, se dividen (en el intestino posterior los del subgénero *Viannia* y en el intestino medio los del subgénero *Leishmania*) y finalmente migran hacia la proboscis. Desde aquí, los promastigotes podrán acceder a los hospederos vertebrados cuando nuevamente el insecto se alimente de sangre.⁽²³⁾

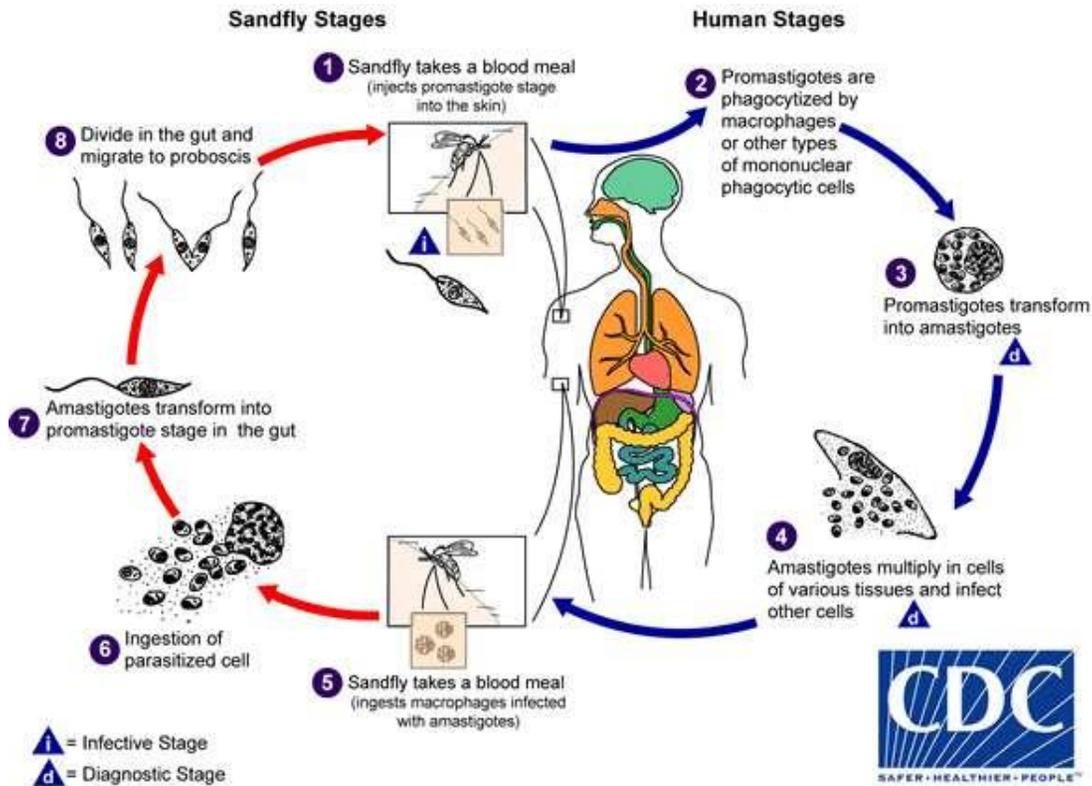


Figura:1 Ciclo biológico de la leishmaniasis humana.^(12, 24)

1.1.3 Formas de leishmaniasis. Manifestaciones clínicas

Las manifestaciones clínicas de la enfermedad dependen de las propiedades del hospedero y de las propiedades del parásito (Anexo 3). En el caso del parásito es importante la especie de *Leishmania* y su perfil isoenzimático. Existen dos formas clínicas principales de leishmaniasis:

- Leishmaniasis visceral (LV)

Se caracteriza por la aparición de fiebre, pérdida de peso, hepatoesplenomegalia, linfadenopatía, pancitopenia e hipergammaglobulinemia entre 3 a 8 meses después de la picada del vector infectado. Puede existir hiperpigmentación de la piel por lo cual se le conoce como enfermedad negra o “kala azar”. Usualmente tiene un curso crónico y puede llevar a la muerte, generalmente por infecciones secundarias. Es causada principalmente por *L. donovani*, *L. infantum* y *L. chagasi*; sin embargo, *L. amazonensis* y *L. tropica* también la pueden producir.

Una variante es la leishmaniasis post-kala azar que es causada por *L. donovani* en la India y África. Aparece un tiempo después de un cuadro resuelto de leishmaniasis

visceral y se caracteriza por máculas, pápulas o nódulos que inician en una distribución peri-oral y luego se diseminan a otras áreas del cuerpo.⁽²²⁾

- Leishmaniasis cutánea ⁽²⁵⁾:

La leishmaniosis cutánea es la manifestación más común de la enfermedad, dándose tanto en infectados por las especies del “Viejo” como del “Nuevo” mundo. En Eurasia y África suele estar causada por *L. major* y *L. tropica* y en menor medida por *L. infantum* o *L. aethiopica*, mientras que en el Nuevo Mundo suele venir dada por las especies *L. amazonensis*, *L. mexicana* y pertenecientes al subgénero *Viannia* donde se encuentran las especies *L. braziliensis* y *L. guyanensis* que se consideran “complicadas” dado el riesgo que tienen de desarrollar la forma mucocutánea. Afecta a la piel, formando lesiones, generalmente en las áreas más expuestas de la misma (brazos, cuello...etc.), dado que los flebotomos no suelen ser capaces de atravesar tejidos con sus piezas bucales. Las picaduras en sí (aún en ausencia de infección por el parásito) se resuelven lentamente, hasta varias semanas, con un importante componente pruriginoso.⁽²⁶⁾

Leishmaniasis cutánea localizada (LCL):

Las lesiones cutáneas causadas por *Leishmania* comienzan como una pápula de color rosado que crecen y pasan a formar una lesión nodular (a veces con un reblandecimiento del centro de la lesión) que suele llevar a una ulceración indolora con un endurecimiento periférico. Tanto en las leishmaniasis del Viejo como del Nuevo Mundo, las úlceras están cubiertas por escaras hiperqueratósicas en el primer caso y por una cubierta fibrinosa en el segundo. Pueden aparecer varias lesiones con una apariencia variable, llegando a darse en algunos casos lesiones satelitales en torno a las úlceras, o incluso diseminarse a través de las vías linfáticas llevando a adenopatías regionales. Asimismo, es común que aparezcan infecciones secundarias de origen bacteriano, frecuentemente asociadas a drenaje purulento y celulitis local. La superinfección bacteriana puede llevar a un agrandamiento de la lesión primaria y a la persistencia de la misma.

En el desarrollo normal de la enfermedad cutánea localizada se suele dar una curación gradual, en un proceso que lleva de meses a años, dependiendo tanto de la especie infectante como del tamaño de lesión, dejando cicatrices atróficas

Revisión Bibliográfica

deprimidas o queloides, (en los pacientes predispuestos a ello). Tras la resolución de la lesión primaria puede darse una reactivación de la misma durante el año siguiente a la curación.

Leishmaniasis Cutánea Diseminada (DCL):

Esta manifestación es relativamente poco frecuente, dándose sobre todo al comienzo de infecciones causadas por *L. aethiopica*, *L. mexicana* y *L. amazonensis*. Comienzan con una lesión localizada que no se ulcera, diseminándose los amastigotes a macrófagos en otras zonas de la piel. Generalmente se forman lesiones en forma de placas o nódulos blandos, normalmente en la cara y en la superficie de los miembros extensores, aunque puede diseminarse al resto del cuerpo. Se detecta un gran número de parásitos mientras que la respuesta linfocítica es escasa, las causas de este déficit de respuesta pueden deberse a una anergia ante los antígenos del parásito o a una inmunosupresión como la causada por el SIDA. El tratamiento eficaz de esta manifestación es complicado.^(22, 26)

Leishmaniasis Mucocutánea (MCL):

La leishmaniasis mucocutánea, también conocida como espundia, es la forma más grave de la leishmaniasis cutánea, siendo causada por las especies del subgénero *Viannia*, siendo *L. braziliensis*, *L. guyanensis*, *L. panamensis*, aunque también se ha descrito en pacientes infectados con especies del subgénero *Leishmania*, como *L. amazonensis* (presente en América del Sur) y *L. infantum* y *L. aethiopica* (leishmaniasis del Viejo Mundo) en pacientes inmunodeprimidos. La clínica de esta afección se caracteriza principalmente por una destrucción de las mucosas, con taponamiento nasal, hemorragias en las mucosas, aumento de las secreciones y desprendimiento del tejido necrosado. En muchos casos se dan deformidades, inflamación y dolor, otro signo característico es la erosión de tejidos mucosales, generalmente en la nariz boca y cartílago nasal, aunque también puede afectar a los carrillos, faringe, paladar, epiglotis, laringe, tráquea y genitales. Los desenlaces más comunes son una desfiguración causada por la destrucción del cartílago nasal y los tejidos circundantes. No es frecuente que afecte a la función respiratoria, asimismo su letalidad es baja.⁽²⁶⁾

Leishmaniasis Cutánea Post-Kala-Azar (PKDL):

Revisión Bibliográfica

Se trata de una erupción cutánea observada en pacientes que han sido sometidos a tratamiento contra la leishmaniasis visceral causada por *L. donovani* (endémica en India, Etiopía y Sudán), tras un período de tiempo tras interrumpir el tratamiento o lograr la curación clínica, también se ha asociado *L. infantum* y *L. amazonensis*. Suele presentarse en forma de máculas eritematosa o hipopigmentadas, que en ocasiones evolucionan a placas o nódulos. Los factores de riesgo para la aparición de esta forma de la enfermedad pueden englobar el estado inmunológico del paciente, la cepa del parásito y la eficacia de la terapia. Se da hasta en un 60% de los casos de leishmaniasis visceral en Sudán, mientras que en el Asia meridional la incidencia se encuentra entre un 5 y un 10%.^(22, 26)

1.1.4 Distribución geográfica

La leishmaniasis está considerada por la OMS en la categoría I de las enfermedades infecciosas (situación emergente e incontrolada), y se reporta endémica al menos en 98 países de todos los continentes, excepto Australia. Se calcula que viven en zonas activas de leishmaniasis alrededor de 350 millones de personas, reportándose alrededor de 2 millones de nuevos casos cada año.

Algunos reportes indican que el número global de casos se ha incrementado en la última década, según lo reportado en Brasil, Bolivia, Colombia, Perú y Afganistán. Tal incremento se explica en parte por el desarrollo de medios diagnósticos y el aumento en la notificación de casos, pero en lo fundamental, es el resultado del inadecuado control de vectores y reservorios en las zonas endémicas, de la mayor detección de leishmaniasis cutánea asociada con infecciones oportunistas y el virus de inmunodeficiencia humana (VIH), así como de la emergencia en la resistencia al tratamiento, entre otros factores.⁽¹¹⁾

1.1.5 Co-infección *Leishmania*/VIH

La asociación de la infección del VIH con la infección por *Leishmania* constituye una nueva enfermedad, extraordinariamente grave y de frecuencia creciente. Descrita por primera vez en España en 1985, se ha constituido como una amenaza real, en particular en el sudoeste de Europa, donde un 25-70% de los casos de leishmaniasis visceral están relacionados con la infección, y un 2-9% de los casos de SIDA desarrollan leishmaniasis visceral. Estudios recientes han puesto de

manifiesto que las cepas latentes del VIH que replican muy lentamente en cultivos celulares, se activan en presencia de *Leishmania*. Esta constatación *in vitro* sugiere que también pueda ocurrir *in vivo*, siendo una explicación plausible de la patogenia de la coinfección. En la figura 2 se muestra la distribución mundial de la leishmaniasis, indicándose las zonas en que existe coinfección *Leishmania*/VIH.

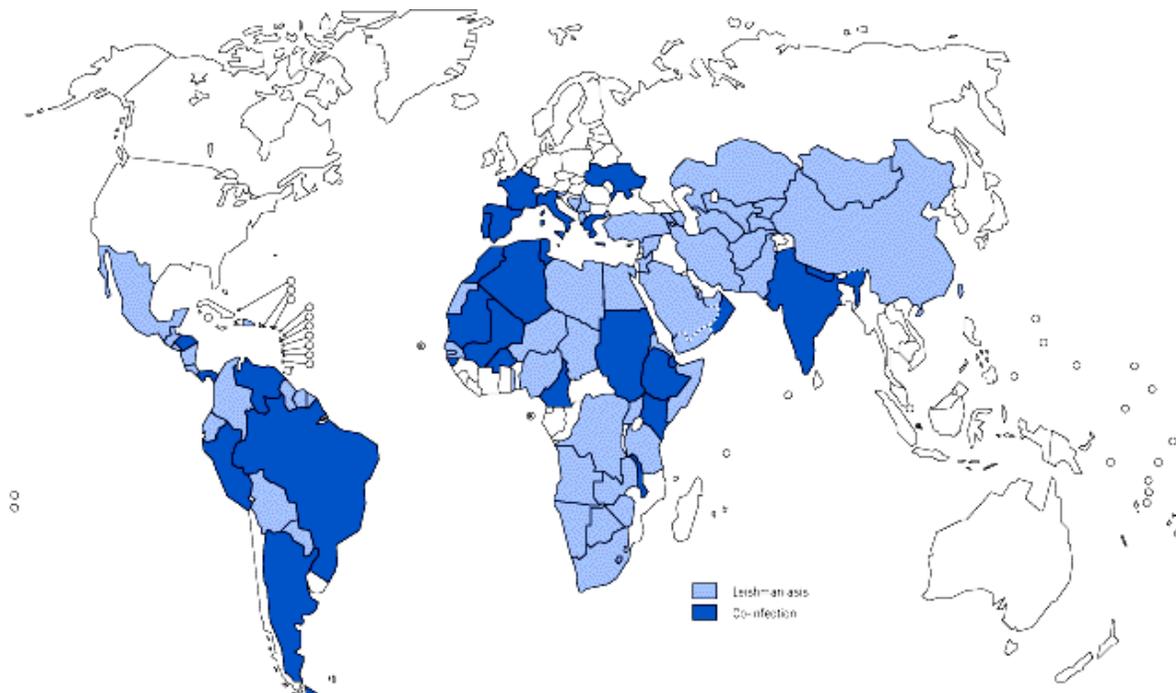


Figura 2 Distribución global de leishmaniasis y de coinfección *Leishmania*/VIH. En azul claro se muestran los casos de leishmaniasis, y en azul oscuro los casos de coinfección de *Leishmania*/VIH en el período 1990-1998. (Organización Mundial de la Salud)⁽²⁷⁾

De los 1.700 casos de coinfección comunicados a la Organización Mundial de la Salud ⁽²⁾, procedentes de 33 países, hasta 1998, 1.440 fueron de la zona: España (835), Italia (229), Francia (259) y Portugal (117). La mayoría eran varones, jóvenes (20-40 años) y adictos a drogas por vía parenteral. En América la mayor parte de los casos procedían de Brasil. En relación con las infecciones por *Leishmania* en inmunocompetentes, la coinfección presenta una serie de características diferenciales:

1. La picadura por el vector infectado en inmunocompetentes no supone necesariamente el desarrollo de la enfermedad, mientras que en

Revisión Bibliográfica

inmunodeprimidos (por infección VIH u otras causas), existe un alto riesgo de desarrollo de leishmaniasis visceral, de forma rápida y grave.

2. El SIDA y la leishmaniasis visceral se insertan en un círculo vicioso de refuerzo mutuo. El primero incrementa el riesgo de leishmaniasis visceral entre 100 y 1.000 veces en áreas endémicas. Las dos enfermedades producen una disminución sinérgica de la respuesta inmunitaria, al destruir las mismas células.
3. El uso de agujas intravenosas en adictos a drogas permite la transmisión directa de persona a persona, lo que constituye un nuevo patrón de transmisión de la enfermedad en el que un nuevo ciclo antroponótico artificial se añade al ciclo zoonótico convencional. Además, el colectivo de pacientes coinfectados se constituye en un nuevo reservorio no existente previamente.
4. El diagnóstico de la coinfección presenta dificultades específicas, debidas fundamentalmente a la destrucción de las mismas células y por consiguiente de la respuesta inmunitaria por ambos agentes infecciosos: sintomatología inusual, en la que no siempre están presentes los signos y síntomas habituales, falsas negatividades serológicas en el 42,6 % de los pacientes coinfectados, enfermedades asociadas como toxoplasmosis y tuberculosis y dificultad de diagnóstico parasitológico. La punción medular es la técnica más sensible y segura para la búsqueda de amastigotes y, si no es posible realizarla, se deben buscar parásitos en sangre periférica. También se han utilizado punciones hepáticas y esplénicas.
5. El tratamiento presenta así mismo serias dificultades, por disminución de la eficacia de los medicamentos, recaídas por resistencias al tratamiento (más del 50 % comparado con el 10 % en inmunocompetentes) y acumulación de efectos secundarios.

La educación sanitaria, en particular a los colectivos de riesgo, adictos a drogas fundamentalmente y declaración de los nuevos casos son factores importantes en la prevención, que deben estar imbricados en los esfuerzos de colectividades y organizaciones implicados en el tema. Desde 1994, una red de vigilancia epidemiológica auspiciada y supervisada por la OMS viene reuniendo datos

inicialmente en 13 países. Esta red se ha hecho extensiva a 28 instituciones y continúa creciendo, estableciendo laboratorios de referencia y hospitales, con una infraestructura capaz de diagnosticar y tratar a los pacientes coinfectados. Los centros de seguimiento siguen protocolos proporcionados por la OMS y el UNAIS (Programa de las Naciones Unidas sobre el Sida) para llegar a conclusiones homogéneas.⁽²⁸⁾

1.1.6 Tratamiento y resistencia

A principios del siglo XX, el tratamiento de la leishmaniasis estaba conformado por diversas sustancias que incluían el tártaro emético esterilizado, curaciones con pomadas fenicadas y aristol con resultados aceptables, pero en esa misma época Kolle, al ensayar una emulsión oleosa de trióxido de antimonio y casi de manera concomitante Martindale, al formular y usar las primeras soluciones a base de anhídrido antimonioso y óxido de antimonio, obtuvieron resultados verdaderamente provechosos; estas preparaciones las ensayó en México Incháustegui en las primeras décadas del siglo pasado con resultados benéficos.⁽⁹⁾ Los antimonios pentavalentes: estibogluconato de sodio (Pentostam®, Glaxo Wellcome, Reino Unido) y antimoniato de meglumina (Glucantime®, Rhone-Poulanc Rorer, Francia), continúan siendo, desde los años 40 hasta la actualidad, el único tratamiento eficaz con el que se logran resultados clínicos y microbiológicos satisfactorios contra la leishmaniasis. Estos compuestos pueden causar serios daños, que son generalmente reversibles, y se reportan entre los efectos colaterales: dolores musculares, falla renal, hepatotoxicidad y cardiotoxicidad. Tanto las drogas utilizadas como la atención médica que requiere su administración, encarecen el tratamiento, y los reportes de pacientes no respondedores se incrementan, ya sea por estar infectados con cepas resistentes o por inmunosupresión (por ej. causada por HIV). Por estas razones, en la pasada década se prestó la mayor atención al desarrollo de alternativas para la dosificación, esquemas de administración o modos de aplicación de los tratamientos.

A pesar de esto, una cuidadosa revisión de los datos acumulados en los últimos años evidencia que los antimoniales pentavalentes, administrados de forma parenteral o intralesional, continúan siendo la terapia de elección para las formas

Revisión Bibliográfica

cutáneas. Otras alternativas terapéuticas incluyen la miltefosina y la anfotericina B, en especial para la forma mucocutánea, Algunas de sus formulaciones lipídicas (AmBisome, Amphocil), que resultan menos tóxicas, se han aplicado en estos casos, y también para tratar la enfermedad visceral en diversas regiones. También es importante señalar que hay evidencias de que la respuesta al tratamiento de las formas cutáneas depende de la especie infectante, lo que se ha demostrado tanto *in vitro* como *in vivo*.^(4, 29)

Los antimoniales trivalentes (Repodral®, Antiomalina®) se indican por vía parenteral, a dosis de 2 a 3 mL (0.02 a 0.03 g) en días alternos hasta completar una serie de 12 a 20 días; los pentavalentes, como Glucantime®, se prescriben a dosis de 10 a 60 mg/kg durante 12 días a tres semanas (o hasta obtener datos de curación clínica y parasitológica) y Pentostán® (estibogluconato sódico), a dosis de 20 mg/kg/día, durante 30 días, no necesariamente consecutivos, cuya periodicidad en las aplicaciones es determinada por la aparición de efectos colaterales. La vía de administración es intravenosa cuando los pacientes estén internados porque la vía intramuscular es muy dolorosa; sin embargo, ésta es la única alternativa factible para el tratamiento a gran escala. Los efectos adversos más frecuentes descritos son reacción local, anorexia, náusea, vómito, mialgias, artralgias, aumento de enzimas hepáticas, de urea, creatinina y alteraciones electrocardiográficas, como la inversión de la onda T, prolongación del segmento Q-T, depresión del segmento S-T y bradicardia sinusal.⁽²⁹⁾

En la forma difusa es útil la pentamidina, que es una segunda alternativa. Los efectos adversos incluyen: nefrotoxicidad, hepatotoxicidad, hipertensión, hipoglucemia, hiperglucemia (raramente), alteraciones electrocardiográficas, abscesos glúteos, parestesia centrofacial, cefalea, epigastralgia y vértigo. Para el tratamiento de las formas cutáneo-mucosas la duración del régimen terapéutico con antimoniales por vía parenteral es de 28 días, con lo que se logra una tasa de curación de alrededor de 75% de los pacientes con formas leves o moderadas, esta tasa es menor en pacientes con formas clínicas más graves en quienes se administra anfotericina B como tratamiento de rescate. Se ha documentado también el tratamiento de la enfermedad con antimoniales por vía intralesional como la

Revisión Bibliográfica

primera línea terapéutica contra leishmaniasis cutánea localizada provocada por *Leishmania major*, *Leishmania tropica*, *Leishmania braziliensis* o *Leishmania panamensis*, o ante riesgo de toxicidad cardiovascular.

En algunos estudios se ha observado que la combinación de antimoniales con lidocaína, además de aminorar las molestias locales, favorece la fragmentación y pérdida de definición de la membrana citoplasmática y otros organelos de los parásitos dada su capacidad anfófila. Como tratamientos alternos de la forma cutánea, la diaminodifenilsulfona (dapsona) ha mostrado utilidad, a dosis de 3 mg/kg/día durante tres semanas. En casos provocados por *Leishmania mexicana* hay respuesta a ketoconazol a dosis de 200-600 mg/día durante cuatro a seis semanas con o sin crioterapia concomitante o a itraconazol, 200-400 mg/día, durante uno a dos meses. Por el contrario, el fluconazol ha demostrado en estudios ser ineficaz como tratamiento contra esta enfermedad.⁽⁹⁾

En la forma visceral, un tratamiento efectivo y seguro es la miltefosina oral. En la forma anérgica se prescriben la anfotericina B y su forma liposómica. El fármaco también produce buena respuesta en pacientes inmunodeprimidos, pero los intervalos de recaídas en estos pacientes son altos. La anfotericina B es una alternativa extremadamente eficaz, pero tóxica, indicada incluso en las formas resistentes a los antimoniales o en pacientes que resultan intolerantes a estos. La anfotericina B liposomal es un tratamiento efectivo contra la leishmaniasis visceral en adultos inmunocompetentes y niños, incluso en sujetos con enfermedad resistente. Otra opción la constituye el factor de transferencia, y de manera más reciente, la paromomicina, que ya ha completado la fase III en estudios experimentales efectuados en la India y África para su aprobación en estos casos. A este respecto, en el estado de Bihar, en la India, en donde las tasas de curación actuales son de 35%, debido a intolerancia a los antimoniales o a la falta de recursos para sostener los tratamientos, se ha sugerido el potencial desarrollo de resistencia a la miltefosina y a la anfotericina liposomal por parte de los parásitos en este lugar. Asimismo, se ha reportado falla al tratamiento con antimoniales pentavalentes en Nepal, por lo que en la actualidad el tratamiento de primera línea en este país es la anfotericina B liposomal, de manera especial contra la forma visceral.⁽²⁹⁾

Revisión Bibliográfica

Otras opciones farmacológicas incluyen la rifampicina, 600 a 1200 mg/día por más de dos meses sola o con isoniazida; interferón γ a dosis de 50 a 100 mm/m² de superficie corporal/día por vía subcutánea durante 10 a 15 días (aunque no lleva a la curación en monoterapia); factor estimulante de colonias de granulocitos-macrófagos humanos recombinante (rHGM-CSF) por vía subcutánea durante 10 días o intralesional; alopurinol asociado o no con antimonioales; metronidazol, 250 mg tres veces al día en ciclos de 10 a 15 días; trimetoprim-sulfametoxazol a dosis de 160/800 mg dos veces al día durante cuatro semanas; azitromicina en esquemas de 500 mg/día durante 3, 5 y 10 días o 1 g/día durante dos días y pentoxifilina a dosis de 1200 mg/día durante 30 días.⁽⁹⁾

Estudios clínicos con paromomicina y miltefosina han mostrado utilidad y actualmente la combinación de miltefosina-termoterapia se considera una alternativa eficaz en el tratamiento de leishmaniasis cutánea. Localmente pueden prescribirse antisépticos como el sulfato de paromomicina a 15% y el cloruro de metilbenzetonio a 12% en ungüento dos veces al día durante 10 días a 3 semanas o solución de sulfato de bleomicina a 1% por vía intralesional. Otra alternativa de tratamiento local ante leishmaniasis cutánea localizada es la inyección de anfotericina intralesional, que ha demostrado efectividad a dosis de 2.5 mg/mL en un estudio realizado en la India.⁽⁹⁾

La resistencia a los medicamentos se define como la disminución de la eficacia de los compuestos en una población de células (es decir, los parásitos de *Leishmania*) que antes eran susceptibles a ellos. Esta definición supone que se conoce la susceptibilidad inicial de los organismos a los fármacos, lo que no siempre es cierto para las cepas de campo de *Leishmania* aisladas de los pacientes. Existen al menos tres mecanismos celulares directamente asociados con la expresión de la quimiorresistencia: amplificación génica de las enzimas diana del fármaco, cambios estructurales y funcionales de las enzimas diana del fármaco y disminución de los niveles intracelulares del fármaco debido a su extrusión por transportadores específicos.⁽¹²⁾ Sin embargo, la quimiorresistencia es un fenómeno multifactorial. Generalmente se asocia con un aumento de la expresión de las proteínas transportadoras del fármaco (multirresistencia, o MDR, también conocida como

glicoproteína) y de la proteína asociada a la multirresistencia (MRP). Los mecanismos de la quimiorresistencia pueden ser múltiples y no exclusivos de un solo tipo de medicamento; los parásitos pueden utilizar múltiples mecanismos para ser resistentes a un medicamento o pueden utilizar un mecanismo para ser resistentes a varios medicamentos que originan los llamados fenómenos de reacción cruzada.

En la leishmaniasis, la quimiorresistencia puede ser natural, o adquirida, que se desarrolla cuando los parásitos están expuestos a dosis subóptimas de fármacos.

De forma similar a lo que ocurre en las células cancerosas de *Leishmania*, la quimiorresistencia se ha asociado con una disminución de la acumulación celular de fármacos, debido a la extrusión por transportadores de membrana tipo ABC. Además, la expresión de la quimiorresistencia está asociada a cambios en los mecanismos bioquímicos y fisiológicos que son fundamentales para la supervivencia y la infectividad de los parásitos. La investigación sobre los cambios inducidos por los medicamentos en la fisiología del parásito puede ciertamente proporcionar pistas para predecir el éxito de una terapia dada.⁽¹²⁾

1.1.7 Vacunas para tratar la leishmaniasis

El objetivo de una vacuna contra la leishmaniasis es inducir en el hospedero una respuesta protectora y de memoria para un posterior encuentro con el parásito. El desarrollo de vacunas es una estrategia que se ha venido trabajando con el fin de prevenir la leishmaniasis en zonas endémicas y evitar su expansión a zonas no endémicas. Se han desarrollado varias generaciones de vacunas contra la leishmaniasis. La primera se centró en parásitos muertos o extractos crudos, fueron desarrolladas a partir de 1930 en Brasil y se han usado en países como Colombia y Ecuador desde la década del 40 del pasado siglo. Estas vacunas han sido preparadas con *Leishmania amazonensis* y/o *Leishmania major*, lisadas por autoclave y con adyuvante de Freund completo. Ejemplos de estas vacunas de primera generación constituyen la *Leishmune*®, que consiste en una preparación enriquecida de glicoproteína de promastigotes de *Leishmania donovani*, También se han usado en perros vacunas contra *Leishmania infantum* atenuadas con gentamicina.⁽³⁰⁾

Revisión Bibliográfica

La segunda generación de vacunas se basó principalmente en la *leishmanización*, consiste en la inoculación deliberada de formas vivas de *Leishmania* atenuados por medios físicos, químicos o genéticos, procedentes del exudado de lesiones cutáneas de personas infectadas, en sitios anatómicos no expuestos de individuos sanos. Esta práctica se ha llevado a cabo principalmente en Israel, Irán y la desaparecida URSS. Su empleo se ha limitado por razones éticas y la aparición de varios efectos adversos, como el desarrollo de lesiones persistentes, psoriasis e hipersensibilidad y persistencia del parásito.⁽³⁰⁾ Otro enfoque de las vacunas de segunda generación es el uso de virus o bacterias vivos recombinantes que expresan antígenos de *Leishmania*. Son ejemplos de proteínas del parásito expresadas por bacterias con este propósito: la proteasa de superficie gp63 de *L. major* y el antígeno de superficie clonado en *Salmonella thypymurium* mutante (patógeno intracelular capaz de penetrar en el macrófago humano), sola o en asociación con IL-2, IFN- γ , TNF- α o antígeno LACK.

La tercera generación incluye, como el procedimiento de vacunación más reciente, la codificación de genes para un antígeno de protección, clonado en un vector. Este método implica la inoculación de DNA insertado en un plásmido bacteriano que actúa como vector. El gen del antígeno vacunal queda bajo el control de un potente promotor mamífero. Cuando se inocula en un animal este plásmido, construido por ingeniería genética, es capturado por las células del hospedero, en las cuales el DNA se transcribe a RNAm y se traduce en la proteína endógena de la vacuna. De esta generación la sustancia más probada como candidato para vacuna recombinante es la proteína de superficie B1 acilada hidrófila (HASPB1), que confiere protección contra *L. donovani*. También se ha empleado el candidato antigénico NH36, que confiere protección contra la infección por *L. chagasi*, *L. mexicana* y *L. amazonensis*. Al mismo tiempo se han empleado las proteínas ribosomales recombinantes L3 y L5 de *L. major* para el desarrollo de vacunas contra *Leishmania*. La tercera generación tiene grandes ventajas con respecto a las anteriores: son vacunas simples, se pueden producir a gran escala, son estables a temperatura ambiente, lo que facilita su almacenamiento y transporte, y, finalmente, se pueden combinar varios antígenos y administrarlos en una sola dosis y por tanto

brindar protección contra más de una especie de *Leishmania*, con lo cual se reduce el número de vacunas.⁽³⁰⁾

1.2 *Leishmania major*. Generalidades

La *Leishmania major* es una de las especies consideradas patógenas para el hombre en el Viejo Mundo y la causa fundamental de leishmaniasis cutánea en regiones del norte de África, semidesiertos del Este Medio, India Oriental y Sudán, Pakistán y Sub-Sahara siendo además la única especie de *Leishmania* detectada en Mali.⁽⁶⁾ Es el agente causal de la leishmaniasis cutánea zoonótica (ZCL, siglas en inglés) la cual se caracteriza por diferentes presentaciones clínicas que dependen en gran parte del agente hospedero. Entre los principales reservorios del Viejo Mundo se encuentran los gerbos y roedores y los principales vectores de transmisión los flebótomos de las especies *P. papatasssi* y *P. duboscqi*.⁽¹²⁾

En condiciones normales, la infección por *Leishmania spp.* tiene una fase aguda y sintomática en la cual los promastigotes del parásito se reproducen dentro de los macrófagos del huésped y se convierten en amastigotes con la capacidad de sobrevivir indefinidamente dentro de los lisosomas de esos fagocitos, pudiendo años más tarde producir una reactivación de la infección con síntomas mucho más intensos que los de la infección inicial.⁽³¹⁾ Los fosfoglicanos, son moléculas ubicadas en la superficie del parásito *Leishmania*, constituyentes de una envoltura densa que protege al parásito y a la vez le confiere virulencia. Su pérdida deja al parásito desprotegido frente al complemento y a otros mecanismos de defensa del huésped y le impide sobrevivir dentro de los macrófagos indefinidamente. La importancia de los fosfoglicanos para la supervivencia parasitaria fue examinada por investigadores de la Universidad de Nueva York en un estudio con parásitos de *Leishmania* desprovistos de fosfoglicanos mediante una delección genética intencionada (*Leishmania major* lpg2⁻). Los investigadores observaron el efecto de la delección sobre la supervivencia del parásito en los flebótomos que transmiten la infección al ser humano y observaron que al cabo de 3 días dicha supervivencia era 10 veces menor que la de parásitos normales y que pasado ese período el parásito desaparecía por completo. Los investigadores también observaron el

Revisión Bibliográfica

comportamiento de los parásitos en macrófagos, donde tampoco pudieron sobrevivir. En cambio, cuando los parásitos mutantes fueron inyectados en ratones genéticamente susceptibles a la leishmaniasis, estos últimos permanecieron asintomáticos y se pensó que el parásito había desaparecido como en los demás casos. Más tarde se descubrió, sorprendentemente, que había parásitos en muestras de tejido obtenidas de estos ratones, deduciéndose que algún tipo de célula distinta de un macrófago le permite al parásito sobrevivir y proliferar sin causar síntomas de enfermedad. En este momento los investigadores están dedicados a buscar esa célula y a estudiar las implicaciones de estos hallazgos desde el punto de vista de la factibilidad de elaborar una vacuna contra la leishmaniasis. En última instancia, *Leishmania lpg2⁻* parece ser un modelo promisorio para estudiar los mecanismos que conducen a la infección asintomática persistente en el ser humano.⁽³¹⁾

Cada vez son más los casos clínicos en los cuales los parásitos de leishmaniasis han desarrollado resistencia, por lo cual es necesario el diseño de nuevos medicamentos, más eficientes, efectivos y seguros. En respuesta a esta necesidad, recientemente, varios compuestos con bases azólicas han sido evaluados por su actividad leishmanicida in-vitro mostrándose como agentes terapéuticos prometedores. Particularmente, Tahghighi y colaboradores, reportaron diecisiete 5-(5-nitrofurán-2-ilos) y treinta y tres 5-(5-nitrotiofen-2ilo)-1,3,4-tiodiazol-2-aminas con una amina acíclica en la posición C2 del anillo tiodiazólico, su actividad leishmanicida fue evaluada *in vitro* contra la especie *Leishmania major* (L. major).⁽³²⁾

Una gran cantidad de estudios realizados acerca de la *Leishmania* concuerdan en que, la respuesta inmune mediada por células tanto en humanos como en modelos experimentales, desempeña un papel muy importante en el proceso de infección por el parásito, así como en el control y la resolución de la leishmaniasis las respuestas inmunológicas del huésped durante la infección se han estudiado a fondo y se ha determinado que la resistencia y susceptibilidad a *L. major* está mediada por Th1 y Th2, subconjuntos de células T CD4+, respectivamente. Las células Th1 secretan citoquinas protectoras IFN- γ , IL-2 y ejecutan respuestas CMI,

Revisión Bibliográfica

y las células Th2 producen citoquinas IL-4, IL-10 e IL-13 y ayudan en la producción de anticuerpos de la inmunidad humoral.⁽¹²⁾

En un estudio realizado con ratones BALB/c previamente infectados con *L. braziliensis* y *L. major* (ambas especies inductoras de leishmaniasis cutánea) pudo apreciarse una respuesta de hiper-sensibilidad de tipo retardado a *L. major* y la producción de grandes cantidades de interferón g (IFN-g) pero bajas cantidades de interleuquina-4. Esto sugiere que existen epítomos de antígenos de reacción cruzada entre *L. braziliensis* y *L. major* y que una pre-infección con *L. braziliensis* provoca o induce una respuesta protectora Th1 al parásito en ratones BALB / c para *L. major*.⁽³³⁾

En la leishmaniasis los parásitos viven a expensas de su anfitrión. Para asegurar la supervivencia del genoma del parásito, es importante que el huésped sobreviva al menos hasta lograr la transmisión exitosa del parásito. El nivel de tetrahidrobiopterina (H4B) juega un papel crucial en el desarrollo metacíclico del parásito, la forma que es infectiva para el huésped mamífero. Stephen Beverley y colaboradores han demostrado que los niveles bajos H4B causan metaciclogénesis elevada. Tras la inyección en ratones de parásitos de *L. major*, estos parásitos mostraron ser hipervirulentos, causando más lesiones que en los controles, así como un mayor número de parásitos por lesión. Los autores sugieren que, para limitar el daño al huésped, el parásito utiliza PTR 1 para aumentar H4B de nuevo, controlando de este modo su propia virulencia.⁽¹⁰⁾

En general, la leishmaniasis cutánea se caracteriza por una lesión que cura en un período de tiempo. Sin embargo, en casos raros, la infección puede darse de una forma que imposibilite la curación de la enfermedad y persistir durante varios años y no responder a ninguna de las quimioterapias existentes. Una de las cepas que causan esta forma grave de la enfermedad *Leishmania major* Seidman (LMSD) fue aislada hace 45 años de un voluntario del Cuerpo de Paz en Senegal. El paciente exhibió lesiones cutáneas crónicas con la persistencia de los parásitos a pesar de una respuesta inmune mediada por células. Esta respuesta enigmática se reprodujo en ratones cepa C57BL/6, que son típicamente resistentes a *L. major*, y desarrollaron lesiones cutáneas de difícil curación después de la infección con

LMSD, mientras que exhibieron una fuerte respuesta Th1 que se caracteriza por altos niveles de IFN-g e IL-13 y ausencia de IL-4. Se ha demostrado que se requerían de IL-1b, y el reclutamiento de neutrófilos para la susceptibilidad a LMSD, pero el mecanismo para estos procesos no estaba claro. La respuesta ahora se ha proporcionado en un nuevo estudio realizado por Lee y colegas, los cuales identificaron una población de macrófagos dérmicos MRC1hi que fueron infectadas preferentemente por LMSD.⁽³⁴⁾

1.3 Enfoques para la búsqueda de medicamentos destinados a la leishmaniasis

La leishmaniasis no constituye una enfermedad de interés para las compañías farmacéuticas. Solo por iniciativa de instituciones académicas, así como de organizaciones humanitarias y no-gubernamentales sin fines de lucro como la OMS se han logrado importantes progresos en el control de esta y otras enfermedades olvidadas en particular el desarrollo de nuevas alternativas terapéuticas. Por las razones anteriores se insiste en el uso racional y eficiente de los recursos financieros y las capacidades disponibles en este campo.

El reposicionamiento de fármacos constituye una de las opciones para el descubrimiento de nuevas alternativas terapéuticas para las enfermedades parasitarias olvidadas. Ello consiste en estudiar entre los medicamentos aprobados o en fases avanzadas de desarrollo para otras indicaciones, aprovechando la información disponible de los productos. De estos compuestos ya se dispondrá información valiosa como puede ser información físico-química, la analítica, los estudios toxicológicos pre-clínicos, los estudios farmacocinéticos, de estabilidad, de formulación e interacciones; los cuales representan un gasto sustancial en el proceso de desarrollo del fármaco. Para la nueva indicación solo se requiere demostrar la actividad en el laboratorio, así como en modelos animales de la enfermedad, y entonces ejecutar los ensayos clínicos según las nuevas condiciones de uso.⁽²³⁾



Figura 3. Esquema comparativo del camino a recorrer por un fármaco en su desarrollo para la primera indicación (A) y al ser reposicionado (B).

Existen múltiples ejemplos del éxito de esta estrategia. En el caso de la leishmaniasis, la indicación primaria de la anfotericina B y el ketoconazol fue como antimicóticos; la paromomicina, es un antibiótico aminoglucósido; la miltefosina, un antineoplásico; la pentamidina se emplea para el tratamiento de la neumonía por *Pneumocystis carinii* y de la tripanosomiasis africana; y el buparvaquone, originalmente registrado como medicamento para el tratamiento de la teileriosis bovina, se encuentra en fase de desarrollo para la leishmaniosis cutánea.⁽²³⁾

1.4 Estudios *in silico*

Podemos definir como **modelo experimental** a cualquier sistema lógico, físico o biológico capaz de simular total o parcialmente el proceso o fenómeno que pretendemos estudiar.⁽³⁵⁾ Podemos considerar los siguientes tipos de modelos experimentales:

- Los modelos matemáticos computacionales o modelos *in silico*.⁽³⁶⁾
- Los modelos físicos o mecánicos.
- Los modelos biológicos.
 - Celulares y tisulares.
 - Órganos aislados.
 - Animales completos.
 - El ser humano como sujeto experimental.

Los modelos matemáticos computacionales permiten simular situaciones reales, modificando los valores de las variables involucradas en un amplio rango de valores

y contrastarlos con los valores reales. De esta manera nos mostrarán comportamientos que quizá llevaría décadas obtener de la observación o de la experimentación.⁽³⁶⁾

El desarrollo de fármacos es sin dudas un proceso complejo y altamente costoso, que en su forma racional incluye la identificación de compuestos que se unen a un blanco terapéutico o que muestran determinada actividad biológica. El empleo de métodos *in silico* constituye una vía cada vez más utilizada en el desarrollo de fármacos ya que permite disminuir de forma apreciable el tiempo y los costos asociados al desarrollo de los mismos. Los objetivos principales son:

- *Diseño o identificación de nuevos compuestos*, esto es, diseñar nuevas estructuras con efecto biológico en una categoría terapéutica deseada. Una alternativa es buscar dentro de colecciones de compuestos existentes moléculas que tengan una actividad biológica específica.
- *Seleccionar candidatos*. Los métodos computacionales ayudan a distinguir las moléculas sobre las cuales se deben enfocar primero las pruebas experimentales, por ejemplo, compra, síntesis, y evaluación biológica. Sin embargo, los métodos de cómputo no reemplazan a los experimentos.
- *Optimizar Líderes* para mejorar las propiedades deseadas y disminuir los efectos adversos.⁽³⁷⁾

1.5 Estudios QSAR

Los estudios QSAR (*Quantitative Structure-Activity Relationships*, por sus siglas en inglés o Relaciones Cuantitativas Estructura Actividad, en español) engloban las relaciones cuantitativas entre la estructura molecular y la actividad biológica y/o química-física, lo cual constituye uno de los paradigmas de la Química Médica. Estos estudios reúnen un conjunto de técnicas computacionales relacionadas con diseño y visualización espacial virtual de moléculas (también llamado *in silico*), cálculo de propiedades fisicoquímicas moleculares (descriptores), bioinformática y estadística; todo esto con el fin de hacer una predicción teórica de la actividad biológica que permita el diseño teórico de posibles futuros nuevos fármacos (aún no existentes), evitando pasar por el proceso a prueba y error de síntesis orgánica.⁽³⁸⁾

Revisión Bibliográfica

Los estudios QSAR tienen sus primeros antecedentes históricos entre las décadas del 60 y del 80 del pasado siglo cuando surgen las primeras aproximaciones como Las Relaciones Lineales de Energía Libre (LFER), Las Relaciones Lineales de Energía de Solvatación (LSER), los análisis de Hansch, entre otras.⁽³⁹⁾ Para la realización de estudios QSAR se efectúan los siguientes pasos generales:⁽³²⁾

- Selección de un número de moléculas con valores conocidos de actividad biológica.
- Cálculo de los descriptores moleculares correspondientes.
- Obtención, análisis y validación de modelos matemáticos.
- Cribado virtual de base de datos

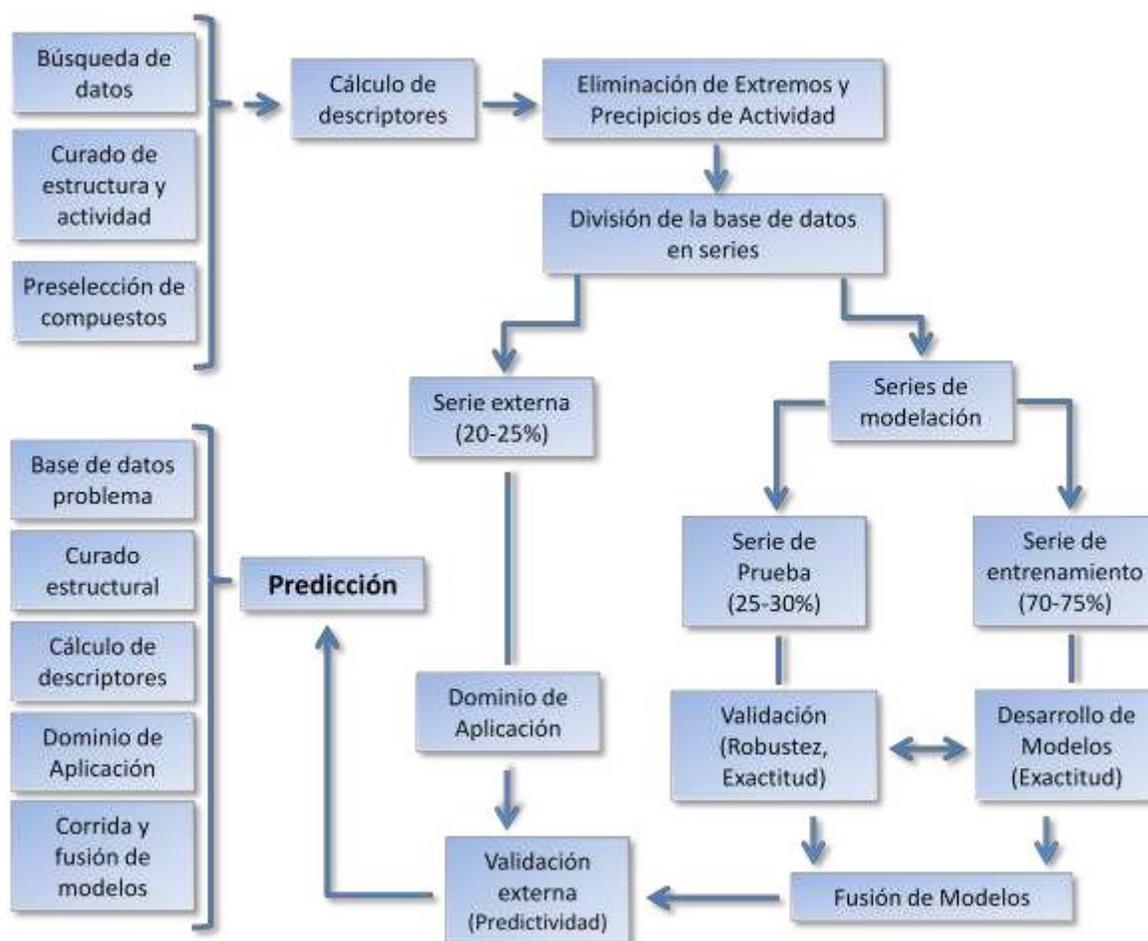


Figura 4. Diagrama para el desarrollo de un modelo QSAR

En la Figura 4 se representa esquemáticamente la secuencia de pasos para solucionar un problema de modelación mediante estudios cuantitativos de relación

estructura-actividad. En los acápites siguientes se profundizará en cada uno de los aspectos. No obstante, puede resumirse que el proceso inicia con la elección correcta de los datos y su minuciosa inspección para eliminar posibles errores.

1.6 Descriptores moleculares

Los Descriptores o Índices Moleculares (DMs o IMs) definidos como el resultado final de un procedimiento lógico-matemático el cual transforma la información química, codificada dentro de una representación simbólica de una molécula, en un número (o conjunto de estos) útil (útiles) o en el resultado de alguno de los experimentos ya estandarizados.^(39, 40) Actualmente existen más de 5000 tipos de descriptores⁽⁴⁰⁾, lo que evidencia el desarrollo y evolución de los mismos por desempeñar un papel fundamental en el desarrollo de modelos para la química, las ciencias farmacéuticas, la política de protección del medio ambiente, la toxicología, la ecotoxicología, la investigación sanitaria y el control de calidad.⁽⁴⁰⁾

Los descriptores de acuerdo con la definición dada se pueden clasificar en dos grandes grupos:

1. Mediciones experimentales: log P, Refractividad Molar, Momento Dipolo, Polarizabilidad, etc.
2. Descriptores Moleculares Teóricos: Se derivan de las representaciones simbólicas de las moléculas y pueden ser clasificados a su vez de acuerdo al tipo de representación como:
 - Descriptores 0D: Descriptores moleculares obtenidos a partir de la fórmula química de la molécula. Entre ellos se encuentran: Número de átomos (A), Peso Molecular (Mw), Número de átomos específicos (Nx), descriptores constitucionales y cualquier función de las propiedades atómicas (carga atómica, radio de Van der Waals, polarizabilidad atómica y constantes hidrofóbicas atómicas).⁽³⁹⁻⁴¹⁾
 - Descriptores 1D: Descriptores moleculares derivados de una lista subestructural, la cual consiste en una lista parcial de fragmentos estructurales incluidos en una molécula. No pueden representar toda la topología de la estructura química, pueden calcularse e

interpretarse fácilmente y se utilizan normalmente para el análisis de similitud/diversidad y el cribado virtual de grandes bases de datos de sustancias químicas. ⁽³⁹⁻⁴¹⁾

- Descriptores 2D: Se trata de una representación bidimensional y/o topológica que, además de la composición atómica, incluye la información sobre la conectividad de los átomos de la molécula en términos de un gráfico molecular. Constituyen los descriptores más comúnmente usados, ejemplo: Perfiles Moleculares, Índices de Información Topológica. ⁽³⁹⁻⁴¹⁾
- Descriptores 3D: Derivan la representación molecular donde la información espacial sobre las posiciones de los átomos se añade a las conectividades del átomo, o sea, cada átomo de la molécula se asocia con un conjunto de coordenadas tridimensionales que determinan su propia posición en el espacio tridimensional. Esta representación se denomina representación geométrica de una molécula en la que la molécula se considera una estructura geométrica rígida. Ejemplo de estos descriptores son los descriptores geométricos, estéricos y de tamaño. ⁽³⁹⁻⁴¹⁾
- Descriptores 4D: Descriptores que tienen en cuenta las propiedades moleculares derivadas de la distribución de los electrones y las interacciones de las moléculas con las sondas que caracterizan el ambiente (campos de interacción molecular). ⁽³⁹⁻⁴¹⁾

1.7 Softwares empleados en los estudios QSAR

1.7.1 DRAGON

El software DRAGON es un software surgido en 1994 por Milano Chemometrics and QSAR Research Group con el nombre "WHIM/3D QSAR". Ha sido concebido para el cálculo de una gran variedad de descriptores moleculares derivados de diferentes representaciones moleculares, permitiendo al usuario elegir aquellos descriptores moleculares que sean más adecuados para su investigación específica. DRAGON permite fusionar descriptores moleculares calculados y propiedades definidas por el

usuario para un conjunto de moléculas, proporcionando un archivo de salida completo que se carga fácilmente con cualquier aplicación de análisis de correlación. Posee incluido un menú que permite el cálculo de la correlación de pares entre los descriptores moleculares y las propiedades experimentales y el análisis gráfico de la distribución de las moléculas en el espacio descriptor y de respuesta.⁽⁴²⁾

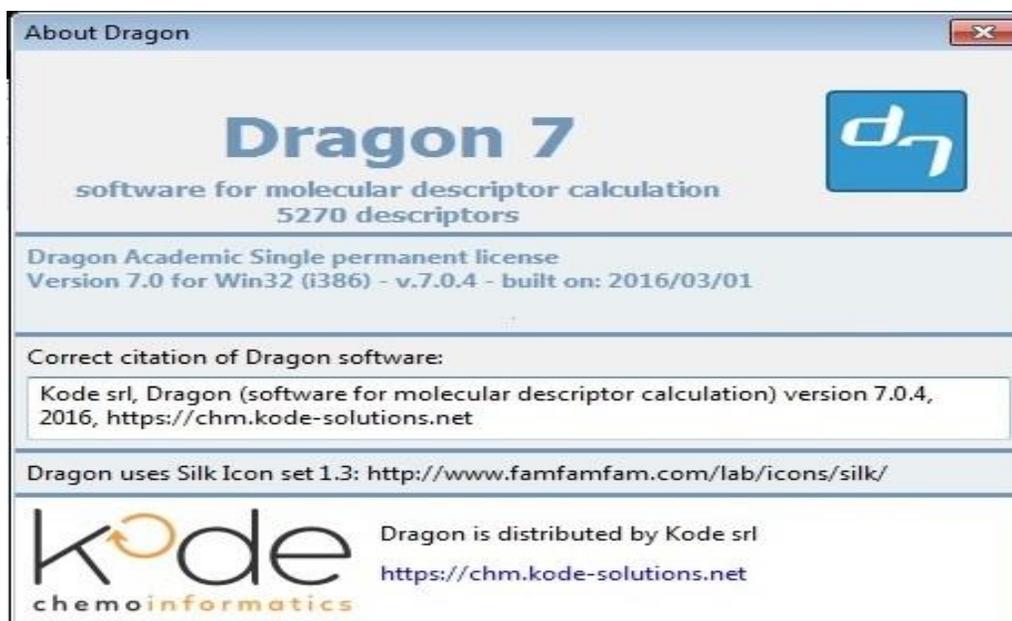


Figura 5. Software DRAGON 5.5

1.7.2 WEKA

Se denomina minería de datos a la aplicación de técnicas en grandes volúmenes de datos para descubrir información útil, aplicable y no trivial, además de obtener relaciones y tendencias ocultas hasta el momento.⁽⁴³⁾ Existen diversos sistemas computacionales que permiten aplicar estas técnicas con muy buenos resultados. Una herramienta de gran utilidad y fácil acceso es el software WEKA (*Waikato Environment for Knowledge Analysis* o en español Entorno para Análisis del Conocimiento de la Universidad de Waikato). WEKA es un entorno para experimentación de análisis de datos que permite aplicar, analizar y evaluar las técnicas más relevantes de análisis de datos, principalmente las provenientes del aprendizaje automático, sobre cualquier conjunto de datos del usuario.⁽⁴³⁻⁴⁵⁾ de tres entornos de trabajo gráficos y un entorno en modo consola, permitiendo la

Revisión Bibliográfica

implementación de algoritmos para preprocesamiento de datos, clasificación, regresión, clustering, selección de atributos, reglas de asociación, etc.⁽⁴⁴⁾

El desarrollo de WEKA se inició en 1993 en la Universidad de Waikato (Nueva Zelanda) siendo la primera versión pública que se liberó Weka 2.1 del año 1996.^{(43,}

⁴⁴⁾ Más tarde el código fue enteramente reescrito en Java lo que permitió su ejecución en cualquier plataforma dotada de una implementación de la máquina virtual de Java. Actualmente la última versión de WEKA a mayo de 2019 es la 3.9.3 y se distribuye como software de libre distribución (licencia GNU GPL).



Figura 6 Logo e interfaz principal del WEKA. Plataforma de software para el aprendizaje automático y la minería de datos.

El software lleva el origen de su nombre por un ave de igual nombre Weka (*Gallirallus australis*) endémica de Nueva Zelanda en peligro de extinción, famosa por su curiosidad y agresividad. De aspecto pardo y tamaño similar a una gallina y alimentación basada en insectos y pequeños frutos.^(43, 46)

1.7.3 ChemOffice

ChemOffice es una colección de herramientas inteligentes e integradas que permite a los científicos e investigadores trabajar con estructuras químicas (en 2D o 3D) así como obtener información sobre sus propiedades químicas y físicas, almacenando esta de tal forma que posteriormente pueda compartirla con el mundo científico. Es útil para visualizar y ganar en profundidad en la comprensión de resultados, así como para correlacionar la actividad biológica con las estructuras químicas. Está compuesto, a su vez, por otras herramientas de trabajo como:

Revisión Bibliográfica

- ChemDraw: herramienta de dibujo y análisis de estructuras químicas tanto 2D como 3D, además de esquemas de reacción e información textual. La información se puede cortar y pegar fácilmente en documentos de Word y otras aplicaciones. Las moléculas pueden ser convertidas a otros formatos de archivo molecular y los análisis (como el cálculo teórico de los espectros de RMN 1H y 13C) son rutinarios.
- Chem3D: potente plataforma para realizar modelos, cálculos moleculares y visualización de compuestos.
- ChemFinder: un gestor de base de datos químicamente inteligente y un motor de búsqueda que crea hojas de cálculo en Excel para su consulta. Proporciona una base de datos de investigación y proporciona enlaces web para más de 100.000 sustancias químicas que se pueden buscar por nombre, nombre parcial o por su estructura. Pueden ser configurado además para entregar las propiedades calculadas.
- E-Notebook Enterprise: permite la fácil compilación de notas de investigación diarias. Las páginas del cuaderno consisten en hojas de cálculo de Excel, documentos de Word, dibujos de ChemDraw y datos espectrales, y se pueden buscar por texto y estructura.

Además, posee otras herramientas como ChemScript, MNova, ChemInfo y BioAssay, interfaces que se complementan las unas a las otras. La integración de estas aplicaciones con los programas de Microsoft Office (Word, Excel, etc.) ofrece soluciones completas para gestionar, modelar y publicar la información química y biológica.⁽⁴⁷⁾

1.7.4 ChemAxon. Principales características y aplicaciones

ChemAxon es una compañía que proporciona plataformas de desarrollo de software químico basadas en Java para las industrias biotecnológica y farmacéutica y crea soluciones de plataforma cruzada basadas en la web para la quimioinformática y la comunicación química. Los productos ChemAxon incluyen herramientas para la visualización y el dibujo de moléculas, la búsqueda y gestión de bases de datos químicas y el descubrimiento de fármacos; estos productos tienen licencia gratuita para uso académico. A continuación, se ofrece la lista de aplicaciones que ofrece:

Revisión Bibliográfica

- Aplicaciones independientes: BioEddie, ChemAxon Assay, ChemCurator, Chemicalize, ChemLocator, InstantJChem, MadFast Similarity Search, Marvin, Marvin Live, Plexus Connect, Plexus Design, etc.
- Paquete de herramientas y componentes: Biomolecule Toolkit, Calculator and Predicts, Chemical Name and Structure Conversion, JChem Engines, Marvin JS, Reactor, etc.
- Integración con terceros: .NET Connectors, JChem for Office, JChem for Share Point, etc.⁽⁴⁸⁾

Marvin: es una colección de herramientas Java para dibujar, mostrar y caracterizar estructuras químicas, subestructuras y reacciones. Contiene limpiador 2D para mejorar la calidad de la visualización de las moléculas y un optimizador de geometría propio capaz de crear estructuras 3D a partir de la conectividad y realizar una búsqueda conformacional. Incluye las aplicaciones Marvin Beans, Marvin Sketch, Marvin View, MarvinSpace.^(47, 48)

JChem: Es una herramienta Java para el desarrollo de aplicaciones que permiten la búsqueda de datos estructurales y no estructurales mixtos. Puede integrar una variedad de sistemas de bases de datos (Oracle, MS SQL Server, DB2, Access) con interfaces web y ofrece una subestructura rápida, y un motor de búsqueda exacto. Las estructuras se almacenan en tablas de base de datos y se pueden combinar datos estructurales y no estructurales. Permite importar y exportar archivos SDF, SMILES, etc.⁽⁴⁷⁾ Incluye las aplicaciones JChemBase, Standardizer, Screen, Reactor, Fragmenter, Serial Molecule Generator, etc.⁽⁴⁷⁾

Standardizer: es una herramienta de canonización de estructura en JChem para convertir de diferentes fuentes a formas de representación estándar. El programa puede automatizar la identificación de mesómeros y tautómeros, ofreciendo muchas más funciones para la preparación de compuestos para su almacenamiento en bases de datos. Permite la preparación de moléculas flexibles para su transformación, esto incluye: manipulaciones de hidrógeno, transformaciones de los enlaces aromáticos, eliminación de disolventes y contra iones, etc.⁽⁴⁷⁾

Instant JChem: es una herramienta lista para usar que permite crear, gestionar y analizar estructuras químicas y base de datos. Soporta la funcionalidad de consulta

clasificación y manejo de grandes volúmenes de datos (cientos de miles de estructuras), tanto en bases de datos locales como remotas, sin ninguna administración adicional. Tiene una amplia gama de funcionalidades, incluyendo vistas de bases de datos personalizables, integración de enumeración de bibliotecas y población dinámica de columnas con propiedades moleculares singulares y combinadas y predicciones de descriptores.⁽⁴⁷⁾

1.7.5 ISIDA/QSPR

El software ISIDA / QSPR es una aplicación que utiliza fragmentos moleculares subestructurales (SMF) y análisis de regresión lineal multidimensional (MLR) para la creación de modelos QSPR y QSAR y la predicción de propiedades físicas, químicas y biológicas. Forma parte del proyecto forma parte del proyecto ISIDA de colaboración entre el Laboratorio de Quimioinformática del Prof. Alexandre Varnek (Universidad de Estrasburgo, Francia) y el Dr. Vitaly Solov'ev (Instituto de Química Física y Electroquímica, Moscú, Rusia). El programa ISIDA / QSPR incluye:

- Editor de EdiSDF para visualizar y editar archivos que contienen fórmulas químicas 2D y 3D en formato de archivos de datos de estructura (SDF).
- Programa FMF para predecir las propiedades físicas, químicas y biológicas utilizando modelos confeccionados por ISIDA / QSPR.
- Módulo MolFrag para el análisis de fragmentos moleculares subestructurales (SMF) y sus contribuciones.⁽⁴⁷⁾

1.8 Inteligencia Artificial

La inteligencia artificial puede definirse como una disciplina académica relacionada con la teoría de la computación cuyo objetivo es emular algunas de las facultades intelectuales humanas en sistemas artificiales para lograr en los mismos procesos de percepción sensorial y el reconocimiento de patrones. Para el diseño de un sistema de inteligencia artificial se requiere de herramientas de disciplinas como el cálculo numérico, la estadística, la informática, el procesado de señales, el control automático o la neurociencia. En la actualidad, la inteligencia artificial se ha consolidado como una disciplina que permite diseñar aplicaciones de gran utilidad práctica en numerosos campos. Actualmente, existe una enorme lista de ámbitos

de conocimiento en los que se emplea, entre los son de especial relevancia la minería de datos, el diagnóstico médico, la robótica, la visión artificial, el análisis de datos o la planificación y logística.⁽⁴⁹⁾

1.8.1 Aprendizaje o adquisición de conocimiento en la Inteligencia Artificial

En la actualidad, gracias a los avances de la informática, es posible almacenar y procesar grandes cantidades de datos a través de las redes computacionales.⁽⁵⁰⁾ En este contexto aparece el concepto de aprendizaje en informática, denominado también aprendizaje de máquina o automático, que corresponde a programas computacionales que buscan optimizar los parámetros de un modelo usando datos previos o datos de entrenamiento. Los modelos pueden ser inductivos, cuando permiten hacer predicciones sobre el futuro o bien descriptivos cuando permiten generar conocimiento a partir de los datos.⁽⁵⁰⁾

El aprendizaje automático usa la teoría estadística para construir modelos matemáticos, pues de esta manera es posible hacer inferencias a partir de una muestra. La ciencia de la computación es requerida en la fase de entrenamiento para la implementación de algoritmos de optimización eficientes. Una vez que un modelo es ajustado, se requiere también eficiencia en su representación y solución algorítmica para la fase de inferencia. Se distinguen tres distintos tipos de aprendizaje: supervisado, no supervisado y por refuerzo. El problema de aprendizaje supervisado consiste en aprender una función a partir de ejemplos de sus entradas y sus salidas. Corresponde a la situación en que se tiene una variable de salida, ya sea cuantitativa o cualitativa, que se desea predecir basándose en un conjunto de características. Se establece un modelo que permite relacionar las características con la variable de salida. Luego se considera un conjunto de datos de entrenamiento en los cuales se observan tanto los valores de la variable de salida como de las características para determinados individuos (personas u otros). Usando tales datos se ajustan los parámetros del modelo, con lo cual es posible predecir valores de la variable de salida para nuevos individuos.^(50, 51)

El aprendizaje no supervisado consiste en aprender a partir de patrones de entradas para los que no se especifican los valores de sus salidas. Corresponde a la situación en que existe un conjunto de datos que contienen diversas características de

determinados individuos, sin que ninguna de ellas se considere una variable de salida que se desee predecir. En este caso la tarea de aprendizaje es describir como están organizados los datos, posibles asociaciones entre ellos o agrupamientos.⁽⁵⁰⁾

⁵¹⁾ El aprendizaje por refuerzo típicamente incluye el subproblema de aprender cómo se comporta el entorno. En vez de que un profesor indique al agente qué hacer, el agente de aprendizaje por refuerzo debe aprender a partir del refuerzo. En el aprendizaje por refuerzo tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

1.9 Clasificación. Algoritmos de aprendizaje automático para clasificación

La clasificación constituye una de las tareas de reconocimiento de patrones en la que se etiqueta a un individuo a partir de ciertas propiedades que lo caracterizan; entendiendo como individuo una entidad de cualquier tipo y debe cumplir que sea una clase discreta. La clasificación, desde el punto de vista del reconocimiento de patrones, involucra tres fases importantes: la definición de las clases, la representación de la información en forma de atributos y el aprendizaje mediante algoritmos. El objetivo del aprendizaje automático para clasificación consiste en inducir una aproximación (modelo o hipótesis) h de una función desconocida f definida desde un espacio de entrada X hacia un espacio discreto y desordenado $Y = 1, \dots, K$, dado un conjunto de entrenamiento S .⁽⁴⁹⁾

Podemos clasificar los algoritmos en función del principio de inducción que usan para adquirir los modelos o reglas de clasificación. Estos métodos están basados en: probabilidades, distancias, reglas o kernels. La tabla 1 recoge los algoritmos de aprendizaje supervisado para clasificación.⁽⁴⁹⁾

Tabla 1. Algoritmos de aprendizaje supervisado para clasificación según el basamento del método.

Base del método	Algoritmo de clasificación
Métodos Probabilísticos	Naive Bayes, Máxima Entropía

Revisión Bibliográfica

Métodos Basados en distancias	kNN, Clasificador lineal basado en distancias, <i>Clustering</i> dentro de clases
Métodos basados en clases	Árboles de decisión, AdaBoost
Métodos basados en kernels	Clasificador lineal con kernel, Kernels para tratamiento de textos, Máquinas de vectores de soporte
Redes neuronales	Perceptrones

1.9.1 kNN

KNN (En inglés, *k nearest neighbours*) es un algoritmo de aprendizaje supervisado en el cual se clasifican los nuevos ejemplos a partir de medida de similitud o distancia. En este algoritmo la clasificación de nuevos ejemplos se realiza buscando el conjunto de los k ejemplos más cercanos de entre un conjunto de ejemplos etiquetados previamente guardados y seleccionando la clase más frecuente de entre sus etiquetas. Se basa en la suposición de que los prototipos más cercanos tienen una probabilidad a posteriori similar.^(49, 51, 52) Para la realización de este algoritmo se emplean tres conjuntos de datos, el conjunto de entrenamiento que utilizaremos para el aprendizaje del clasificador y los conjuntos de validación y de test con los que se comprueban si el clasificador es capaz de generalizar, es decir si presenta buenos resultados al introducir datos no empleados durante el entrenamiento. Los pasos seguidos para la implementación de este algoritmo incluyen:

- El proceso de aprendizaje de este clasificador consiste en almacenar en un vector el conjunto de entrenamiento, junto a la clase asociada a cada muestra de este conjunto. En primer lugar, y con motivo del aprendizaje del algoritmo, se calcula la distancia euclídeana de cada muestra de entrenamiento, a todas las demás que se tienen almacenadas en el vector del punto anterior y de las que conocemos la clase a la que corresponden, se mantienen las K muestras más cercanas y clasificando la nueva muestra de entrenamiento en la clase más frecuente a la que pertenecen los K vecinos obtenidos anteriormente. La segunda tarea para diseñar el clasificador, es realizar el mismo proceso con los datos de validación: se calcula el porcentaje de clasificación sobre los ejemplos

de este conjunto (desconocidos en la tarea de aprendizaje) para conocer su poder de generalización.

- El método kNN tiene como ventajas que el coste del aprendizaje es nulo, no se necesita hacer ninguna suposición sobre los conceptos a aprender y además se pueden aprender conceptos complejos usando funciones sencillas como aproximaciones locales. Las inconvenientes del algoritmo es que Puede producir resultados equivocados si el cálculo de distancia no es el adecuado, el coste de encontrar los k mejores vecinos es grande, no hay un mecanismo para decidir el valor óptimo para k (depende de cada conjunto de datos), su rendimiento baja si el número de descriptores asciende y su interpretabilidad es nula (no hay una descripción de los conceptos aprendidos).⁽⁵²⁾

1.9.2 Árboles de decisión

Un árbol de decisión toma como entrada un objeto o una situación descrita a través de un conjunto de atributos y devuelve una decisión: el valor previsto de la salida dada la entrada. Un árbol de decisión desarrolla una secuencia de test para poder alcanzar una decisión. Cada nodo interno del árbol corresponde con un test sobre el valor de una de las propiedades, y las ramas que salen del nodo están etiquetadas con los posibles valores e dicha propiedad. Cada nodo hoja del árbol representa el valor que ha de ser devuelto si dicho nodo hoja es alcanzado.⁽⁵¹⁾ El proceso de construcción del árbol es un proceso iterativo, en el que, en cada iteración, se selecciona el atributo que mejor particiona el conjunto de entrenamiento. Para realizar este proceso, tenemos que mirar la bondad de las particiones que genera cada uno de los atributos y, en un segundo paso, seleccionar el mejor.

Los árboles de clasificación o de decisión, se caracterizan por su sencillez, su campo de acción abarca diversas áreas como: el diagnóstico médico, juegos, predicciones meteorológicas y control de calidad, este es el modelo de aprendizaje inductivo supervisado no paramétrico más utilizado. Como forma de representación del conocimiento, Presentan dos algoritmos principales: algoritmo ID3 y algoritmo C5.0.⁽⁵³⁾ Este algoritmo tiene la gran ventaja de la facilidad de interpretación del modelo de aprendizaje. Un árbol de decisión nos da la información clara de la toma de decisiones y de la importancia de los diferentes atributos involucrados. Dos de

los grandes inconvenientes que plantea son la elevada fragmentación de los datos en presencia de atributos con muchos valores y el elevado coste computacional que esto implica. Esto hace que no sea un método muy adecuado para problemas con grandes espacios de atributos; como pueden ser aquellos que contienen información léxica, como la categorización de textos o los filtros anti spam. Otro inconveniente a tener en cuenta es que los nodos terminales correspondientes a reglas que dan cobertura a pocos ejemplos de entrenamiento no producen estimaciones fiables de las clases. Tiende a sobreentrenar el conjunto de entrenamiento.⁽⁴⁹⁾

1.9.3 Máquinas vectoriales de soporte

Las máquinas vectoriales de soporte (SVM, del inglés *Support Vector Machine*) tienen su origen en los trabajos sobre la teoría del aprendizaje estadístico y fueron introducidas en los años 90. Dentro de la tarea de clasificación, las SVMs pertenecen a la categoría de los clasificadores lineales. Aunque originariamente las SVMs fueron pensadas para resolver problemas de clasificación binaria, actualmente se utilizan para resolver otros tipos de problemas (regresión, agrupamiento, multclasificación).⁽⁵⁴⁾

Mientras la mayoría de los métodos de aprendizaje se centran en minimizar los errores cometidos por el modelo generado a partir de los ejemplos de entrenamiento (error empírico), el sesgo inductivo asociado a las SVMs radica en la minimización del denominado riesgo estructural. La idea es seleccionar un hiperplano de separación que equidista de los ejemplos más cercanos de cada clase para, de esta forma, conseguir lo que se denomina un *margen máximo* a cada lado del hiperplano. Además, a la hora de definir el hiperplano, sólo se consideran los ejemplos de entrenamiento de cada clase que caen justo en la frontera de dichos márgenes. Estos ejemplos reciben el nombre de *vectores soporte*. Desde un punto de vista práctico, el hiperplano separador de margen máximo ha demostrado tener una buena capacidad de generalización, evitando en gran medida el problema del sobreajuste a los ejemplos de entrenamiento.^(49, 54) Las ventajas de las SVM radican en la relativa facilidad del entrenamiento, no hay óptimo local, como en las redes neuronales, se escalan relativamente bien para datos en espacios dimensionales

altos y el compromiso entre la complejidad del clasificador y el error puede ser controlado explícitamente. La principal desventaja del algoritmo es que se necesita una buena función kernel, es decir, se necesitan metodologías eficientes para sintonizar los parámetros de inicialización de la SVM.^(49, 54)

1.9.4 Redes neuronales

Las redes neuronales han demostrado ser una buena aproximación a problemas donde el conocimiento de estos es impreciso o variante en el tiempo. Su capacidad de aprender convierte a las redes neuronales en algoritmos adaptativos y elaborados a la vez. Están inspiradas en el funcionamiento biológico de las neuronas, este tipo de algoritmos explotan el concepto de disponer de elementos simples de cómputo, las neuronas, interconectadas (sinapsis) de forma más o menos compleja. La idea central se encuentra en el concepto de neurona, a continuación se muestran sus características:

- Recibe un conjunto de señales de entrada procedentes de otras neuronas o de un sistema exterior.
- Estas señales de entrada se transmiten a través de unas conexiones que llevan asociados unos pesos.
- La combinación de señales de entrada y pesos es procesada para dar lugar a unas señales de salida.⁽⁵⁵⁾

Una red neuronal se determina por las neuronas y la matriz de pesos. Hay tres tipos de capas de neuronas: la capa de entrada, la capa oculta y la capa de salida. Un perceptrón multicapa está compuesto por una capa de entrada, una capa de salida y una o más capas ocultas; aunque se ha demostrado que para la mayoría de problemas bastará con una sola capa oculta. Esta red tiene la capacidad de facilitar salidas satisfactorias a entradas que el sistema no ha visto nunca en su fase de entrenamiento; además puede aproximar cualquier tipo de función o relación continua entre un grupo de variables de entrada y salida lo que convierte a las redes perceptrón multicapa en herramientas de propósito general, flexibles y no lineales.⁽⁵⁵⁾

2 CAPÍTULO II: MATERIALES Y MÉTODOS

En este acápite se detalla todo el proceso realizado para lograr el cumplimiento de los objetivos trazados como parte del trabajo de investigación en el cual se lograron identificar nuevas estructuras químicas con potencial actividad contra la forma amastigote de *Leishmania major*. El proceso comienza con la construcción muy detallada de una base de datos, la representación química de los compuestos y el cálculo de descriptores moleculares para obtener modelos, empleando técnicas de inteligencia artificial, capaces de identificar estos compuestos a través del cribado virtual de bases de datos.

2.1 Gestión de la data

Para la construcción de la base de datos se realizó una exhaustiva búsqueda bibliográfica en internet que incluyó la base de datos Pubchem, hasta encontrar varios artículos de revistas de prestigio internacional en los cuales se reportaban ensayos biológicos realizados para determinar la actividad antileishmaniásica de compuestos contra amastigotes de *Leishmania major*.^(25, 48, 56-70)

Se obtuvo una base de datos constituida por 120 compuestos probados frente a amastigotes de *Leishmania major*. La base de datos al comienzo tuvo 201 compuestos probados contra la forma amastigote de *Leishmania major* con sus valores de IC50. Se tuvo en cuenta la metodología llevada a cabo en los ensayos realizados y que existiera una igualdad entre compuestos activos e inactivos. Se determinó la actividad según el parámetro IC50, considerándose activos aquellos que presentaban un valor de $IC_{50} \leq 15 \mu M$ e inactivos los que tuvieran valores de $IC_{50} > 15 \mu M$. Se verificó que no hubiera estructuras duplicadas empleando el programa *ISIDA/QSPR* y la forma *EdiSDF v. 5.03*.

El curado estructural de la base de datos consistió en la eliminación de duplicados, sales y mezclas, compuestos organometálicos, isótopos y definiciones de grupos R, así como la adición de hidrógenos explícitos. Para ello se empleó el módulo *Standardizer* implementado en el programa ChemAxon.⁽⁷¹⁾ Finalmente, luego del

curado y estandarización la base de datos final contó con 120 compuestos con actividad probada contra amastigotes de *Leishmania major* de ellos 69 activos y 51 inactivos con sus respectivos valores de IC50 y otros datos de interés como la fórmula molecular, masa molar y el SMILE canónico correspondiente según su estructura química.

2.2 Estrategias computacionales

Los estudios QSAR (*Quantitative Structure-Activity Relationships*) constituyen una metodología que relaciona numéricamente estructuras (químicas) con sus actividades biológicas y reúnen un conjunto de técnicas computacionales relacionadas con diseño y visualización espacial de moléculas (también llamado *in silico*), cálculo de propiedades fisicoquímicas moleculares (Descriptores), bioinformática y estadística. Todo esto con el fin de hacer una predicción teórica de la actividad biológica que permita el diseño teórico de futuros nuevos fármacos evitando pasar por el proceso a prueba y error de síntesis orgánica.⁽³⁸⁾ A continuación, se representa el esquema general para el desarrollo de estudios QSAR:

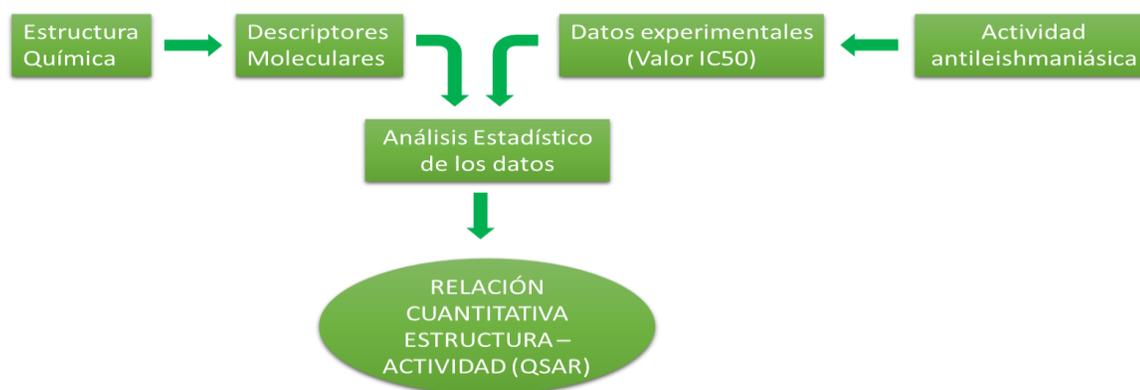


Figura 7. Esquema general para la obtención de un modelo QSAR.

Todos los cálculos se efectuaron en un ordenador (PC) Intel(R) Core™ i5-2500 CPU @ 3.30GHz. Se empleó el programa *DRAGON 7.0 Professional* ⁽⁷²⁾ para el cálculo de los DMs 0D-2D; se calcularon un total de 423 descriptores moleculares donde fueron excluidos los DMs 3D. Los datos se almacenaron en hojas de cálculo, utilizando la herramienta de trabajo *Microsoft Excel 2016* correspondiente al

paquete de herramientas *Microsoft Office Professional Plus 2016*; donde las filas representaron a los compuestos químicos (casos) y las columnas a los DMs (variables). Las matrices fueron guardadas como archivos de extensión .csv (Archivo de valores separados por comas de Microsoft Excel). El análisis estadístico se desarrolló con el paquete de procesamiento *WEKA 3.6* de 2009 y *WEKA 3.7* de 2011 para Windows en el cual se exportaron los archivos de extensión .csv y guardados como archivos de extensión .arff.

2.3 Parametrización de la estructura química

Las estructuras químicas de la base de datos (120 compuestos) fueron representadas usando el módulo *ChemDraw Ultra 12.0* del software *ChemBioOffice Ultra 12.0*, las cuales fueron guardadas como ficheros *.sdf del tipo MDL y correctamente curadas y estandarizadas utilizando el software *Standardizer* para eliminar cargas, añadir hidrógenos explícitos y remover sales con el objetivo de evitar interrupciones que pudiesen afectar el cálculo. Los ficheros obtenidos se ejecutaron en el programa *DRAGON 7.0* ⁽⁷²⁾ donde se calcularon los descriptores moleculares de 0D-2D, eliminando las variables constantes, cercanas a constantes y con una correlación mayor al 90%. Los datos fueron organizados y recogidos en matrices de *Microsoft Excel 2016* con los complementos *JChem para Office* e *Instant Jchem* que integran el manejo de estructuras y capacidad de visualización de las mismas. Los descriptores calculados fueron agrupados según las familias concebidas en el propio programa *DRAGON*, a saber: índices constitucionales, descriptores de anillos, índices topológicos, conteos de caminos y rutas (walk and path counts), índices de conectividad, índices de información, descriptores basados en la matriz 2D, autocorrelaciones 2D, Burden eigenvalues, descriptores tipo P_VSA, índices ETA, índices de adyacencia de bordes, conteo de grupos funcionales, fragmentos centrados en átomos, índices E-state tipo átomo, CATS-2D, pares de átomos 2D, propiedades moleculares e índices de similitud a medicamentos.

2.4 Análisis estadístico de los resultados

Con el objetivo de evaluar la distribución y diversidad existente dentro de los conjuntos de compuestos activos e inactivos respectivamente se realizó un Análisis de Conglomerados (AC) en el paquete de procesamiento *Statistica 8.0*. Luego se procedió a estandarizar todas las matrices de DMs anteriormente calculados. Posteriormente se dividieron en los compuestos activos e inactivos en diferentes subconjuntos mediante dos AC del tipo *k-MCA* (en inglés Análisis de Conglomerados de *k-medias*) conocido también como análisis de agrupaciones. De esta manera quedó asegurada la representatividad de elementos de similar dominio en los subconjuntos obtenidos en cada caso.

Para la confección de la serie de entrenamiento (SE), de predicción (SP) y de validación externa (VE) se seleccionaron de forma aleatoria los compuestos de cada conglomerado para construir las 3 series de la siguiente manera: 65% de la data inicial SE, 25% SP y 10% VE. Los compuestos incluidos en la SP no fueron usados en la determinación de los modelos cuantitativos. Para la selección de atributos se utilizó el software *WEKA 3.6* empleando varios evaluadores de atributos y métodos de búsqueda específicos según el algoritmo de clasificación. Los evaluadores de atributos empleados fueron *CfsSubsetEval* y *WrapperSubsetEval*. Los métodos de búsqueda empleados fueron: *BestFirst*, *EvolutionarySearch*, *GenericSearch*, *GreedyStepwise*, *LinearForwardSelection*, *RankSearch* y *SubsetSizeForwardSelection*. Los algoritmos de aprendizaje automático para clasificación empleados fueron *k*-vecino más cercano (IBk), árbol de clasificación (J48), red neuronal artificial (MLP por sus siglas en inglés *MultiLayer Perceptron*) y máquina de soporte vectorial (SMO por sus siglas en inglés *Sequential Minimal Optimization*)⁽⁷³⁾ para obtener los mejores modelos capaces de predecir la actividad de los compuestos incluidos en la SE y SP.

2.5 Obtención y validación de los modelos

Luego del cálculo de los DMs y de la selección de los mejores atributos se procedió a la obtención de los modelos de clasificación. Se cargaron los archivos guardados en formato *.arff en el software *Weka 3.6* y se seleccionaron las variables que mejor

Materiales y Métodos

describían a cada modelo y se clasificó la serie de entrenamiento de cada modelo para calcular la probabilidad de los compuestos en investigación de ser activos. Para elegir la cantidad máxima de variables en el modelo (dimensión óptima) se consideró que el modelo alcanzara una tasa de falsos positivos mínima en la serie de predicción interna y que la exactitud lograda en la serie de entrenamiento no estuviera acompañada de una disminución marcada de la exactitud en la serie de predicción interna. Los mejores atributos fueron seleccionados según la cantidad de veces que se repetían para los diferentes métodos de búsqueda empleados; para ello se utilizaron técnicas de programación en *JAVA* empleando la herramienta *Python*. Los modelos finales se obtuvieron empleando los atributos seleccionados teniendo en cuenta el principio de parsimonia o Razón de *Ockham*: “*Numquam ponenda est pluralitas sine necessitate*” o “La pluralidad nunca se debe postular sin necesidad de ello”.

Para seleccionar el mejor modelo por cada técnica, el mismo debía tener la mayor significación estadística con la menor cantidad de variables posible pues cuando las explicaciones múltiples están disponibles para un fenómeno, se prefiere la versión más simple.^(74, 75) Para la selección de los mejores modelos se emplearon técnicas de programación en *JAVA* empleando la herramienta *Python* que permitió determinar la mejor combinación de los atributos seleccionados para cada técnica de clasificación. Se calcularon los parámetros estadísticos más utilizados en la literatura médico-estadística: exactitud o porcentaje de buena clasificación (**Q**) en cada uno de los grupos, coeficiente de correlación de *Matthews* (**C**), sensibilidad (**Sens**), especificidad (**Espec**) y razón de falsos positivos (**FPr**), lo que nos permitió evaluar el desempeño de los clasificadores.

$$Q = 100 * (VP + VN) / (VP + FP + VN + FN)$$

$$FPr = 100 * FP / (FP + VN)$$

$$\text{Sensibilidad (Sens)} = 100 * VP / (VP + FN)$$

$$\text{Especificidad (Espec)} = 100 * VN / (VN + FP)$$

$$C = 100 * (VP * VN - FP * FN) / \sqrt{(VN + FN) * (VN + FP) * (VP + FP) * (VP + FN)}$$

Donde VP y VN son los verdaderos positivos y negativos, respectivamente y FP y FN son los falsos positivos y negativos, correspondientemente.

Para demostrar el poder predictivo de los modelos se empleó la serie de predicción obtenida mediante el análisis de conglomerados descrito en el epígrafe 2.4.⁽⁵⁰⁾ En cada caso se determinaron los parámetros estadísticos antes mencionados los cuales permitieron promover la aceptación de estos modelos QSAR. También fue empleada una serie de validación externa que incluyó un grupo de 18 compuestos (10 activos y 8 inactivos) seleccionados al azar a partir de la base de datos primaria de la cual constituyen el 10%, los cuales nunca fueron empleados en el desarrollo de los modelos y cumplían los requerimientos establecidos, por lo que permitirían validar el poder predictivo de los modelos obtenidos para amastigotes. Por otro lado, se realizó la validación interna mediante validación cruzada *10 fold cross validation* (en esta la validación cruzada de 10 iteraciones los datos se dividen en 10 subconjuntos (*fold*s). Uno de los subconjuntos se utiliza como datos de prueba y el resto (10-1) como datos de entrenamiento. El proceso de validación cruzada es repetido durante 10 iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. El error se calcula como la media aritmética de los errores de cada iteración para obtener un único resultado.

2.6 Cribado virtual y sistema multclasificador de ensamblado basado en modelos QSAR

El cribado virtual se realizó a diferentes compuestos de origen natural, de síntesis, fármacos experimentales y el Cuadro Básico de Medicamentos de Cuba (444 compuestos), una base de datos de fármacos licenciados.⁽⁷⁶⁾ En este trabajo de investigación fueron empleados además para el cribado virtual 468 compuestos de síntesis obtenidos en laboratorios de Química Orgánica, de ellos 407 de la Universidad de Rostock, Alemania y 61 del Conservatorio Nacional de Artes y Oficios (CNAM) de París, Francia; los cuales poseen una amplia variabilidad estructural. Además, fueron empleados 444 compuestos correspondientes al Cuadro Básico de Medicamentos de Cuba y 33 compuestos del Departamento de Ciencias Farmacéuticas de la Universidad Central Martha Abreu de Las Villas, Cuba.

Materiales y Métodos

Tras el cribado se procedió a la identificación, ensamblado, análisis y reporte de todas las estructuras de los compuestos identificados como potencialmente activos contra amastigotes de *L. major*, los cuales podrán ser empleados en futuras investigaciones relacionadas con el descubrimiento de nuevos compuestos para el tratamiento de las leishmaniasis causadas por esta especie. Para el ensamblado se empleó un sistema multclasificador ensamblado a partir de la combinación adecuada de varios clasificadores con el fin de mejorar los resultados a obtener.⁽⁷⁷⁾ En esta técnica a cada clasificador se le otorga un voto de igual valor y por el voto de la pluralidad se escogen los compuestos predichos como activos por la mayoría de los modelos. Las probabilidades generadas para cada compuesto por los modelos obtenidos por familias fueron fusionadas con el objetivo de incrementar la calidad en las predicciones y/o aumentar el espacio químico a ser empleado del multclasificador que se utilizaría para las predicciones finales.

3 CAPÍTULO III: RESULTADOS Y DISCUSIÓN

Durante este capítulo se presentarán los resultados del trabajo de investigación que permitieron la construcción de los modelos a partir del empleo de técnicas de Inteligencia Artificial y a su vez la identificación de nuevos compuestos potencialmente activos contra amastigotes de *Leishmania major*.

3.1 Gestión de la data

De un total de 200 compuestos obtenidos a partir de artículos de revistas encontrados en la Web se seleccionaron 120 (Anexo 1) con un alto grado de variabilidad estructural para la construcción de la base de datos final. La base de datos puede considerarse como pequeña ya que, en otros estudios realizados como Flores y colaboradores 2015, Rojas Socarrás 2017 y Salgado Vernier 2018 se construyeron bases de datos de 365, 437 y 336 compuestos respectivamente. La relativa pequeñez de la base de datos se debe a la escasa aplicación de ensayos biológicos en amastigotes de *L. major*, además de que existían reportes principalmente de compuestos activos y pocos compuestos inactivos. No obstante, se aseguró una variabilidad estructural de compuestos que van desde tiazoles, pirazolonas, ácidos carboxílicos y aminas hasta fenoles, cicloalcanos, compuestos insaturados, compuestos con grupos esteroidales y compuestos heterocíclicos substituidos, disubstituidos y trisubstituidos. El parámetro empleado para la clasificación de los compuestos fue la IC₅₀ quedando constituida la base de datos por 69 compuestos activos y 51 inactivos. Al haber sido probada con anterioridad la actividad antileishmaniásica de estos compuestos con procedimientos muy afines son más confiables los resultados a obtener evitando la futura clasificación “errónea” de compuestos.

De la fase preparatoria de las bases de datos consideramos interesante resaltar la importancia del curado estructural y de actividad. La minuciosa identificación y eliminación de compuestos repetidos y la homogeneización de la representación de las subestructuras resultó un paso clave. Luego de la estandarización y curación de la base de datos, se calcularon en el software *DRAGON 7.0* las diferentes familias

Resultados y Discusión

de DMs de 0-2D obteniéndose 4544 descriptores que fueron reducidos a 423 variables tras la remoción de aquellas constantes, cercanas a constantes y con valores perdidos, con una correlación de 0.85 y varianza menor de 5 %.⁽⁷²⁾ Algunos de los parámetros implementados en el software e incluidos para el cálculo fueron las propiedades moleculares, descriptores de búsqueda y análisis subestructural, conteo de fragmentos, huellas digitales de frecuencia e índices de conectividad, entre otros.

Para asegurar la adecuada diversidad estructural de la data y por consiguiente la calidad de los métodos QSAR se realizó un Análisis de Conglomerados (AC) implementado en el paquete de procesamiento *STATISTICA 8.0* del tipo *k*-MCA (siglas en inglés de Análisis de Conglomerados de *k*-Medias, también llamado análisis de agrupaciones o clúster) (Ver Anexo 2). De este modo quedó asegurada la representatividad de elementos en las diferentes series obtenidas. De cada conglomerado se seleccionaron, de forma aleatoria, los compuestos para la conformación de la serie de entrenamiento (65%), predicción (20%) y de validación externa (15%), quedando incluidos en ellas 78, 24 y 18 compuestos respectivamente. El procedimiento empleado se refleja en la figura 8, Anexo 5.4:

3.2 Parametrización de la estructura química. Estrategias computacionales. Análisis estadístico de datos

Luego de la selección de atributos mediante el empleo de diferentes evaluadores y métodos de búsqueda específicos en el software *Weka 7.3*, el resultado obtenido permitió una reducción de las variables más adecuadas para cada técnica, obteniéndose de un total de 423 descriptores moleculares calculados seis variables para IBk, J48 y MLP respectivamente. Se obtuvieron además ocho variables para otro modelo MLP. Elegidas ya las variables adecuadas para cada técnica, se emplearon las mismas para la elaboración de los ficheros de extensión .arff en el *Weka 7.0* para la conformación de los modelos de clasificación finales.

Resultados y Discusión

Cabe mencionar que tras una exhaustiva revisión bibliográfica no se encontraron en la literatura estudios acerca del empleo de las técnicas de Inteligencia Artificial para el desarrollo de modelos de clasificación con especies de *Leishmania*. Por ende se utilizaron tesis de diplomado y maestrías respectivamente realizadas en nuestro centro de estudios “Universidad Marta Abreu de Las Villas” como referencia para la comparación de la capacidad predictiva de los modelos obtenidos.^(7, 78, 79)

Con el empleo de técnicas de clasificación se logró un grupo de modelos con distinto grado de calidad y valor predictivo según las familias de descriptores. De modo general se obtuvieron modelos de seis y ocho variables con una exactitud en la predicción de la serie interna de 87.5% - 83.3%, una tasa de falsos positivos de la serie de entrenamiento menor al 27% y un coeficiente de Matthews de 1.0-0.69. Los algoritmos y parámetros estadísticos correspondientes a los modelos obtenidos se describen a continuación:

Para el modelo **IBk** de seis variables obtenido se utilizó el siguiente algoritmo de clasificación de `weka.classifiers.lazy.IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""` donde *K* (clasificador *K*-vecinos más cercanos) selecciona el valor apropiado de vecinos basado en la validación cruzada y *W* (*windowSize* o en español tamaño de ventana) obtiene el número máximo de instancias permitidas en la serie de entrenamiento, un valor de cero significa que no hay límite para el número de instancias de formación. El modelo clasificó correctamente un 100% de las instancias para la serie de entrenamiento y un 87.50% para la serie de predicción con un área bajo la curva ROC de 1.00 y 0.76 para ambas series respectivamente. En la siguiente tabla se observan los resultados obtenidos para el modelo IBk con respecto a la sensibilidad, especificidad, coeficiente de correlación de Matthews, relación de falsos positivos y exactitud para las series de entrenamiento y predicción, cabe destacar los resultados perfectos de la serie de entrenamiento los cuales indican un sobreentrenamiento para la primera serie.

En 2015 Flores y colaboradores obtuvieron un modelo IBk con algoritmo de clasificación para Weka: `weka.classifiers.lazy.IBk -K3- W0-A weka. core.`

Resultados y Discusión

neighboursearch. LinearNNSearch– A"weka. core. ManhattanDistance-R first-last"- el cual incluyó siete variables a partir de una base de datos de compuestos evaluados contra *Leishmania amazonensis*. Este clasificó correctamente un 83,16% de las instancias para la SE y un 85,25% para la SP, estos resultados son similares a los obtenidos por nuestro modelo, aunque los nuestros lo superan para la SE.⁽⁷⁹⁾ En 2015 Álvarez González, OR realizó un modelo IBk para siete variables con el siguiente algoritmo de clasificación: -K 2 -W 0 -A neighboursearch.LinearNNSearch -A ManhattanDistance -R first-last obteniendo una clasificación correcta de un 89.53% de las instancias para la serie de entrenamiento y un 83.33 % para la serie de predicción, siendo resultados muy similares a los obtenidos en nuestro estudio.⁽⁷⁸⁾ Además, en 2017 Rojas Socarrás también realizó un modelo IBk para seis variables a partir de una base de datos de compuestos contra amastigotes de *Leishmania infantum* con el siguiente algoritmo de clasificación: [WEKA.classifiers.lazy. IBk -K3 -W0 -A WEKA.core.neighboursearch.LinearNN Search-A"WEKA.core.ManhattanDistance-R first-last"-t. El modelo clasificó correctamente un 88,81% de las instancias para la serie de entrenamiento y un 85,05% para la serie de predicción.⁽⁷⁾ Los resultados de estos estudios son similares a los obtenidos en el presente estudio, aunque poseen mejores resultados de especificidad, exactitud y FPr. Los resultados obtenidos en el estudio de especificidad, Área Bajo la Curva, exactitud y FPr son menores como consecuencia del tamaño de la base de datos confeccionada, ya que, a menor número de compuestos de la base de datos, mayor probabilidad de ocurrir errores en la clasificación de los modelos. Los atributos comprendidos en el modelo obtenido en este trabajo incluyen descriptores de anillo, índices topológicos, autocorrelaciones 2D, conteo de grupos funcionales entre otros.

El modelo de clasificación de árbol (**J48**) presentó seis variables y el siguiente esquema algorítmico: weka.classifiers.trees.J48 -C 0.19 -M 2 teniendo que C (*confidenceFactor*: factor de confianza utilizado para la poda es de 0,19, valores menores incurren en un incremento de la poda, y M (número mínimo de instancias por hoja) es de 2 (Figura 10, anexo 5.6). El modelo mostró un 93.59% (SE) y 83.33%

Resultados y Discusión

(SP) correctamente clasificados, y un Área Bajo la Curva ⁽⁵⁰⁾ de 0.88 para la SE y 0.68 para la SP.

A modo de comparación se presenta una tabla resumen de los resultados del modelo de clasificación J48 realizado por otros autores donde se muestran resultados inferiores a los obtenidos en este estudio. ^(7, 77,78, 79)

Tabla 4. Recopilación de los resultados de parámetros estadísticos para el modelo de clasificación árbol J48 para estudios anteriormente realizados.

Autor (es)	# de variable	Serie	C	Q (%)	Sensibilidad (%)	Especificidad (%)	FPr (%)
Rojas Socarrás 2017	7	SE	0,76	87,76	80,77	91,30	6,41
		SP	0,68	84,11	80,39	85,42	12,50
Álvarez González 2015	8	SE	0.81	89.53	100	81.25	19.15
		SP	0.68	83.33	92.86	76.47	25
Flores Balmaseda 2014	8	SE	0.69	82.56	97.44	73.08	29.79
		SP	0.55	76.67	85.71	70	31.25

Leyenda: Q exactitud, C coeficiente de correlación de Matthews, FPr Razón de falsos positivos.

Se obtuvieron dos modelos de clasificación MLP con las mismas seis variables, pero con cambios en el algoritmo de clasificación. En cuanto a los resultados obtenidos para las redes neuronales MLP (*Multi-Layer Perceptron* por sus siglas en inglés o perceptrón multicapas) mostró mejores resultados en cuanto a número de instancias clasificadas correctamente con un 89.74% y 84.62% en la serie de entrenamiento y un 83.33 % para la serie de predicción respectivamente. Además, se obtuvieron valores de Área Bajo la Curva de 0.80 y 0.69 para la SE respectivamente. Resultó relevante la obtención de un 97.44 % de sensibilidad y un 3.03 % de FPr para el primer modelo de clasificación MLP. El esquema algorítmico empleado para los diferentes modelos MLP fue: `weka.classifiers.functions.MultilayerPerceptron -L 0.7 -M 0.8 -N 500 -V 0 -S 0 -E 20 -H 6` y `weka.classifiers.functions.MultilayerPerceptron -L 0.8 -M 0.6 -N 500 -V 0 -S 0 -E 20 -H 9` respectivamente; siendo L (*LearningRate* o intervalo de incremento) que constituye la suma actualizada de los pesos, en este caso 0.7 y 0.8, H (*hiddenLayers* o capas ocultas), define las capas ocultas de la red

Resultados y Discusión

neuronal conformando una lista de números enteros positivos uno para cada capa oculta. Se mostraron valores satisfactorios en los diferentes parámetros estadísticos obtenidos los cuales se representan en la tabla 5 (anexo 5.7) entre los que resaltan los valores de FPr y especificidad de la primera variante MLP.

Los resultados obtenidos por otros autores se observan en la tabla 6, donde se observan los resultados inferiores de Rojas Socarrás 2017 en los parámetros estadísticos obtenidos. (7,77, 78, 79)

Tabla 6: Parámetros estadísticos obtenidos para modelo MLP (Perceptrón Multicapas).

Autor (es)	# de variables	de Serie	C	Q (%)	Sensibilidad (%)	Especificidad (%)	FPr (%)
Rojas Socarrás 2017	7	SE	0,66	83,22	79,23	83,06	13,46
		SP	0,67	83,18	74,51	88,37	8,93
Flores Balmaseda 2014	7	SE	0.84	91.86	87.18	94.44	4.26
		SP	0.64	80	57.14	100	0

Leyenda: Q exactitud, C coeficiente de correlación de Matthews, FPr Razón de falsos positivos.

Los histogramas de los atributos para la serie de entrenamiento se muestran en la figura 12, anexo 5.7 estos incluyeron descriptores de autocorrelación 2D, contador de grupos funcionales, pares de átomos 2D y fragmentos atómicos centrados.

La técnica máquina vectorial de soporte (**SMO**) reflejó un modelo constituido por ocho variables y utilizó un esquema algorítmico: `weka.classifiers.functions.SMO -C 20.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K weka.classifiers.functions.SMO -C 20.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.RBFKernel -C 250007 -G 1.7" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"` en la cual C es el parámetro de complejidad y tiene un valor de 250007, y se emplea un incremento numérico de -1 que nos indica el número de pliegues de validación cruzada utilizados para generar datos de entrenamiento para modelos logísticos y G es el valor gamma, en este caso es 1,7. La SE reportó un 88.46% de instancias clasificadas correctamente como activas e inactivas con respecto al total y la serie

Resultados y Discusión

de predicción un 87.50% con un área bajo la curva ROC de 0.78 para la SE y 0.76 para la SP respectivamente. El resto de los parámetros estadísticos quedan plasmados en la tabla 7, anexo 5.8, resaltando los altos valores de especificidad para ambas series.

Los histogramas de los atributos incluidos en este modelo se muestran en la figura 13, anexo 5.8:

La máquina vectorial de soporte (SMO) con algoritmo [-C 90.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K.RBFKernel -C 250007 -G 0.9] para ocho variables obtenida por Flores y colaboradores mostró un número de instancias clasificadas correctamente de 91.86 % en la SE y 86.67 % para la SP, resultados ligeramente superiores a los obtenidos en nuestro estudio.⁽⁷⁹⁾ En los resultados obtenidos por Rojas Socarrás (2017),⁽⁷⁾ se utilizó el esquema logarítmico WEKA.classifiers.functions.SMO -C50.0 -L0.001 -P1.0E-12 -N0 -V-1 -W 1- K WEKA.classifiers. functions.supportVector. RBFKernel -C 250007-G 3.9 para siete variables, alcanzado un 81,47% de clasificación correcta para la SE y 78.50% para la SP; estos resultados reflejan cierta analogía con los obtenidos en el presente estudio para el modelo de clasificación SMO.

Resumiendo los resultados obtenidos para la SE y SP de los diferentes modelos de clasificación empleados en el presente estudio, se presenta el gráfico #1, anexo 5.9. Se puede observar que el modelo de mejores resultados fue el J48 con 93.59% y 83.33 % de clasificación correcta de la SE y SP respectivamente. El modelo con los resultados más discretos fue el IBk (100% SE y 87.5% SP).

3.3 Validación de los modelos obtenidos

Con el objetivo de evaluar la robustez y poder predictivo de los modelos obtenidos se evaluaron las SP correspondientes para amastigotes de *L. major*, logrando valores adecuados de los parámetros estadísticos obtenidos, los cuales se observan en la tabla 8 para cada modelo:

Tabla 8. Parámetros estadísticos obtenidos para la SP de los modelos

Resultados y Discusión

Modelo	Serie	C	Q	Sensib	Specif	False Postive Rate
IBK	Predicción	0.76	87.50	100.00	82.35	30.00
J48	Predicción	0.68	83.33	78.57	91.67	10.00
MLP v.1	Predicción	0.68	83.33	78.57	91.67	10.00
MLP v.2	Predicción	0.68	83.33	78.57	91.67	10.00
SMO	Predicción	0.76	87.50	100.00	82.35	30.00

Leyenda: Q exactitud; C coeficiente de correlación de Matthews; Sensib Sensibilidad; Specif Especificidad; FPR Razón de falsos positivos.

La validación externa realizada para los cinco modelos lanzó resultados similares a los obtenidos por Rojas Socarrás en 2017 e inferiores a Flores Balmaseda 2015.^(7, 78-80) En lo que a la validación cruzada se refiere, la tabla 9, anexo 5.9 recoge los parámetros estadísticos obtenidos por cada modelo y en el gráfico 2, anexo 5.9 se visualizan mejor estos resultados donde se define que el mayor porcentaje de clasificación se logró con el modelo SMO. Los valores de exactitud que se obtienen resultan mayores para el árbol de clasificación J48 y para la máquina vectorial de soporte, el valor de falsos positivos de manera general resultó ser alto para todos los modelos con valores superiores al 12 %.

Los valores de la validación cruzada resultaron superiores que a los de la validación externa con valores del coeficiente de correlación de Matthews superiores a 0.56 y valores de la razón de falsos positivos inferiores al 42 % donde los resultados más bajos se le atribuyen al modelo de red neuronal variante 2 con un 12.12%.

3.4 Cribado virtual de base de datos

Para la identificación de nuevos compuestos con potencial actividad contra amastigotes de *Leishmania major* se evaluaron un total de 1501 compuestos de diferente origen que incluía el Cuadro Básico de Medicamentos de Cuba, compuestos de síntesis y de origen natural obteniéndose como resultado una amplia variedad estructural por cada uno de los modelos. El modelo IBK clasificó la mayor cantidad de compuestos como activos con 634 compuestos lo que representa el 64.1 % del total de compuestos cribados. Los compuestos clasificados como activos e inactivos por cada modelo individual puede observarse en el gráfico 3, anexo 6.

Resultados y Discusión

El resultado del cribado teniendo en cuenta los porcentajes de compuestos identificados como activos e inactivos por los modelos puede apreciarse en el gráfico 4 anexo 6. En él se observa que el 11% de los 989 compuestos empleados para el cribado es identificado como activos por la totalidad de los modelos (114), un 25% por tres modelos (266), 23% activos por al menos dos (223) y un 11% por un solo modelo (109), mientras que el 3% del total es hallado inactivo por todos ellos (32). El ensamblado de voto mayoritario simple de todos los modelos obtenidos es el elegido para la identificación final de los compuestos potencialmente activos contra amastigotes de *Leishmania major*, este permite reducir los errores individuales de cada modelo por separado y realizar un análisis más exhaustivo que tiene en cuenta el algoritmo individual incluida en cada técnica de inteligencia artificial. La totalidad de estos compuestos identificados como potencialmente activos pueden ser evaluados experimentalmente para corroborar su actividad contra amastigotes de *Leishmania major* en posteriores estudios.

4 Conclusiones

1. Se confeccionó una base de datos que incluyó 120 compuestos evaluados experimentalmente contra la forma amastigote del parásito *Leishmania major* con una amplia variabilidad estructural.
2. Se obtuvieron cinco modelos QSAR desarrollados con técnicas de inteligencia artificial (IBk, J48, MLP y SMO) alcanzando buenos porcentajes de clasificación, sensibilidad y especificidad tanto para las series de entrenamiento como las de predicción.
3. La validación de los modelos obtenidos para amastigotes demostró su alto grado de predictibilidad.
4. Se identificaron 114 compuestos con potencial actividad contra *Leishmania major* a través del cribado virtual empleando el ensamblado de los modelos obtenidos.

Recomendaciones

1. Realizar un ensamblado con los modelos obtenidos con anterioridad para amastigotes de *Leishmania major* y los alcanzados en el presente estudio para reducir a través del cribado virtual el número de compuestos a evaluar en ensayos *in vitro* e *in vivo*
2. Evaluar experimentalmente los compuestos identificados como potencialmente activos para corroborar su actividad.
3. Trabajar en la obtención de nuevos modelos por otras técnicas que permitan complementar los resultados obtenidos hasta el momento

Referencias Bibliográficas

Referencias Bibliográficas

1. Montalvo AM, Fraga J, Monzote L, García M, Fonseca L. Diagnóstico de la leishmaniasis: de la observación microscópica del parásito a la detección del ADN. *Revista Cubana de Medicina Tropical*. 2012;64:108-31.
2. (OMS) OMDIS. Leishmaniasis 2019 [updated 14/3/2019].
3. Jiménez Camacho DC. Leishmaniasis. *Revista Médica de Costa Rica y Centroamerica* 2015;614:129-33.
4. Chongo Alfaro ML, García Echegoyen R. Leishmaniasis y transfusión. Artículo de revisión. *Revista Mexicana de Medicina Transfusional*. 2010;3:842-7.
5. Salud SDPYPDL. Manual para el diagnóstico, tratamiento y control de la leishmaniasis. In: *Enfermedades CNdPPyCd*, editor. México 2015.
6. [Available from: <https://www.uniprot.org/proteomes/UP000000542>].
7. Rojas Socarrás S. Empleo de estudios “*in silico*” en la identificación de nuevos compuestos con potencial actividad contra *Leishmania infantum*. Santa Clara: Universidad Central Marta Abreu de Las Villas; 2017.
8. Piccolo Johanning L, Pérez Elizondo E, Álvarez Morales L, Wang Zúñiga C, Sancho Torres M. Leishmaniasis: Opciones terapéuticas en la población pediátrica. *Medicina Legal de Costa Rica Edición Virtual*. 2018;35.
9. Torres-Guerrero E, Arenas R. Leishmaniasis. Alternativas terapéuticas actuales. *Revista Mexicana de Dermatología*. 2018;62:400-9.
10. Shiff C, Hudson Keenihan S, Schettters T, Taverne J, Ramster B. *Leishmania major*: un parásito compasivo? *TRENDS in Parasitology*. 2001;17(8).
11. Montalvo Álvarez AM. Leishmaniasis. Aspectos de interés sobre un parasitismo exótico para Cuba. *Revista Cubana de Higiene y Epidemiología*. 2010;48(1):78-92.
12. Kumar A. *Leishmania and Leishmaniasis*. Springer. 2013.
13. <https://www.who.int/es/news-room/fact-sheets/detail/leishmaniasis> 2016 [
14. Amela C, Suarez B, Isidoro B, Sierra J, Santos S, Simón F. Evaluación del riesgo de transmisión de *Leishmania infantum* en España. 2012.
15. Cardona Sánchez D. Moléculas bioactivas contra *Leishmania (viannia) panamensis*. Actividad y optimización molecular. 2006.
16. Montalvo Alvarez A. Tipificación de especies de *Leishmania* de importancia médica basada en el gen que codifica la proteína HSP70 citoplasmática. La Habana: Instituto de Medicina Tropical “Pedro Kourí”; 2011.
17. González C, Cabrera O, Munstermann L, Ferro C. Distribución de los vectores de *Leishmania infantum* (Kinetoplastida: Trypanosomatidae) en Colombia. *Ciencias Biomédicas*. 2006;26:64-72.
18. Feliciangeli M. Sobre los flebótomos (Diptera: Psychodidae: Phlebotominae), con especial referencia a las especies conocidas en Venezuela. 2006;26(2):61-80.
19. Miraballes Ferrer M. *Leishmaniasis: Enfermedad emergente en Uruguay*. Uruguay: Universidad de la República; 2013.
20. Salud OMDI. Control de las leishmaniasis. Ginebra; 2010.
21. Moreno Mateos D. Estrategias celulares y moleculares en el cribado de nuevos agentes anti-leishmaniásicos. Madrid: Universidad de Alcalá; 2010.

Referencias Bibliográficas

22. Ríos Yuil JM, Sousa O. Inmunología en la infección por *leishmania*: Conceptos actuales. Revista Médico Científica. 2011;23(1):19-31.
23. Sifontes Rodríguez S. Descubrimiento racional de compuestos químicos para el tratamiento de la leishmaniosis y la enfermedad de Chagas. Madrid: Universidad Complutense de Madrid; 2017.
24. Prevention CCfDCa. Laboratory Identification of Parasitic Diseases of Public Health Concern. Leishmaniasis - Lyfe Cycle 2015 [
25. Voak AA, Gobalakrishnapillai V, Seifert K, Balczó E, Hu L, Hall BS, et al. An Essential Type I Nitroreductase from *Leishmania major* Can Be Used to Activate Leishmanicidal Prodrugs. THE JOURNAL OF BIOLOGICAL CHEMISTRY. 2013;288(40):28466–76.
26. Pérez Pérez F. Actualidad y perspectivas en el tratamiento tópico y local de la Leishmaniasis Cutánea.
Autor:
Tutor: . Madrid: Universidad Complutense de Madrid; 2015.
27. Ortiz Gómez A. Farnesil difosfato sintetasa de *Leishmania major*: Estudios de una potencial diana terapéutica. Granada: Instituto de Parasitología y Biomedicina “López- Neyra”; 2005.
28. García Almagro D. Leishmaniasis cutánea. Actas Dermo-Sifiliográficas. 2005;96(1).
29. Leishmaniasis (cutánea y visceral). In: Health TCfFSP, editor. Louisiana2009.
30. Martínez-Silva E, Ruiz de Ramos Y, Bastidas-Pacheco G. Vacunas contra *Leishmania*. Iatreia. 2016;29(2).
31. Späth G, al. e. Persistence without pathology in phosphoglycan-deficient *Leishmania major*. Revista Panameña de Salud Pública. 2003;301:1241-9.
32. Ensuncho AE, Robles JR, Figueredo SF. Modelado de las Relaciones Cuantitativas Estructura-Actividad (QSAR) de los Derivados 5-(Nitroheteroaril) - 1,3,4-tiadiazol con Actividad Leishmanicida. Información Tecnológica. 2017;28(2):191-202.
33. Resolution of an Infection with *Leishmania braziliensis* Confers Complete Protection to a Subsequent Challenge with *Leishmania major* in BALB/c Mice [Available from: https://www.researchgate.net/publication/13259508_Resolution_of_an_Infection_with_Leishmania_braziliensis_Confers_Complete_Protection_to_a_Subsequent_Challenge_with_Leishmania_major_in_BALBc_Mice.
34. Sohrabi Y, Lipoldová M. Mannose Receptor and the Mystery of Nonhealing *Leishmania major* Infection. Trends in Parasitology. 2018;1747:1-3.
35. Del Cañizo López J, López Martín D, Lledó García E, García Barreno P. Diseño de modelos experimentales en investigación quirúrgica. ACTAS UROLÓGICAS ESPAÑOLAS 2008;32(1):13.
36. Lorena Fina B, Lombarte M, Rigalli A. Investigación de un fenómeno natural: ¿estudios *in vivo*, *in vitro* o *in silico*? Actualizaciones en Osteología. 2013;9(3):294-9.
37. Saldívar González F, Prieto Martínez F, Medina-Franco JL. Descubrimiento y desarrollo de fármacos: un enfoque computacional. Educación Química. 2017;28:51-8.

Referencias Bibliográficas

38. Lozano Aponte J, Scior T. ¿Qué sabe Ud. acerca de...QSAR? Revista Mexicana de Ciencias Farmacéuticas. 2012;23(2).
39. Morales Rodriguez Y. Aplicación de métodos "In Silico" para la predicción de permeación de compuestos químicos. Santa Clara: Universidad Central Marta Abreu de Las Villas; 2007.
40. Consonni V, Todeschini R. Molecular Descriptors. 2010. In: Recent Advances in QSAR Studies Methods and Applications [Internet]. New York; [29-102].
41. Mauri A, Consonni V, Todeschini R. Molecular Descriptors. Milano2015.
42. Mauri A, Consonni V, Pavan M, Todeschini R. Dragon software: An easy approach to molecular descriptor calculations. MATCH Communications in Mathematical and in Computer Chemistry. 2006;56:237-48.
43. Calleja Gómez AJ. Minería de datos con weka para la predicción del precio de automóviles de segunda mano. Valencia: UNIVERSIDAD POLITÉCNICA DE VALENCIA; 2010.
44. García Gutiérrez JA. Comenzando con Weka: Filtrado y selección de subconjuntos de atributos basada en su relevancia descriptiva para la clase. Madrid: University of Malaga; 2016.
45. Hernández Orallo J, Ferri Ramírez C. Curso de Doctorado Extracción Automática de Conocimiento en Bases de Datos e Ingeniería del Software. Valencia: Universitat Politècnica de València; 2006.
46. García Morate D. Manual de Weka. 2013.
47. Bunin BA, Bajorath J, Siesel B, Morales G. Chemoinformatics: Theory, practice, & products. Netherlands: Springer; 2007.
48. Software solutions and services for chemistry & biology 2019 [Available from: <https://chemaxon.com/>].
49. Benitez R, Escudero G, Kanaan S, Masip Rodo D. Inteligencia Artificial Avanzada. Barcelona2014.
50. Ponce Gallegos JC, Torres Soto A, Quezada Aguilera FS, Silva Sprock A, Martínez Flor EU, Casali A, et al. Inteligencia Artificial. México2014.
51. Russell S, Norvig P. Inteligencia artificial. Un enfoque moderno. Madrid2012.
52. García Cambroner C, Gómez Moreno I. Algoritmos de aprendizaje: Knn & kmeans. In: Madrid UCId, editor. Madrid2009.
53. Solarte Martínez GR, Ocampo CA. Técnicas de clasificación y análisis de representacion del conocimiento para problemas de diagnóstico. Scientia et Technica Año XV. 2009;42.
54. Carmona Suárez EJ. Tutorial sobre Máquinas de Vectores Soporte (SVM). In: (UNED) UNdEaD, editor. Madrid2014.
55. Sánchez Anzola N. Máquinas de soporte vectorial y redes neuronales artificiales en la predicción del movimiento USD/COP spot intradiario. ODEON. 2015(9):113-72.
56. National Center for Biotechnology Information. PubChem BioAssay Database; AID=1255041, <https://pubchem.ncbi.nlm.nih.gov/bioassay/1255041> (accessed Feb. 18, 2019). 2019.
57. Asgharpour Sarouey L, Khanaliha K, Rahimi-Moghaddam P, Khorrami S, Saaid Dayer M, Tabatabaie F. In vitro Effects of Ketotifen and Cromolyn Sodium on Promastigote and Amastigotes of Leishmania major. Jundishapur Journal of Microbiology. 2018.

Referencias Bibliográficas

58. BERMAN JD, KING M, EDWARDS N. Antileishmanial Activities of 2,4-Diaminoquinazoline Putative Dihydrofolate Reductase Inhibitors. *Antimicrobial Agents And Chemotherapy*. 1989;33(11):1860-3.
59. Bettiol E, Samanovic M, Murkin AS, Raper J, Buckner FS, Rodriguez A. Identification of Three Classes of Heteroaromatic Compounds with Activity against Intracellular *Trypanosoma cruzi* by Chemical Library Screening. *Chemical Library Screening*. 2009;3(2).
60. Bringmann G, Thomale K, Bischof S, Schneider C, Schultheis M, Schwarz T, et al. A Novel *Leishmania major* Amastigote Assay in 96-Well Format for Rapid Drug Screening and Its Use for Discovery and Evaluation of a New Class of Leishmanicidal Quinolinium Salts. *Antimicrobial Agents and Chemotherapy*. 2013;57(7):3003–11.
61. Buckner FS, Wilson AJ. Colorimetric assay for screening compounds against leishmania amastigotes grown in macrophages. *American Society of Tropical Medicine and Hygiene*. 2005;72(5):600-5.
62. Firdessa R, Good L, Amstalden MC, Chindera K, Kamaruzzaman NF, Schultheis M, et al. Pathogen- and Host-Directed Antileishmanial Effects Mediated by Polyhexanide (PHMB). *PLOS Neglected Tropical Diseases* 2015.
63. Gamage SA, Figgitt DP, Wojcik SJ, Ralph RK, Ransijn A, Mauel J, et al. Structure-Activity Relationships for the Antileishmanial and Antitrypanosomal Activities of 1'-Substituted 9-Anilinoacridines. *Journal of Medicinal Chemistry*. 1997;40(16).
64. Glaser J, Schultheis M, Hazra S, Hazra B, Moll H, Schurigt U, et al. Antileishmanial Lead Structures from Nature: Analysis of Structure-Activity Relationships of a Compound Library Derived from Caffeic Acid Bornyl Ester. *Molecules*. 2014;19.
65. Khraiweh M, Leed S, Roncal N, Johnson J, Sciotti R, Smith P, et al. Antileishmanial Activity of Compounds Derived from the Medicines for Malaria Venture Open Access Box against Intracellular *Leishmania major* Amastigotes. *American Society of Tropical Medicine and Hygiene*. 2016;94(2):340–7.
66. Kolodziej H, Kiderlen AF. Antileishmanial activity and immune modulatory effects of tannins and related compounds on *Leishmania* parasitised RAW 264.7 cells. *Phytochemistry*. 2005;66:2056–71.
67. Maximin Ahuaa K, Ioset J-R, Ransijn A, Mauel J, Mavi S, Hostettmann K. Antileishmanial and antifungal acridone derivatives from the roots of *Thamnosma rhodesica*. *Phytochemistry* 2004;65:963–8.
68. Ponte-Sucre A, Faber JH, Gulder T, Kajahn I, Pedersen SEH, Schultheis M, et al. Activities of Naphthylisoquinoline Alkaloids and Synthetic Analogs against *Leishmania major*. *Antimicrob Agents And Chemotherapy*. 2007;51(1):188-94.
69. Tana N, Kalogaa M, Radtkea OA, Kiderlenc AF, Oksuz S, Ulubelen A, et al. Abietane diterpenoids and triterpenoic acids from *Salvia cilicica* and their antileishmanial activities. *Phytochemistry*. 2002;881–4.
70. Vasquez MA, Iniguez E, Das U, Beverley SM, Herrera LJ, Dimmock JR, et al. Evaluation of α,β -Unsaturated Ketones as Antileishmanial Agents. *Antimicrob Agents Chemother*. 2015;59:3598–601.
71. ChemAxonLtd. JS, Version 6.1.0. 2013.

Referencias Bibliográficas

72. Kode srl. Dragon (software for molecular descriptor calculation). 7.0.4 ed2016.
73. Nascimento DS, Coelho AL. Bagging heterogêneo evolutivo: Caracterização e análise comparativa com ensembles homogêneas de redes neurais RBF. Anais do IX Simpósio Brasileiro de Automação Inteligente. 2009:6.
74. Ariew R. Maquinilla de afeitar de Ockham: Un análisis histórico y filosófico del principio de parsimonia de Ockham. . Universidad de Illinois: Champane-Urbana; 1976.
75. Wold S, Eriksson L. Statistical Validation of QSAR Results. Chemometrics Methods in Molecular design. Weinheim (Germany): Van de Waterbeemd, H., VCH; 1995. p. 309-18.
76. Alfonso-Orta I, et al. Formulario Nacional de Medicamentos. Cuba2014. Available from: <http://fnmedicamentos.sld.cu/index.php?P=Home>.
77. Baldi P BS, Chauvin Y, Andersen CA, Nielsen H, . Assessing the Accuracy of Prediction Algorithms for Classification: an Overview. Bioinformatics. 2000;16:412–24.
78. Alvarez González OR. Identificación de nuevos compuestos antileishmaniásicos a través de estudios *in silico*. Santa Clara: Universidad Central Marta Abreu de Las Villas; 2015.
79. Flores Balmaseda N. Identificación de nuevos compuestos antileishmaniásicos a través de estudios *in silico*. Santa Clara: Universidad Central Marta Abreu de Las Villas; 2015.
80. Salgado Vernier O. Empleo de estudios “*in silico*” en la identificación de nuevos compuestos con potencial actividad contra promastigotes de *Leishmania major*. Santa Clara: Universidad Central Marta Abreu de Las Villas; 2018.

4.1 Manifestaciones clínicas de la leishmaniasis



Figura 16. Leishmaniasis cutánea localizada previa al tratamiento



Figura 17. Evolución de leishmaniasis mucocutánea en un paciente infectado con VIH.



Figura 18. Apariencia física de pacientes con leishmaniasis visceral