



**UNIVERSIDAD MICHOACANA DE
SAN NICOLÁS DE HIDALGO**



**INSTITUTO DE INVESTIGACIONES QUÍMICO
BIOLÓGICAS**

“Diseño por QSAR de compuestos potencialmente activos contra

***Trypanosoma cruzi*”**

**Tesis que para obtener el grado de
Maestra en Ciencias Químicas**

Presenta

Q.F.B. Juana Suárez García

ASESOR DE TESIS:

D.C. Luis Chacón García

MORELIA, MICHOACAN

Diciembre, 2012

El presente trabajo se llevo a cabo en el Laboratorio de Diseño Molecular del Instituto de Investigaciones Químico Biológicas de la Universidad Michoacana de San Nicolás de Hidalgo, bajo la asesoría del D.C. Luis Chacón García.

Una parte de este trabajo fue presentada en la 8^a reunión de la Academia Mexicana de Química Orgánica, realizada en Marzo del 2012.

DEDICATORIAS

Gracias a Dios por todas las bendiciones que me ha dado a lo largo de mi vida.

Dedico con todo mi amor, admiración y agradecimiento, a las personas a quienes debo todo lo que soy mis padres: Mario Suárez Colín y Martha García Victoria.

A mis hermanos Yetza, Sarai, Mari, Erica, Andrés, Marco Antonio y José Manuel, a quienes les agradezco por su cariño, comprensión y apoyo, saben que los quiero mucho, mucho.

A quienes ahora ya no están conmigo pero siempre me animaron a echarle muchas ganas mis abuelitos: Manuel e Irene, Jose y Aurelia.

AGRADECIMIENTOS

A mi asesor el D.C. Luis Chacón García, por compartir sus conocimientos, brindarme su amistad y sobre todo por darme la oportunidad de trabajar en su equipo y con ello convivir con gente muy buena.

A mi comité tutorial D.C. Claudia Contreras Celedón, D.C. Betzabe González Campos, D.C. Erick Cuevas Yáñez y D.C. Mario Gómez Hurtado.

A mis profesores del posgrado.

A mis compañeros y amigos de posgrado Q.F.B. Moran, Q.F.B. Ramón.

De manera muy especial a mis amigas M.C. Fany Elisa y M.C. Rosalía a quienes estimo y admiro mucho por su fortaleza, gracias por su cariño y amistad.

A la familia Carrillo Pérez y Manrique Soto, por su cariño y considerarme como parte de ellos.

A mis amigos de toda la vida, Aza, Meli, Sandra, Meche, Osky, Junior, José Luis, Liz e Itzel Mejia, gracias por su amistad, por su apoyo y sobre todo su cariño es un honor ser su amiga.

A todas las personas maravillosas que pude haber conocido y con quienes compartí estos 2 últimos años, Gaby, Ale, Mario Valle, Chava, Melchor, Anita, César, Nancy, Fátima, Viri, Luis José, Santiago, Migue y Rebe, de corazón gracias por su amistad.

De manera especial quiero agradecer a Lira, Ana Lilia y Anahi, no solo por su amistad y cariño, sino por su paciencia para enseñarme QSAR, mil gracias.

A mis buenos y muy estimados amigos Mario Solorio y Yunuen Torres, gracias por estar ahí.

Al laboratorio de Productos naturales en especial a la D.C. Rosa Elva N. del Rio Torres, Yuni, Edna y Anita.

A las personas que conviven y nos apoyan todos los días Vicky, Blanquita,
Contadora Sarita, Rocío, Cheli y Q.F.B. Manuel.

Agradezco el apoyo brindado por CONACyT.

ÍNDICE

ÍNDICE.....	I
INDICE DE TABLAS	III
INDICE DE FIGURAS.....	V
ABREVIATURAS.....	VI
RESUMEN	VII
ABSTRACT	IX
I.- INTRODUCCIÓN	1
II.- ANTECEDENTES	3
II. 1.- <i>Trypanosoma</i>	5
II.2.- Ciclo Biológico:.....	5
II.3.- Tratamiento:.....	7
II.4.- Efectos tóxicos y adversos:	8
II.5.- Diseño de fármacos.	14
II.6.- Búsqueda de compuestos:	14
II.7.- Diseño de fármacos asistido por computadora.	16
II.8.- QSAR	17
II.8.1.- La elección de la serie.	20
II.8.2.- Obtención del modelo:	20
II.8.3Análisis del modelo.....	21
II.8.4.- Parámetros estadísticos:.....	23
II.8.5.- Técnicas de validación:	26

II.8.6.- Descriptores:.....	28
III.- OBJETIVO GENERAL:.....	31
III.1.-Objetivos particulares:.....	31
IV.- METODOLOGÍA:	32
V.- RESULTADOS:	34
V.1.- Obtención de modelos QSAR de las benzacepinas serie 7a-w con actividad contra <i>trypanosoma cruzi</i>	34
V.2.- Análisis de los modelos sin correlación:.....	38
V.3.- Obtención de modelos QSAR para las benzacepinas serie 7 a-w citotóxica frente a células vero.....	52
V.4.- Modelos seleccionados para el análisis de Toxicidad:.....	58
VI.- DISCUSIÓN DE RESULTADOS.....	67
VI.1 QSAR de compuestos activos contra <i>Trypanosoma cruzi</i>	67
VI.2.-Propuesta de nuevos análogos potencialmente activos contra <i>Trypanosoma cruzi</i>	73
VI.3.- Descriptores del modelo para predicción de actividad:	78
VI. 4 Citotoxicidad:.....	86
VII.- CONCLUSIONES:	102
VIII.- BIBLIOGRAFIA:	103

INDICE DE TABLAS

TABLA 1. COMPARACIÓN DE CARACTERÍSTICAS DIFERENCIALES ENTRE 3 ESPECIES DE INSECTOS TRIATOMINOS EN MÉXICO.....	4
TABLA 2. ACTIVIDAD ANTIPARASITARIA Y CITOTÓXICA DE LOS DERIVADOS CIS-2ARIL-4-HIDROXITETRAHIDRO-1-BENZOACEPINAS, CONTRA <u>T. CRUZI</u>	13
TABLA 3. FAMILIAS DE DESCRIPTORES, SU NOMENCLATURA Y NÚMERO DE DESCRIPTORES QUE LO CONFORMAN OBTENIDOS POR EL PROGRAMA DRAGON.....	29
TABLA 4. RESUMEN DE LOS 50 “MEJORES MODELOS” CON SUS RESPETIVOS DESCRIPTORES Y SU VALIDACIÓN ESTADÍSTICA.....	35
TABLA 5.1 CORRELACIÓN DE LOS DESCRIPTORES DEL MODELO 3.....	38
TABLA 6. VALORES DE LOS DESCRIPTORES DE LAS COMPUESTOS, ASÍ COMO LA ACTIVIDAD CALCULADA PARA EL MODELO 3.....	42
TABLA 7. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 3.....	43
TABLA 8. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 3.....	43
TABLA 9. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, ASÍ COMO LA ACTIVIDAD CALCULADA POR EL MODELO 5.....	44
TABLA 10. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 5.....	45
TABLA 11. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 5.....	45
TABLA 12. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, ASÍ COMO LA ACTIVIDAD CALCULADA PARA EL MODELO 6.....	46
TABLA 13. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 6.....	47
TABLA 14. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 6.....	47
TABLA 15. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, ASÍ COMO LA ACTIVIDAD CALCULADA PARA EL MODELO 9.....	48
TABLA 16. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 9.....	49
TABLA 17. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 9.....	49
TABLA 18. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, ASÍ COMO LA ACTIVIDAD CALCULADA PARA EL MODELO 12.....	50
TABLA 19. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 12.....	51
TABLA 20. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 12.....	51
TABLA 21. SERIE DE COMPUESTOS A UTILIZAR PARA EL QSAR DE CITOTOXICIDAD.....	52
TABLA 22. LOS 50 MODELOS CON SUS DESCRIPTORES Y PRINCIPALES PARÁMETROS ESTADÍSTICOS.....	53
TABLA 23.1 CORRELACIÓN DE LOS DESCRIPTORES DEL MODELO 2.....	55
TABLA 24. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, SE MUESTRA LA CITOTOXICIDAD CALCULADA Y EXPERIMENTAL.....	59
TABLA 25. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 3.....	60
TABLA 26. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 3.....	60
TABLA 27. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, SE DESCRIBE LA CITOTOXICIDAD CALCULADA Y LA EXPERIMENTAL OBTENIDA A PARTIR DEL MODELO 4.....	61
TABLA 28. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 4.....	62
TABLA 29. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 4.....	62
TABLA 30. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, SE MUESTRA LA CITOTOXICIDAD CALCULADA Y EXPERIMENTAL OBTENIDA A PARTIR DEL MODELO 10.....	63
TABLA 31. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 10.....	64
TABLA 32. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 10.....	64
TABLA 33. VALORES DE LOS DESCRIPTORES DE LOS COMPUESTOS, SE MUESTRA LA CITOTOXICIDAD CALCULADA Y EXPERIMENTAL OBTENIDA A PARTIR DEL MODELO 17.....	65
TABLA 34. VALORES DE CORRELACIÓN ENTRE LOS DESCRIPTORES DEL MODELO 17.....	66

TABLA 35. VALORES OBTENIDOS EN LA ESTIMACIÓN LINEAL PARA EL MODELO 17.	66
TABLA 36. RESUMEN DE RESULTADOS DE LOS MEJORES MODELOS ENCONTRADOS EN EL ANÁLISIS QSAR DE BENZACEPINAS ACTIVAS CONTRA <i>TRYPANOSOMA CRUZI</i>	68
TABLA 37. VALORES DE ACTIVIDAD CALCULADA Y EXPERIMENTAL DE LOS COMPUESTOS 7A-T UTILIZANDO EL MODELO 3.	69
TABLA 38. VALORES DE ACTIVIDAD CALCULADA Y EXPERIMENTAL DE LOS COMPUESTOS 7A-T UTILIZANDO EL MODELO 5.	69
TABLA 39. VALORES DE ACTIVIDAD CALCULADA Y EXPERIMENTAL DE LOS COMPUESTOS 7A-T UTILIZANDO EL MODELO 6.	70
TABLA 40. VALORES DE ACTIVIDAD CALCULADA Y EXPERIMENTAL DE LOS COMPUESTOS 7A-T UTILIZANDO EL MODELO 9.	70
TABLA 41. VALORES DE ACTIVIDAD CALCULADA Y EXPERIMENTAL DE LOS COMPUESTOS 7A-T UTILIZANDO EL MODELO 12.	70
TABLA 42. RESIDUOS DE LA DIFERENCIA ENTRE LA SUMATORIA DE LA ACTIVIDAD EXPERIMENTAL Y LA CALCULADA DE LOS MEJORES MODELOS.	72
TABLA 43. RESULTADOS DE ACTIVIDAD CALCULADA POR LOS MODELOS 3 Y 5 PARA LA SERIE DE COMPUESTOS PROPUESTO.	76
TABLA 44. VALORES DE LOS DESCRIPTORES DEL MODELO 5 DE LAS MOLÉCULAS 7E, 7MB Y 7VC.	82
TABLA 45. TABLA DE ELECTRONEGATIVIDADES DE DIFERENTES FUENTES.	84
TABLA 46. LOS MEJORES MODELOS Y SUS PRINCIPALES PARÁMETROS ESTADÍSTICOS.	87
TABLA 47. PRESENTA LOS VALORES DE CITOTOXICIDAD CALCULADOS CON EL MODELO 3,4, 10 Y 17 PARA LA SERIE PROPUESTA.	88
TABLA 48. COMPUESTOS PARA UTILIZADOS PARA LA VALIDACIÓN DE LOS MODELOS.	94
TABLA 49. VALORES DE MASA, VOLUMEN DE VAN DER WAALS, ELECTRONEGATIVIDAD Y POLARIZABILIDAD.	95

INDICE DE FIGURAS

FIGURA 2. ESTRUCTURAS DE LOS FÁRMACOS PARA EL TRATAMIENTO DE LA ENFERMEDAD DE CHAGAS.....	8
FIGURA 3. ESTRUCTURA GENERAL DE LOS DERIVADOS CON ACTIVIDAD CONTRA <i>TRYPANOSOMA CRUZI</i>	9
FIGURA 4. REPRESENTACIÓN DE LA METODOLOGÍA QSAR.	19
FIGURA 5. REPRESENTACIÓN DE UNA GRAFICA DE RELACIÓN LINEAL ENTRE 2 VARIABLES.....	22
FIGURA 6. DIAGRAMA DE A METODOLOGÍA PARA LA OBTENCIÓN DE LOS MODELOS Y VALORES ESTADÍSTICOS.	33
FIGURA 7. COMPUESTOS SOMETIDOS A EXPLORACIÓN PARA OBTENCIÓN DE LOS MODELO QSAR.....	35
FIGURA 8. SE REPRESENTA LA RELACIÓN ENTRE LOS VALORES DE LA ACTIVIDAD EXPERIMENTAL Y LA CALCULADA.	43
FIGURA 9. REPRESENTA LA RELACIÓN ENTRE LOS VALORES DE LA ACTIVIDAD EXPERIMENTAL Y LA CALCULADA.	45
FIGURA 10. SE REPRESENTA LA RELACIÓN ENTRE LOS VALORES DE LA ACTIVIDAD EXPERIMENTAL Y CALCULADA.	47
FIGURA 11. SE REPRESENTA LA RELACIÓN ENTRE LOS VALORES DE LA ACTIVIDAD EXPERIMENTAL Y LA CALCULADA.	49
FIGURA 12. SE REPRESENTA LA RELACIÓN ENTRE LOS VALORES DE LA ACTIVIDAD CALCULADA Y LA EXPERIMENTAL.....	51
FIGURA 13. . SE REPRESENTA LA RELACIÓN ENTRE LOS VALORES DE LA CITOTOXICIDAD EXPERIMENTAL Y CALCULADA.	60
FIGURA 14. REPRESENTACIÓN DE LA RELACIÓN ENTRE LOS VALORES DE LA CITOTOXICIDAD EXPERIMENTAL Y LA CALCULADA.	62
FIGURA 15. . REPRESENTACIÓN DE LA RELACIÓN ENTRE LOS VALORES DE LA CITOTOXICIDAD EXPERIMENTAL Y LA CALCULADA.	64
FIGURA 16. REPRESENTACIÓN DE LA RELACIÓN ENTRE LOS VALORES DE LA CITOTOXICIDAD EXPERIMENTAL Y LA CALCULADA.	66
FIGURA 17. ACTIVIDAD EXPERIMENTAL CONTRA OBSERVADA DE LOS 5 MODELOS.....	72
FIGURA 18. SERIE PROPUESTA DE DERIVADOS DE CIS-2-ARIL-4-HIDROXITETRAHIDRO-1-BENZACEPINAS. ..	76
FIGURA 19. ESTRUCTURAS DE LAS MOLÉCULAS MÁS ACTIVAS.....	81
FIGURA 20. REPRESENTACIÓN DE BARRAS DE LOS VALORES DEL DESCRIPTOR MOR21V, PARA LAS MOLÉCULAS 7MB: 1, 7VC: 2 Y 7E: 3, RESPECTIVAMENTE.	82
FIGURA 21. VALORES DEL DESCRIPTOR MOR10E PARA LOS COMPUESTO 7VC: 1, 7E:2 Y 7MB: 3.....	83
FIGURA 22. REPRESENTACIÓN DE LOS VALORES DEL DESCRIPTOR MOR21P PARA LOS COMPUESTOS 7E: 1, 7VC: 2VC Y 7MB: 3.....	83
FIGURA 23. VALORES DEL DESCRIPTOR MOR10U PARA LOS COMPUESTOS, 7VC: 1, 7E: 2 Y 7MB: 3.	84
FIGURA 24. REPRESENTACIÓN DEL NÚCLEO COMÚN.....	85
FIGURA 25. VALORES DEL DESCRIPTOR GATS8V PARA LOS COMPUESTOS 7C: 1, 7PB: 2, 7MA: 3 Y 7J: 4. .	95
FIGURA 26. VALORES DEL DESCRIPTOR GATS2P PARA LOS COMPUESTOS 7MA: 1, 7C: 2, 7J: 3 Y 7PB: 4. .	96
FIGURA 27. VALORES DEL DESCRIPTOR RDF070M PARA LOS COMPUESTOS 7J: 1, 7PB: 2, 7MA: 3 Y 7C: 4.97	
FIGURA 28. COMPARACIÓN DE LA ACTIVIDAD CALCULADA CON EL MODELO 3 Y LA CITOTOXICIDAD CALCULADA CON EL MODELO 3.....	98
FIGURA 29. COMPARACIÓN DE LA ACTIVIDAD CALCULADA CON EL MODELO 5 Y LA CITOTOXICIDAD CALCULADA CON EL MODELO 3.....	99
FIGURA 30. COMPARACIÓN DE LA ACTIVIDAD CALCULADA CON EL MODELO 3 Y LA CITOTOXICIDAD CALCULADA CON EL MODELO 4.....	100
FIGURA 31. COMPARACIÓN DE LA ACTIVIDAD CALCULADA CON EL MODELO 3 Y LA CITOTOXICIDAD CALCULADA CON EL MODELO 4.....	101

Abreviaturas

QSAR	Estudio de relación cuantitativa estructura-actividad.
AM1	Método semiempíricos Austin Model 1.
MNDO	Método semiempíricos Modified Neglected Diatomic Overlap
AGs	Algoritmo genético
Q^2	Capacidad predictiva del modelo.
R^2	Indicador de la precisión del modelo (Coeficiente de determinación).
Q^2_{boot}	Coeficiente de determinación de validez cruzada.
F	Parámetro de Fisher.
s	Desviación estándar.
Kx	Correlación entre descriptores.
n	Número de datos
Å	Angstrom.
ps	Picosegundos
IC	Concentración inhibitoria.
CC	Concentración citotóxica
K	Kelvin
μM	Micromolar

RESUMEN

“Diseño por QSAR de compuestos potencialmente activos contra
Trypanosoma cruzi.”

Existen enfermedades conocidas y diagnosticadas para las cuales no existe un tratamiento exitoso y que además caben en el grupo de enfermedades emergentes, según la OMS, de las cuales urge encontrar mecanismos de prevención y tratamientos eficientes. Uno de estos casos es la enfermedad de Chagas, causada por protozooario *Trypanosoma cruzi*, se ha distribuido a lo largo del continente americano. Actualmente, hay dos fármacos utilizados para el tratamiento esta enfermedad: el Nifurtimox y el Benznidazol, efectivos en la fase aguda de la enfermedad pero ineficientes en la fase crónica. En vista, tanto de la limitante de medicamentos disponibles como del problema de salud pública, es necesaria la búsqueda de compuestos activos que actúen tanto la fase aguda como en la crónica de la enfermedad y a su vez causen menos efectos secundarios y colaterales indeseables.

Una herramienta que permite al químico medicinal optimizar su trabajo es la química computacional, que tiene por objetivo simular procesos químicos que arrojan datos que permiten predecir comportamientos y características de un determinado compuesto al que se le desconoce su actividad. Existen varios métodos de diseño de fármacos, entre ellos el QSAR, (Quantitative structure-activity relationships). Esta metodología, se basa en un proceso por el cual la estructura química esta matemáticamente correlacionada con su actividad biológica.

En el presente trabajo se llevó a cabo el estudio de QSAR de una serie de tetrahydrobenzoacepinas con actividad contra *Trypanosoma cruzi* *in vitro* contra epimastigotes y amastigotes, actividad aunada a baja toxicidad en líneas celulares.

Empleando la metodología QSAR, se obtuvieron modelos matemáticos para explicar la actividad y la citotoxicidad de la misma serie de compuestos. Los modelos que predicen la actividad contienen descriptores Mor21v, Mor10e,

Mor24e, Mor10u y Mor21p, que convierte la estructura tridimensional a un número de variables fijas por medio de una transformación usando estudios de difracción de electrón. Para los modelos de citotoxicidad los descriptores que la explican fueron RDF015m, RDF060u, RDF070m que proporcionan información sobre las distancias interatómicas en la molécula completa y los descriptores GATS8v y GATS2p que se relacionan con la autocorrelación espacial.

Con estos descriptores se propusieron compuestos activos con valores de toxicidad bajos, interpretando las propiedades que miden cada uno de ellos y se determinó que la actividad y baja toxicidad está en función a estos descriptores.

ABSTRACT

“Diseño por QSAR de compuestos potencialmente activos contra
Trypanosoma cruzi.”

There are known and diagnosed diseases for which there is no successful treatment and also fit into the group of emerging diseases, according to OMS, of which mechanisms urgent to find effective prevention and treatment. One such case is Chagas disease, caused by protozoan *Trypanosoma cruzi*, is distributed throughout the Americas. Currently, two drugs used to treat this disease: nifurtimox and benznidazole, effective in the acute phase of the disease but inefficient in the chronic phase. In view of both the limiting the drugs available as a public health problem, it is necessary to search for active compounds that act both the acute and the chronic disease and in turn cause less side effects and undesirable side.

A tool that allows the medicinal chemist optimize your work is computational chemistry, which aims to simulate chemical processes that provide data that predict behaviors and characteristics of a given compound to which activity is unknown. There are several methods of drug design, including QSAR (Quantitative Structure-Activity Relationships). This methodology is based on a process in which the chemical structure is mathematically correlated with biological activity.

The present work was carried out QSAR study of a series of tetrahydrobenzoacepinas with activity against *Trypanosoma cruzi* in vitro against epimastigotes and amastigotes, activity combined with low toxicity in cell lines.

Using the QSAR methodology were obtained mathematical models to explain the activity and cytotoxicity of the same series of compounds. The models predict that contain descriptors Mor21v activity, Mor10e, Mor24e, Mor21p Mor10u and which converts the three-dimensional structure to a variable number of fixed by a transformation using electron diffraction studies. For cytotoxicity models that

explain descriptors were RDF015m, RDF060u, RDF070m that provide information about the interatomic distances in the complete molecule and GATS2p GATS8v descriptors that relate to the spatial autocorrelation.

These descriptors are proposed active compounds with low toxicity values, interpreting the properties measured and each found that the activity and low toxicity is a function of the descriptors.

I.- INTRODUCCIÓN

La manifestación de patógenos, desconocidos o nuevos para el humano puede dar lugar a problemas serios de salud y económicos a nivel mundial. Un ejemplo reciente, fue la aparición del virus de la influenza tipo A H1N1 a finales de la primera década del siglo XXI que hizo reflexionar al mundo entero sobre la vulnerabilidad de los humanos ante un patógeno del cual se desconoce tratamiento o vacuna. Asimismo, existen enfermedades bien conocidas y diagnosticadas para las cuales no existe un tratamiento adecuado y que pertenecen al grupo de enfermedades “emergentes” de acuerdo a la Organización Mundial de la Salud (OMS) de las cuales es necesario encontrar mecanismos de prevención y tratamientos eficientes. Uno de estos casos es la enfermedad de Chagas¹, estudiada formalmente por primera vez en 1909, y nombrada así en honor al médico e infectólogo brasileño Carlos Chagas. Hasta ese entonces no fue considerada un problema de salud pública sino hasta la década de los 60's².

Durante décadas, la enfermedad de Chagas ha afectado a la población de las zonas rurales de América Latina. En los últimos años la migración masiva ha hecho que la enfermedad se convierta en alerta mundial³.

Una vez que se contrae la enfermedad, los síntomas son silenciosos y a menudo tardan años en aparecer. La infección puede permanecer inactiva durante décadas. En la mayoría de las personas infectadas aparecen complicaciones que pueden ocasionar discapacidades y la muerte. Actualmente hay 2 fármacos aprobados para el tratamiento de la enfermedad: el Nifurtimox y el Benznidazol⁴. Ambos compuestos son efectivos en la fase aguda de la enfermedad pero resultan ineficientes en la fase crónica, que es la más notable, además de que los tratamientos son prolongados, y tienen limitantes por los daños que causa su uso constante. En vista, tanto de la limitante de medicamentos disponibles como del

problema de salud pública, es urgente la búsqueda de compuestos activos que actúen no solo en la fase aguda sino que además presente actividad en la fase crónica de la enfermedad y que a su vez causen menores efectos secundarios y colaterales indeseables.

Dentro de la búsqueda de compuestos biológicamente activos, la evolución del conocimiento proporciona cada día más herramientas que permiten al Químico Medicinal optimizar su trabajo. Una de estas herramientas es la química computacional, que tiene por objetivo simular procesos químicos a partir de los cuales es posible obtener datos que permiten predecir comportamientos y características de un determinado compuesto del cual se desconoce su actividad biológica.

Existen varios métodos de diseño de fármacos, entre ellos el QSAR (Quantitative structure-activity relationship). Esta metodología, se basa en un proceso por el cual se correlaciona matemáticamente la estructura química con su actividad biológica. Un estudio QSAR está fundamentado en la premisa de que la actividad está en función de las propiedades estructurales del compuesto activo. Para obtener modelos adecuados, se requiere de una serie de compuestos biológicamente activos, congéneres entre ellos, que actúen bajo el mismo mecanismo de acción y de los cuales se cuente con sus datos de actividad. La obtención de buenos modelos de QSAR depende de varios factores como la calidad de los datos biológicos, la selección de los descriptores moleculares y los métodos estadísticos. Un buen modelo QSAR puede ser capaz de hacer predicciones de la actividad biológica de compuestos aun no sintetizados, llevando consigo la ventaja de reducir el número de compuestos a sintetizar, reduciendo costos y aumentando la posibilidad de obtener compuestos exitosos⁵.

En el presente proyecto se propone el estudio QSAR de una serie de tetrahydrobenzazepinas con actividad contra *Trypanosoma cruzi* *in vitro* y a partir del mismo, el diseño de derivados potencialmente activos.

II.- ANTECEDENTES

La enfermedad de Chagas, causada por el parásito *Trypanosoma cruzi*, se ha distribuido a lo largo del continente americano, principalmente en zonas de alta pobreza y marginación rural de centro y Sudamérica y se estima que hay aproximadamente 35 millones de personas expuestas al riesgo de la infección⁶. Aunque la enfermedad estaba establecida casi exclusivamente en zonas rurales, actualmente debido a la migración a las grandes ciudades, el establecimiento de la enfermedad de Chagas está cambiando su perfil epidemiológico. De acuerdo a reportes sanitarios, hay evidencias de que la enfermedad se ha presentado desde EUA hasta Argentina, incluyendo México⁷.

En México se consideran como áreas endémicas probables de esta enfermedad aquellas ubicadas de cero a 1500 m sobre el nivel del mar, dado que a esta región altitudinal se han encontrado, en zonas rurales, altos índices de insectos triatominos infectados por *T. cruzi*. La cota altitudinal de riesgo podría extenderse hasta 1800 m sobre el nivel del mar. Los Estados de la frontera norte y de la península de Baja California han registrado solo casos esporádicos. Por el contrario, los Estados del centro y sur: Guerrero, Oaxaca, Morelos, Chiapas, Michoacán y Jalisco producen el mayor número de casos confirmados, en esta zona el principal transmisor es *Triatoma pallidipennis*, el insecto triatomo de hábitos domésticos que se encuentra en el País; en Oaxaca y Chiapas además se encuentra el *Rhodnius prolixus*, eficaz transmisor⁶. Investigaciones han resaltado la importancia de 3 especies de insectos triatominos en la transmisión de *T. cruzi* en México, dichas especies se resumen en la siguiente Tabla 1⁸.

Tabla 1. Comparación de características diferenciales entre 3 especies de insectos triatominos en México.

Especie	<i>Triatoma barberi</i>	<i>Triatoma pallidipennis</i>	<i>Triatoma dimidiata</i>
Estado	Querétaro	Morelos	Veracruz
Ecotopo domiciliar	Paredes	Pisos	Piso, bajo camas
Ecotopo silvestre	NE ¹	Tecorrales	NE
Ciclo vida (días)²	523	171	510
Atracción por la luz	Si	No	Si
Defecación	Durante la alimentación	20-30 min después de alimentarse	10-20 min después de alimentarse
Densidad poblacional³	50.5	29.5	11
Índice de infección	56.6	29	9
Índice de metaciclologenia	76.6	15	26

1 No se recolectaron, 2 Observaciones de laboratorio, 3 Promedio anual

La enfermedad de Chagas representa un problema de salud pública en 17 Países Latinoamericanos, con un total estimado por la Organización Mundial de la Salud de 100 millones de personas expuestas a la enfermedad y 20 millones de personas infectadas⁹. Cabe resaltar, que esta enfermedad está relacionada, con la pobreza y malas condiciones de vivienda, además de los procesos de colonización, migraciones y transfusiones sanguíneas. La transmisión del parásito puede llevarse a cabo de diversas maneras como por ejemplo la transmisión vectorial, por la leche materna, vía trasplacentaria. Aunque el 80% de las infecciones en humanos se presenta, mediante la introducción de las heces fecales del insecto a la picadura; menos comunes en su transmisión, son por transfusiones sanguíneas, por consumo de cárnicos contaminados o bien por contaminación accidental en el laboratorio⁵.

“Diseño por QSAR de compuestos potencialmente activos contra *Trypanosoma cruzi*”

En la región de Sudamérica, específicamente en el territorio Colombiano se han reportado 19 especies de insectos triatominos capaces de transmitir al parásito *Trypanosoma cruzi*, de las cuales destacan *Rhodnius prolixus*, *Triatoma dimidiata*, *Triatoma maculata*, *Rhodnius robustus*, *Rhodnius brethesi*, *Triatoma venosa* y *Rhodnius pallescens*. El hecho de contar con muchos vectores, sumando la condición de país subdesarrollado, expone a la población a un mayor riesgo de contraer la enfermedad⁹.

II. 1.- Trypanosoma

La tripanosomiasis se denomina como enfermedad africana del sueño; enfermedad de Chagas en el sur de E.U.A., México, América central y del sur, y la tripanosomiasis asintomática en centro y Sudamérica.

En África, el agente etiológico de la tripanosomiasis es el *Trypanosoma brucei*, mientras que el *Trypanosoma cruzi*, produce la enfermedad de Chagas (en América). El género *trypanosoma* es un hemoflagelado que se presenta en la sangre de mamíferos como tripomastigotes alargados maduros¹⁰.

II.2.- Ciclo Biológico:

El ciclo biológico del agente causal, *Trypanosoma cruzi* se presenta en la naturaleza con 4 estadios morfológicos principales que son: tripomastigotes, promastigote, epimastigote y amastigote. El ciclo biológico del parásito se inicia cuando un insecto triatomino punza a un mamífero y le extrae sangre que contienen tripomastigotes circulantes, los cuales pasan al intestino del insecto triatomino, se transforman en epimastigotes, se multiplican por fisión binaria y en pocos días ya hay tripomastigotes metacíclicos¹¹.

Cuando el parásito atraviesa las conjuntivas o pequeñas heridas en la piel, pasa al torrente sanguíneo y se introduce a las células de diferentes tejidos, donde

transformaran en amastigote intracelular, a partir de la fisión binaria, hasta formar las bolsas de amastigotes, que posteriormente se romperán y los amastigotes se transformaran en tripomastigotes que circula a través de la sangre hasta que se introduce a otra célula. (Figura 1)¹²

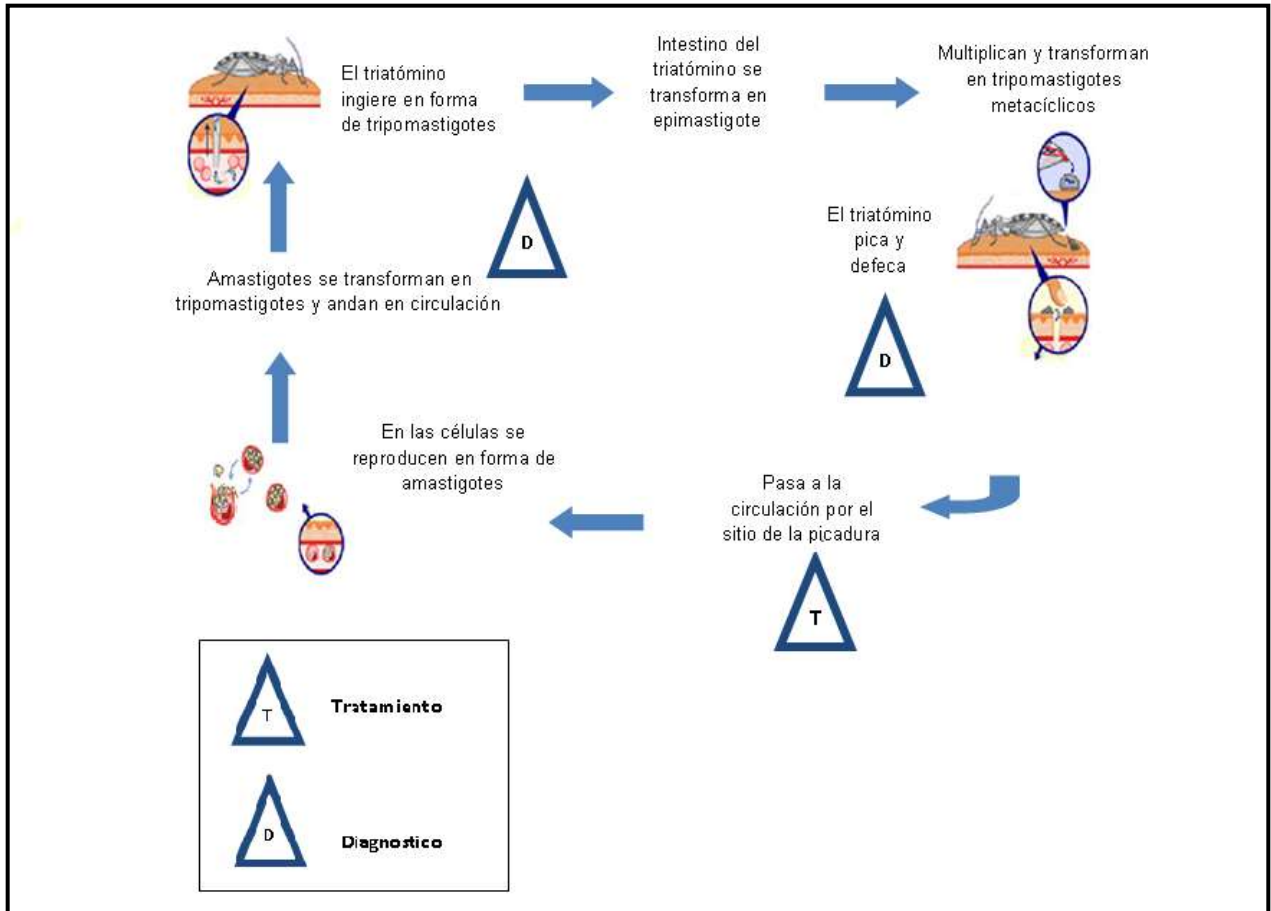
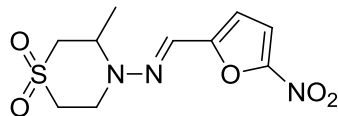


Figura 1. Estadios del *Trypanosoma cruzi*.

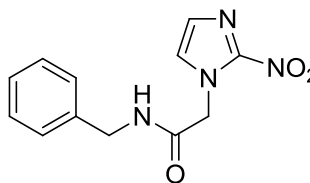
II.3.- Tratamiento:

Investigaciones revelan que derivados de nitrofuranos y nitroimidazoles son eficaces en el combate de infecciones por *Trypanosoma cruzi*, por lo que se han emprendido estudios de innumerables análogos, uno de ellos que resulto exitoso, desde el punto de vista de la química medicinal, fue el Nifurtimox (Lampit ®) desarrollado en 1960 por Bayer, que requiere de 1 a 2 meses de tratamiento y el Benznidazol desarrollado en 1964 por Roche (Rochagan ®). Se administra por vía oral y se obtiene mejores resultados en la etapa aguda que en la crónica de la enfermedad.¹³

Nifurtimox y el Benznidazol (figura 1), son tripanocidas contra las formas de tripomastigote y amastigote de *T. cruzi*¹⁴. A bajas concentraciones (1µM) dañan a los amastigotes intracelulares *in vitro* e inhiben su desarrollo. La exposición continua a la concentración mencionada del compuesto alarga considerablemente el ciclo intracelular. Los tripomastigotes son menos sensibles; concentraciones de 10 µM de Nifurtimox inhiben pero no eliminan la penetración en las células de vertebrados¹³. La actividad tripanocida del Nifurtimox al parecer depende de su capacidad de mostrar reducción parcial del anión radical nitro, seguida de autoxidación hasta generar el nitrofurano primario y formar un anión radical superóxido y otras especies de reactivas oxígeno como el peróxido de hidrógeno y el radical hidroxilo. La reacción de los radicales libres con macromoléculas celulares culmina en daño celular que incluye peroxidación de lípidos y daño de membrana, inactivación de enzimas, lesión de DNA y mutagénesis¹⁴.



Nifurtimox



Benznidazol

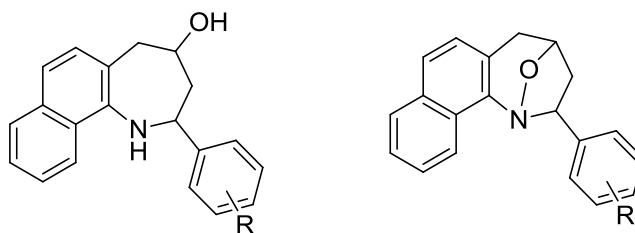
Figura 1. Estructuras de los fármacos para el tratamiento de la enfermedad de chagas.

II.4.- Efectos tóxicos y adversos:

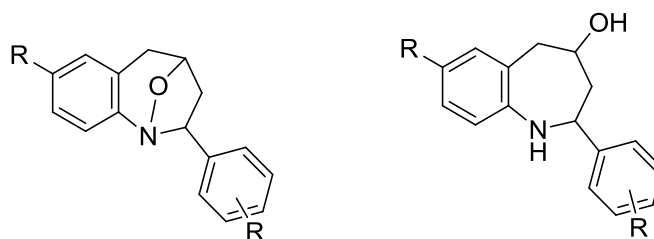
Desafortunadamente los tratamientos comunes contra este parásito tienen numerosas limitaciones, tales como poca eficacia, efectos tóxicos secundarios, largos periodos de tratamiento de alto costo y pérdida de eficacia debido a la resistencia del parásito. Los niños toleran el Nifurtimox y el Benznidazol mejor que los adultos. A pesar de ello, es frecuente que surjan efectos adversos por los fármacos; varían desde reacciones de hipersensibilidad como dermatitis, fiebre, ictericia, infiltrados pulmonares y anafilaxia, hasta complicaciones dependientes de la dosis y la edad, atribuibles más bien a las vías gastrointestinales y los sistemas nerviosos periférico y central. Es frecuente observar náuseas y vómitos, mialgias y debilidad. Después de un tratamiento largo son especialmente comunes la neuropatía periférica y los síntomas de vías gastrointestinales; estas últimas complicaciones a veces causan pérdida ponderal e impiden la repetición o la continuación del tratamiento. Debido a la alta toxicidad del Nifurtimox se prefiere el Benznidazol¹⁴. Por esta razón varios investigadores han enfocado sus esfuerzos en la síntesis de nuevos compuestos que pretenden ser más eficaces y menos tóxicos.

Como parte de la investigación desarrollada en nuestro grupo de trabajo, se cuenta con la colaboración del Laboratorio de Síntesis Orgánica de la Universidad Industrial de Santander, Colombia, dirigido por el Dr. Alirio Palma Rodríguez,

quienes han descrito recientemente la síntesis de una serie de compuestos novedosos derivados de tetrahidronafto[1,2-b]azepinas y tetrahidrobenzoazepinas (Figura 2) con una notable actividad contra las formas de epimastigotes y amastigotes del *Trypanosoma cruzi*. Su actividad, aunada a la baja toxicidad en líneas celulares hace de estos compuestos fármacos potenciales para el tratamiento de la enfermedad de Chagas. Aunque no hay estudios formales sobre su mecanismo de acción, se ha considerado que este es a través de la inhibición de la hidrofolato reductasa del parásito¹⁵.



Derivados de tetrahidronafto[1,2] azepinas

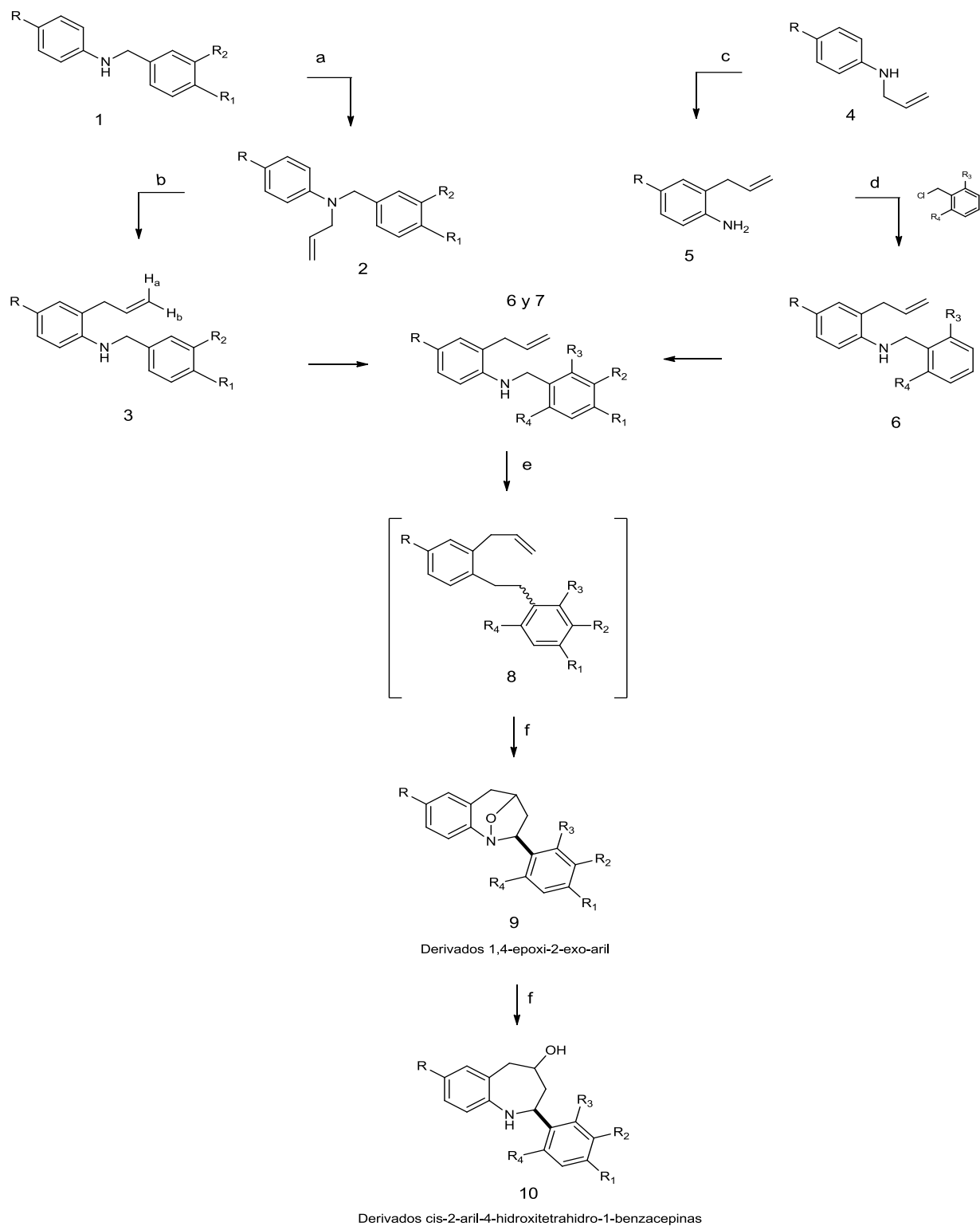


Derivados de tetrahydro-1-benzazepinas

Figura 2. Estructura general de los derivados con actividad contra *Trypanosoma cruzi*.

Los compuestos elegidos para desarrollar el presente trabajo fueron los derivados de las tetrahydrobenzodiazepinas; obtenidos bajo una estrategia sintética, diseñada por Palma y colaboradores, donde se involucran dos reacciones clave; una transposición amina-Claisen y una ciclo adición intramolecular 1,3 dipolar.

Esquema 1



Esquema 1. Síntesis de derivados *cis*-2-aril-4-hidroxitetrahidro-1-benzacepinas. (Modificado de Gómez-Ayala et al. 2010)

Se obtuvieron 23 derivados del 1,4-epoxi-cicloaducto, 1,4-epoxi-2-exo-arilo (serie 6 **a-w**) y 23 derivados de los amino alcoholes cis-2-arilo-4-hidroxitetrahydro-1-benzacepinas (**serie 7 a-w**). Los epoxiaductos 6 a-w no resultaron ser activos en general y solo nueve de ellos presentaron actividad moderada (IC_{50} de 7.68-35.39 μ M) frente a epimastigotes. Sin embargo, en lo referente a los aminoalcoholes 7 a-w, 13 de ellos presentan concentraciones inhibitorias de buenas a moderadas (IC_{50} de 6.10-30.06 μ M) en epimastigotes. Los valores de IC_{50} en epimastigote y CC_{50} en células Vero de cada uno de los aminoalcoholes, así como la actividad de referencia con Nifurtimox, se resume en la Tabla 2. Esta serie de compuestos cumple con los requerimientos para llevar a cabo un estudio QSAR:

- 1.- Todos los análogos deben pertenecer a series congéneres
- 2.- Todos los análogos deben ejercer el mismo mecanismo de acción
- 3.- Todos los análogos se deben unir de manera similar
- 4.- Se deben poder predecir los efectos de reemplazamientos isostéricos
- 5.- La afinidad de unión está correlacionada con las energías de interacción
- 6.- La actividad biológica está relacionada con la afinidad de la unión

Tabla 2. Actividad antiparasitaria y citotóxica de los derivados cis-2-aryl-4-hidroxitetrahydro-1-benzoacepinas, contra *T. cruzi*.

7

Compuesto	R	R ₁	R ₂	R ₃	R ₄	IC ₅₀ epimastigote	CC ₅₀ / Cel.Vero
7 a	H	H	H	H	H	31.82	>300
7b	OCF ₃	H	H	H	H	26.59	123.55
7c	Br	H	H	H	H	15.71	42.14
7d	Cl	H	H	H	H	49.58	190.89
7e	F	H	H	H	H	12.37	>300
7f	H	Cl	H	H	H	65.44	>300
7g	H	F	H	H	H	26.37	>300
7h	H	H	OMe	H	H	30.06	>300
7i	H	H	Me	H	H	15.64	>300
7j	H	H	Cl	H	H	19.62	295.36
7k	Me	Me	H	H	H	6.1	>300
7l	Cl	Cl	H	H	H	16.45	76.6
7m	Cl	Br	H	H	H	27.21	57.69
7n	F	Cl	H	H	H	45.93	193.8
7o	H	H	H	Cl	H	49.73	>300
7p	H	H	H	F	H	62.97	>300
7q	Me	H	H	Cl	H	22.43	82.37
7r	Cl	H	H	Cl	H	15.34	106.01
7s	Cl	H	H	F	H	17.31	182.39
7t	H	H	H	Cl	Cl	48.88	>300
7u	H	H	H	Cl	F	>100	>300
7v	Cl	H	H	Cl	Cl	14.84	49.91
7w	Cl	H	H	Cl	F	14.46	92.02
Nifurtimox						2.33	61.42

II.5.- Diseño de fármacos.

El desarrollo de un nuevo fármaco o medicamento, comprende una serie de fases que conllevan largos periodos de tiempo, desde su obtención hasta su registro, además de un costo. El descubrimiento debe ser diferenciado del desarrollo, el descubrimiento abarca; la síntesis, aislamiento de la fuente natural o la obtención biotecnológica hasta la fase preclínica incluida la toxicología; mientras que la fase del desarrollo implica los estudios clínicos hasta el registro farmacéutico, todo el proceso desde su obtención hasta su registro engloba un costo promedio 231 millones de dólares por cada nuevo medicamento así como años de investigación. Lo más preocupante es que solo 10,000 moléculas ensayadas pasan a la fase de desarrollo y una de cada 100,000 logra superar los ensayos clínicos, para solo 3 de cada 10 moléculas logran el registro. Por tal motivo el diseño racional de fármacos constituye una herramienta indispensable¹⁶.

Para el diseño, es importante conocer la base biológica de la enfermedad, cuando se conoce este proceso en su base molecular y se pueden definir las moléculas implicadas en el mismo, es posible diseñar medicamentos que interactúen con la molécula responsable.

Para conocer a profundidad el proceso fisiológico, es necesario conocer la estructura tridimensional de las moléculas objetivos, con esto se comprenden y se miden sus propiedades fisicoquímicas por medio de descriptores moleculares.

II.6.- Búsqueda de compuestos:

La búsqueda de nuevos fármacos sigue algunas estrategias generales:

1.- Mejora de fármacos ya existentes: Esta estrategia consiste en hacer modificaciones a la estructura química del medicamento original con el objetivo de mejorar su actividad, o disminuir sus efectos adversos, indicando que en ocasiones estas transformaciones llevan a descubrir una propiedad nueva. A estos

medicamentos se les conoce como medicamentos “Me-too”. Un ejemplo clásico es la modificación de la Codeína que llevo al Dextrometorfano, el ingrediente activo de la mayor parte de los antitusivos¹⁷.

2.-Examen sistemático de grupos de compuestos: Radica en el rastreo (screening) a nuevas moléculas que pueden ser de origen natural o sintético y la investigación de su actividad en animales o en los diversos ensayos biológicos, sin partir de una hipótesis o cualquier referencia sobre su función farmacológica, dichos análisis se apoyan más de ensayos *in vitro* que *in vivo*.

3.- Explotación de información biológica: Se parte de los conocimientos de la etnofarmacología para la búsqueda de prototipos. La observación de efectos colaterales, en ocasiones arroja un resultado de interés farmacológico que puede proporcionar información para el diseño de un fármaco.

4.- Aproximaciones racionales: En la actualidad se cuenta con información de la etiología de las enfermedades y avances tecnológicos como la computadora que permiten el diseño de un fármaco asistido por esta; se usan 2 tipos de aproximaciones, el diseño directo y el diseño indirecto de moléculas activas¹⁸.

El diseño directo se lleva a cabo cuando se conoce la estructura de la molécula diana, ya que en la computadora se puede establecer la conformación y configuración adecuada, así como posibles sustituyentes, es decir, en la computadora podemos establecer condiciones favorables de interacción entre el posible fármaco y la molécula diana.

El diseño indirecto se aplica cuando no se conoce la estructura de la macromolécula hospedadora, lo cual es bastante común, y consiste en la comparación de una serie de moléculas que presenten actividad selectiva en un receptor específico, con la finalidad de identificar las características estructurales que puedan tener y con estos datos construir una molécula más eficaz¹⁸. Es en

este punto donde caben los estudios QSAR que han demostrado gran utilidad en el descubrimiento de fármacos.

II.7.- Diseño de fármacos asistido por computadora.

Para el caso de diseño de fármacos asistido por computadora se engloban métodos teóricos y técnicas computacionales para modelar o imitar el comportamiento de las moléculas.

El modelado molecular implica:

Mecánica molecular: Donde las ecuaciones son usadas para seguir las leyes de la física clásica y aplicarlas al núcleo molecular considerando a los electrones fuera, es decir la molécula es tratada como una serie de esferas conectadas por los enlaces. Las energías calculadas por la mecánica molecular no representan una cantidad absoluta por que es utilizada cuando se comparan diferentes conformaciones de la misma molécula. La mecánica molecular es más rápida y menos extensa que la mecánica cuántica.

La mecánica molecular es usada en las siguientes operaciones o cálculos:

- Optimización de energía
- Identificación de conformaciones estables
- Calcular la energía para conformaciones específicas
- Generación de diferentes conformaciones
- Estudio del movimiento molecular

Mecánica cuántica: Usa la física cuántica para predecir las propiedades de una molécula por las interacciones consideradas entre los electrones y núcleos de las moléculas. Para hacer los cálculos posibles, se deben hacer varias aproximaciones. Los métodos de mecánica cuántica pueden ser subdivididos en métodos *ab initio* o *semiempíricos*. El método semiempírico, al ser un método parametrizado, es más rápido aunque menos preciso y puede optimizar

estructuras de tamaño relativamente grande. Los programas semiempíricos son el MINDO/3, MNDO, MNDO-4, AM1 y PM3, estos métodos son rápidos después de usar aproximaciones y hacer uso de los parámetros almacenados.

La mecánica cuántica es apropiada para calcular o siguiente:

- La energía de la órbita molecular y coeficientes
- Calor de formación para conformaciones específicas
- Cargas parciales atómicas calculadas de los coeficientes de la órbita molecular
- Potenciales electrostáticos
- Momento dipolar
- Estado de transición geométricas y energías
- Energía de disociación de enlaces

La elección del método va a depender de los cálculos que se necesiten y del tamaño de la molécula¹⁹.

II.8.- QSAR

Los métodos que relacionan la estructura química con la actividad biológica, asistidos por computadora, pueden dividirse en 2 categorías: el método QSAR (Quantitative Structure-Activity Relationship), en español se traduce como Estudios cuantitativos de relación estructura-actividad y los Métodos de Modelación Molecular.

El QSAR es una de las herramientas más importantes para el diseño y descubrimiento de fármacos, así como para explorar la interacción ligando-receptor/enzima, especialmente cuando no se conoce alguno de estos.²⁰

Para desarrollar un estudio QSAR se requiere que la actividad biológica de la molécula o fármaco, este en función de las características estructurales de la molécula; el método más comúnmente usado es el de Hansch, que es un método

extratermodinámico basado originalmente en efectos electrónicos, hidrofóbicos y estéricos. El modelo matemático de Hansch se explica con la siguiente ecuación:

$$\ln A = f_h(X_h) + f_e(X_e) + f_s(X_s) + Cte$$

Donde:

A: es la actividad

$f_h(X_h) + f_e(X_e) + f_s(X_s)$: son funciones o parámetros fisicoquímicos de la molécula.

Para que los métodos sean válidos deben considerarse ciertas suposiciones englobadas en el paradigma de Hansch¹⁶:

1.- La actividad biológica es función de la estructura del fármaco.

2.- La estructura del fármaco implica ciertas propiedades globales como la hidrofobicidad, carga neta, solubilidad etc., y ciertas propiedades locales como la distribución de la hidrofobicidad, carga y volumen en determinadas posiciones de la molécula.

3.- Estas propiedades globales y locales pueden ser cuantificadas mediante ciertos parámetros.

4.- Siempre existe una función que relaciona los cambios de actividad biológica con los cambios en las propiedades globales y locales, aunque puede que no sea sencilla ni evidente¹⁶.

La metodología que sigue el QSAR es la siguiente:

Se requiere de un prototipo cabeza de serie a partir de la cual se diseña una serie de exploración, que no son más que análogos del prototipo mencionado, estos son evaluados, es decir, se establecerá su actividad experimental tomando en cuenta que los miembros de la serie tienen un núcleo común y diferentes sustituyentes, estos fragmentos van a ser identificados por descriptores que serán utilizados como variables independientes que expresaran una actividad. La actividad se maneja en forma logarítmica y esta representa a la variable dependiente; de esta relación se desprenderá un modelo matemático, que servirá para la predicción o la obtención de productos óptimos¹⁶. Figura 3

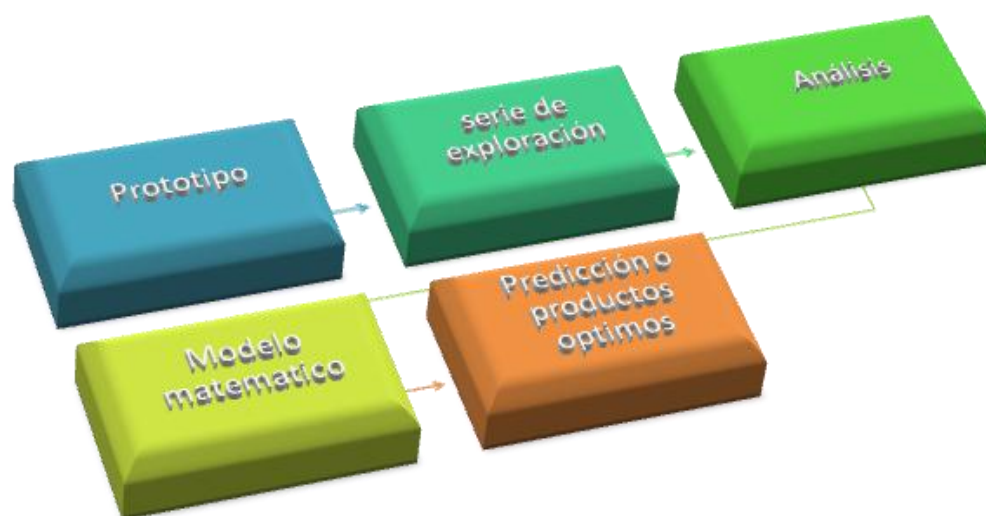


Figura 3. Representación de la metodología QSAR.

II.8.1.- La elección de la serie.

Para la elección de la serie se toman en cuenta los siguientes criterios:

A) Las moléculas deben presentar diferentes características químicas entre sí, ya que no tiene sentido estudiar cómo influye una sola característica en la actividad.

B) Se debe manejar un número de población o moléculas de 10 o más, ya que estadísticamente la población es un conjunto de elementos que tienen características en común y son dignas de estudio.

II.8.2.- Obtención del modelo:

Un modelo se puede obtener mediante diferentes métodos, uno de ellos son los algoritmos genéticos o AGs. Un algoritmo es una serie de pasos matemáticos para resolver un problema específico. Los algoritmos genéticos forman parte de la inteligencia artificial y se basan en las reglas de la evolución biológica: selección natural basada en mutaciones, elitismo, sobrevivencia del “más fuerte”. Los algoritmos genéticos fueron creados en la década de los 60’s por John Holland, como modelo para el estudio del fenómeno de adaptación natural y para el desarrollo de mecanismos que permitieran incorporar este fenómeno a los sistemas de computo²¹. El primer paso es crear una población de modelos productos de regresión lineal. Estos modelos de regresión se acoplan entre sí, mutan, se cruzan, se reproducen y entonces evolucionan a través de generaciones sucesivas hacia una solución óptima²².

Los AGs proporcionan una alternativa de métodos de optimización tradicional, usando la técnica de la búsqueda del más fuerte y ubicar una solución óptima local. La estrategia para la selección se basa en la evolución de una población de modelos, es decir, un conjunto de modelos clasificados según alguna función objetivo. En su tratamiento matemático, cada individuo

se denomina cromosoma y es un vector binario, en el que cada posición (un gen) corresponde a una variable. Cada cromosoma representa un modelo dado por un subconjunto de variables. Esta metodología resulta bastante interesante en estudios QSAR cuando se cuenta con una cantidad de variables (miles) que, de otra forma, resultan inmanejables. En su origen, los estudios QSAR utilizaban una cantidad de variables que se limitaba a factores estéricos (Taft), lipofílicos (log de P) y electrónicos (Hammet) pero la evolución de los métodos computacionales llevó a obtener descriptores como polarizabilidad, de electronegatividad, entre otros. Sin embargo, en la actualidad es posible obtener miles de descriptores teóricos moleculares, para un solo compuesto, en una computadora portátil, lo que hace que un análisis de regresión entre todos los descriptores se vuelva una tarea infinita. En este sentido, la inteligencia artificial ha sido de gran ayuda y los AGs resultan sumamente valiosos en un estudio QSAR de este tipo. Un software que ha resultado exitoso en este tipo de estudios, especializado en QSAR y QSPR (quantitative structure-property relationship) es Moby Digs ®.²³

El resultado de los AGs es una serie de modelos matemáticos producidos por regresiones lineales en donde sobreviven los modelos estadísticamente más aptos o dicho de otra forma, los modelos estadísticamente significativos. Una vez obtenidos los respectivos modelos, que incluyen la ecuación de Hansch, su validación estadística y los descriptores involucrados a los cuales se les atribuye la actividad, se procede al análisis de cada modelo.

II.8.3 Análisis del modelo.

Existen diferentes técnicas para desarrollar un estudio QSAR. El método tradicional incluye el tratamiento estadístico de los datos por métodos multivariados, que incluye el análisis de regresión, en el que se valora la actividad biológica como variable dependiente del conjunto de descriptores moleculares que constituyen las variables independientes. El método más

utilizado es la regresión, en la que se obtendrá la ecuación de la recta, un plano, o de un hiperplano, según sea el número de variables independientes incluida en la expresión²⁴.

La regresión, es un método que se utiliza para predecir el valor de una variable en función de otra u otras. En el análisis de regresión se usa el diagrama de dispersión, en cuyo eje horizontal (eje X), se representan las variables independientes y en el vertical (eje Y) las dependientes. Al representar la relación entre 2 variables, la gráfica de dispersión puede tomar varias formas que van desde algunas funciones matemáticas sencillas hasta las más complejas; la relación más simple sucede cuando se forma una línea recta llamada relación lineal²⁵ (ver Figura 4).



Figura 4. Representación de una grafica de relación lineal entre 2 variables.

El propósito de un análisis de regresión es describir la relación entre una variable dependiente (respuesta o actividad) y una variable independiente (s) (predictores o descriptores), por lo tanto es usada para describir o predecir.

Un modelo de regresión lineal es descrito con la siguiente forma algebraica:

$$y_1 = b_0 + \sum b_1 X_1 + e_1 \quad \text{o} \quad y = X \cdot b + e$$

Donde:

X = Variable independiente

y = Variable dependiente

e = residual

Donde:

y = Actividad

b = Coeficiente de regresión

X = Matriz modelo (son las variables usadas en el modelo)

La precisión predictiva del QSAR, que es uno de los objetivos de este tipo de análisis, es típicamente afectada por 2 aspectos: La selección de descriptores que representen suficientemente la información estructural y la elección de un modelo de predicción específica. El modelo desarrollado puede ser usado para predecir la actividad o propiedad de nuevos compuestos. El QSAR representa modelos predictivos derivados de la aplicación de herramientas estadísticas correlacionando la actividad biológica (incluyendo efectos terapéuticos deseables y efectos secundarios indeseables) o productos químicos (drogas/ sustancias tóxicas/ contaminantes ambientales) con descriptores representantes de estructuras moleculares y/o propiedades²⁰.

II.8.4.- Parámetros estadísticos:

La calidad de un modelo se valora de acuerdo a las siguientes medidas estadísticas:

Coeficiente de determinación, R^2 : La razón de la variación explicada entre la variación total se denomina coeficiente de determinación. Si existen cero variaciones explicadas, la razón es cero. Si hay cero variaciones

inexplicadas, la razón es uno. En otros casos, la razón está entre cero y uno. Dado que la razón nunca es negativa, entonces se denota como R^2 .²⁶

Coefficiente de correlación múltiple al cuadrado R^2 : Que es la varianza total de la respuesta se explica por un modelo de regresión. Se puede calcular a partir de la suma de los cuadrados modelo MSS o de la suma de los cuadrados residual RSS:

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Donde SST es la suma total de cuadrados alrededor de la media. Un valor de uno indica un ajuste perfecto, es decir, un modelo con término de error cero.

En Moby Digs, R^2 se expresa como una cantidad por ciento²³.

F de Fisher: La distribución F, determina si 2 conjuntos tienen diferentes grados de diversidad; por lo tanto representa los niveles de significancia; es decir, determina si la relación observada entre las variables dependientes e independientes se produce por el azar.

Desviación estándar: Es la medida de dispersión más importante y se representa con la letra S, con ella se puede ubicar la posición de una observación con relación a una serie de datos, es decir, nos ayuda a medir la dispersión de los valores respecto a la media.

Q^2 : Nos habla de la capacidad predictoria, deriva de la validación cruzada no es más que la habilidad del modelo, para realizar la predicción.

La Q^2 es la varianza explicada en la predicción:

$$R_{cv}^2 = Q^2 = 1 - \frac{PRESS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_{i/i})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Donde PRESS es la suma de los cuadrados de error de predicción y TSS es la suma total de los cuadrados²³.

Kx: Coeficiente total de correlación muestra una evaluación de la relación existente entre las variables²⁴.

Es la correlación total de los factores predictivos de los modelos:

$$Kx = \frac{\sum_{j=1}^p \left| \frac{\lambda_j}{\sum \lambda_m} - \frac{1}{p} \right|}{\frac{2 * (p - 1)}{p}}$$

Donde λ , son los valores propios obtenidos a partir de la matriz de correlación el conjunto de datos X (n, p), siendo n el numero de objetos y p, el numero de variables²³.

Bootstrap (Q²boot):

La metodología bootstrap debe su nombre y formulación a Bradley Efron (1979). Son básicamente técnicas de simulación que reutilizan los datos observados para constituir un universo del cual extraer repetidas muestras²⁷.

Por medio de esta técnica de validación, el tamaño original de un conjunto de datos (n) se conserva para el conjunto de ensayo, por la selección de n objetos con repetición; de este modo el conjunto de ensayo por lo general se compone de objetos repetidos y la evaluación del conjunto de objetos excluidos. El modelo es calculado sobre el conjunto de ensayo y las respuestas se calculan en el conjunto de evaluación. Todas las diferencias al cuadrado entre la respuesta verdadera y la respuesta calculada de los objetos del

conjunto de evaluación son recopiladas en PRESS (suma de los cuadrados de los residuos predictivos), es la suma de cuadrados de las diferencias entre la respuesta observada y la estimada por técnica de validación.

$$PRESS = \sum_{i=1}^n (y_i - \widehat{y}_{i/l})^2$$

Donde:

$\widehat{y}_{i/l}$: La respuesta del objeto *i-th* estimado, mediante el uso de un modelo obtenido sin utilizar el objeto *i-th*.

Este procedimiento de construcción del conjunto de ensayo y evaluación del conjunto se repite miles de veces, PRESS se suma y el poder de predicción promedio es calculado y la capacidad de predicción promedio es calculado por (Q²boot). Por lo tanto, la validación se realiza mediante la generación aleatoria del conjunto de ensayo con repetición de muestras y entonces la evaluación de la respuesta predictiva de las muestras no incluida en el conjunto de ensayo²³.

II.8.5.- Técnicas de validación:

Estas constituyen una herramienta fundamental para la evaluación de la validez de los modelos obtenidos mediante métodos de regresión multivariada. Las técnicas de validación se utilizan para comprobar el poder de predicción de los modelos, es decir, para dar una medida de su capacidad para realizar predicciones confiables. Una condición necesaria para la validez de un modelo de regresión es que el coeficiente de correlación múltiple R² sea lo más cerca a 1 y el error estándar de la estimación debe ser pequeño; sin embargo estas condiciones no son suficientes para la validez de un modelo. Las técnicas de validación disponibles en Moby Digs son:

- Leave-one-out
- Y-scrambling

Leave-one-out:

El más simple y general procedimiento de validación cruzada es la técnica de leave-one-out (LOO technique), donde cada objeto es excluido, uno a la vez. En este caso, dan n objetos, n modelos reducidos tienen que ser calculados.

Para cada conjunto de datos reducidos, el modelo se calcula y las respuestas para el objeto eliminado se predicen a partir del modelo. Las diferencias al cuadrado entre la respuesta verdadera y la respuesta predicha para el objeto excluido se añaden a PRESS.

El procedimiento de validación cruzada leave-one-out se determina mediante la siguiente ecuación:²⁸

$$Q_{tr}^2 = 1 - \frac{\sum_{i=1}^{N_{tr}} (y_{i,exp} - y_{i,pred})^2}{\sum_{i=1}^{N_{tr}} (y_{i,exp} - \bar{y}_{i,exp})^2}$$

Donde:

N_{tr} = El número total de objetos del conjunto de entrenamiento

$y_{i,exp}$ = Valor experimental

$y_{i,pred}$ = Valor calculado

$\bar{y}_{i,exp}$ = El promedio del valor de la respuesta del conjunto de entrenamiento

Y-scrambling:

Esta técnica de validación es adoptada para comprobar los modelos de correlación con el azar, es decir, modelos donde las variables independientes son

al azar correlacionadas con las variables de respuesta. El ensayo se realiza mediante el cálculo de la calidad del modelo (por lo general R^2 o, mejor, Q^2) modificar aleatoriamente la secuencia del vector de respuesta y , es decir, asignando a cada objeto de una respuesta seleccionada aleatoriamente de las respuestas verdaderas. Cada aleatorización se caracteriza en términos de la correlación de la respuesta codificada con los datos imperturbables ($R^2_{yy'}$). Si el modelo original no tiene correlación azar, existe una diferencia significativa en la calidad del modelo original y la asociada con un modelo obtenido con respuestas aleatorias. El procedimiento se repite varios cientos de veces. Una vez que el modelo de validación se ha realizado los parámetros de aleatorización Y -($a(R^2)$) y un (Q^2) se calculan como las intersecciones de las ecuaciones:

$$R^2 = b_0 + bR^2(yy') \quad b_0 = a(R^2)$$

$$Q^2 = b_0 + bQ^2(yy') \quad b_0 = a(Q^2)$$

Los modelos que son inestables, son caracterizados por valores de intersección altos. Modelos estables tienen valores de intersección bajos²³.

II.8.6.- Descriptores:

Hasta el momento se ha incursionado en el procedimiento estadístico de QSAR, y en este punto resulta importante definir las variables independientes como propiedades moleculares que son la causa molecular de la actividad biológica. Estos son los descriptores moleculares son información química resultado de lógica y procedimientos matemáticos, por lo tanto deben contener información química y ser derivados de procedimientos bien establecidos que les confieran la posibilidad de calcular propiedades para cualquier conjunto de

moléculas²⁹. Es evidente que un descriptor o un pequeño número de descriptores no pueden representar totalmente la complejidad molecular o el modelo de todas las respuestas físico-químicas y las interacciones biológicas, por lo tanto se entiende que se trabaja con los modelos aproximados, aclarando que no es sinónimo de inútil, sino de conveniente.

El campo de los descriptores involucra diferentes teorías, la definición de cada uno de ellos requiere de conocimientos generales de álgebra, química computacional, estadística y quimiometría. Los descriptores moleculares calculados por DRAGON están divididos en 20 familias³⁰.Tabla 3

Tabla 3. Familias de descriptores, su nomenclatura y número de descriptores que lo conforman obtenidos por el programa DRAGON.

<i>ID Block</i>	<i>Block description</i>	<i>No. descriptores</i>
1	descriptores constitucionales	48
2	descriptores topológicos	119
3	“walk and path counts”*	47
4	Indicadores de conectividad	33
5	indicadores de información	47
6	autocorrelaciones “D	96
7	Indicadores “edge adjacency”*	107
8	Descriptores “Burden eigenvalue”	64
9	Indicadores topológicos y de carga	21
10	indicadores basados en eigenvalores	44
11	Perfiles Randico-moleculares	41
12	descriptores geométricos	74
13	Descriptores RDF	150
14	Descriptores 3D-MoRSE	160
15	Descriptores WHIM	99
16	Descriptores GETAWAY	197
17	Conteo de grupos funcionales	154
18	Fragmentos “átomo-central”	120
19	Descriptores de carga	14
20	Propiedades moleculares	29

* Nótese que algunos de ellos no cuentan con traducción al español y se suelen utilizar con su nombre original.

El método de QSAR no solo se ha utilizado para la predicción de la actividad de una serie de compuestos, sino que también se ha aprovechado que analiza la estructura para vincularla con la citotoxicidad ya que todo está en base a las características físico químicas de la estructura de la molécula, caso del trabajo de Marcus T. Scotti y colaboradores, titulado Quantitative structure–activity relationship of sesquiterpene lactones with cytotoxic activity³¹ y el trabajo de Liane Saíz-Urra, Maykel Pérez González y Marta Teijeira, titulado QSAR studies about cytotoxicity of benzophenazines with dual inhibition toward both topoisomerases I and II: 3D-MoRSE descriptors and statistical considerations about variable selection³²

III.- Objetivo General:

Obtener un modelo matemático a partir de un estudio QSAR que permita proponer estructuras novedosas, con potencial actividad contra Trypanosoma cruzi.

III.1.-Objetivos particulares:

1. Llevar a cabo un estudio QSAR con técnicas de algoritmos genéticos que permita explicar la actividad de una serie de benzacepinas activas contra Trypanosoma cruzi in vitro.
2. Elaborar un estudio QSAR que permita explicar la citotoxicidad de dichos compuestos activos contra células vero.
3. Proponer nuevos compuestos análogos activos contra Trypanosoma cruzi e inactivos contra células vero, utilizando el o los modelos matemáticos obtenidos.

IV.- METODOLOGÍA:

La metodología utilizada fue la siguiente (Figura 5):

1.- Modelado molecular y optimización geométrica. Cada uno de los compuestos de la serie de exploración fue modelado y se obtuvo su geometría óptima usando mecánica molecular (MM) y métodos semiempíricos (AM1) con el software **HyperChem 8.0.3**, bajo los siguientes parámetros:

Para el cálculo de dinámica molecular la simulación es a 300 K, en vacío; con un periodo de recopilación de datos de 1ps. En el caso del cálculo de optimización geométrica se utiliza como base el algoritmo Polak-Ribiere³³, con parámetros de condiciones de terminación de 0.1 kcal/ (Å mol) o 525 ciclos como máximo.

2.- Las moléculas modeladas se exportan al software **DRAGON versión 5.3-2005**, el cual fue usado para el cálculo de los descriptores que serán usados para evaluar la relación de las propiedades moleculares y su actividad. El programa calcula 1664 descriptores organizados en 20 familias.

3.- Se eligen las familias a utilizarse que, en este caso, se consideran todas, siendo para el caso de la familia 20 el marcaje el valor de la actividad experimental de los compuestos, para ejecutarlos en el software **Moby Digs versión 1.0-2004** que calcula modelos de regresión, usando algoritmos genéticos para la selección de variables para obtener modelos de predicción. Donde se cargan 2 archivos o evoluciones con 10 pops cada uno en el caso de la primera corrida o evolución para las 2 siguientes evoluciones solo se agrega un archivo con 10 pops, colocando los siguientes parámetros:

Tamaño de la población: 100 en la primera evolución, 70 en la segunda evolución y 50 en la tercera evolución.

Variables máximas permitidas en el modelo: se coloca el número 2, esto es para las 3 evoluciones que se corrieron.

El factor reproducción/mutación: tiene el valor de 0.7 para las 3 evoluciones.

En el parámetro que indica el máximo de descriptores por modelo: se indica 7 para la primera evolución, 5 para la segunda y 3 para la tercera.

El número de ciclos que conforman es 2000 para las 3 evoluciones.

El número de ciclos para ingreso de variables es de: 20,000.

Cuando se finaliza cada evolución se indica cuantos modelos son los que se necesitan guarda, quedando que al terminar la primera evolución se solicitaran 10 por cada pop, para la segunda evolución 50 por cada pop y en la tercera evolución 50 por cada pop.

4.- Se lleva a cabo una corroboración de los datos estadísticos de manera manual utilizando el software Microsoft Excel ®.³⁴

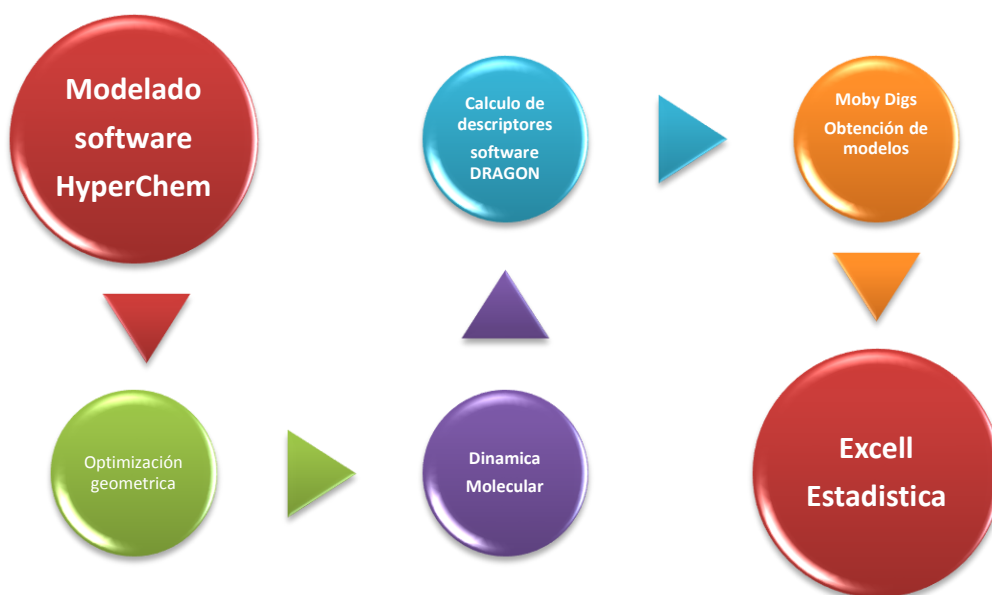
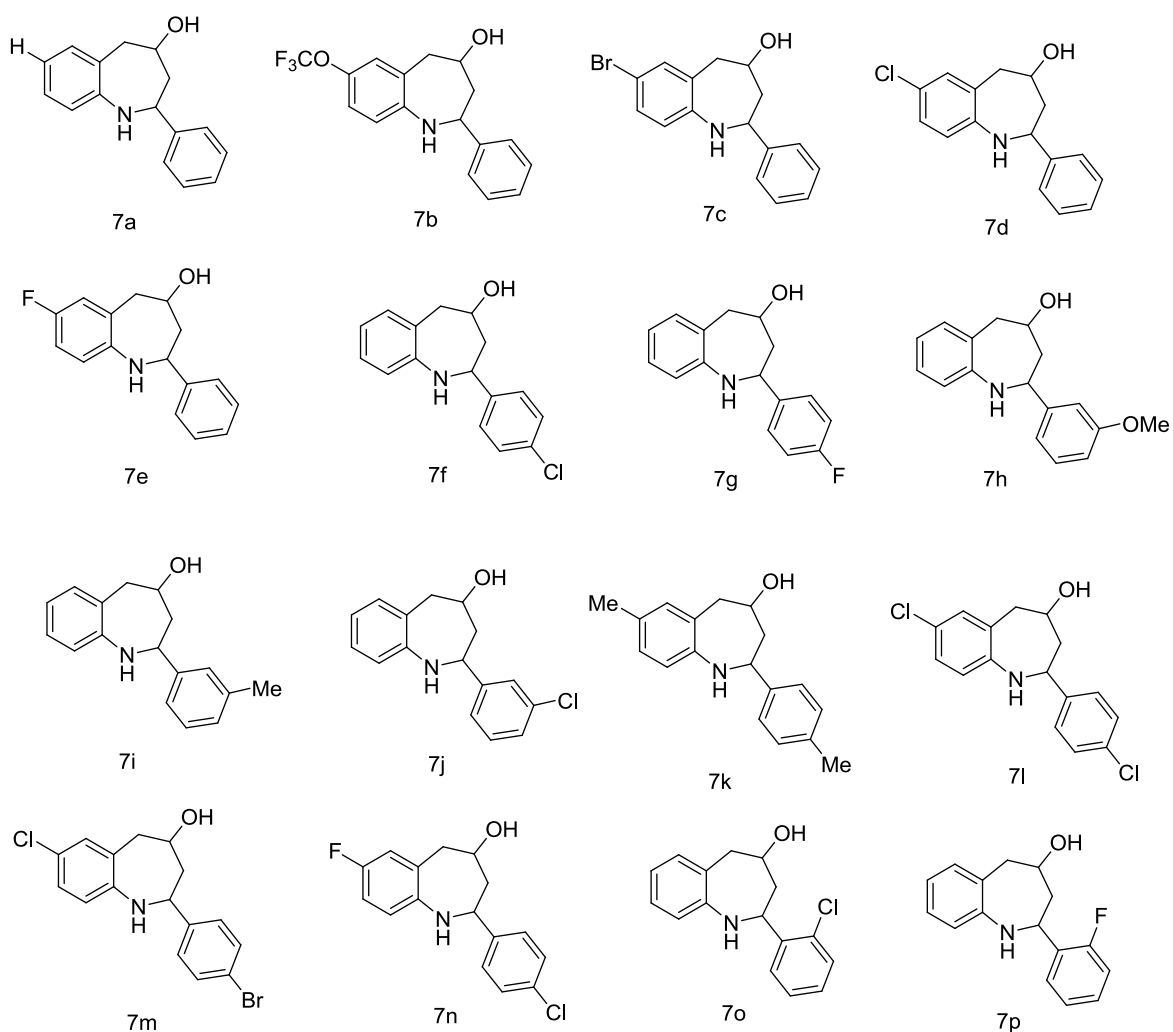


Figura 5. Diagrama de a metodología para la obtención de los modelos y valores estadísticos.

V.- RESULTADOS:

V.1.- Obtención de modelos QSAR de las benzacepinas serie 7a-w con actividad contra *trypanosoma cruzi*.

Los compuestos sometidos al estudio, tomados de la base de datos reportada, se resumen en la Figura 6:



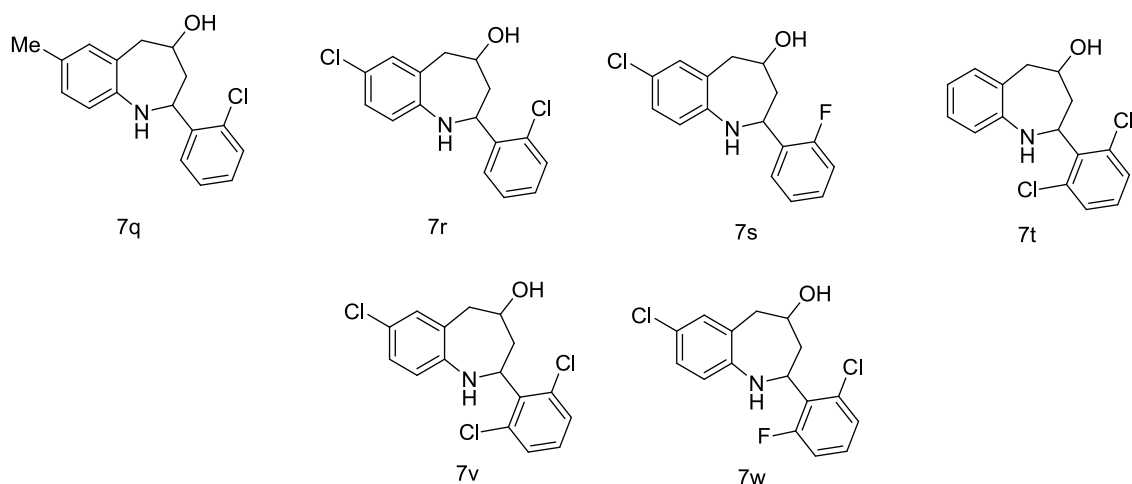


Figura 6. Compuestos sometidos a exploración para obtención de los modelo QSAR.

Se obtuvieron 50 modelos de regresión (Tabla 4) que, de acuerdo a las restricciones introducidas al programa, fueron los mejores. Cada modelo contiene determinados descriptores (propiedades de las moléculas), cuya participación en la ecuación está validada estadísticamente por los valores de la R^2 , Q^2 boot, Kx, F y s^{35} . Además de contar con 5 moléculas que fueron excluidos para posteriormente validar la técnica al aplicar los mejores modelos y observar la relación entre los valores calculados y los experimentales.

Tabla 4. Resumen de los 50 “mejores modelos” con sus respectivos descriptores y su validación estadística.

Modelos	R^2	Q^2	Q^2 boot	Kx	F	s
X5v BEHp7 DP08	90.1	84.26	77.06	51.36	33.38	0.079
X5v BEHp7 DP11	89.82	83.93	77.19	53.67	32.34	0.08
Mor21v Mor10e Mor24e	90.28	83.52	76.5	36.52	34.06	0.078
X5v BEHp7 DP12	89.54	83.44	75.38	54.19	31.38	0.081
Mor10u Mor24e	89.81	83.29	77.9	38.15	32.33	0.08

Mor21p						
Mor10e Mor24e Mor21p	91.03	83.01	79.38	38.44	37.19	0.075
X5v BEHp7 DP13	89.23	82.89	74.83	54.63	30.38	0.083
X5v BEHp7 DP15	88.59	81.78	74.91	55.32	28.47	0.085
SPAN Mor10u Mor19e	89.4	80.71	77.83	8.12	30.91	0.082
IC0 Mor21v Mor10e	86.9	79.99	74.36	31.88	24.32	0.091
AAC Mor21v Mor10e	86.9	79.99	73.79	31.88	24.32	0.091
SPAN Mor10u Mor19u	88.31	78.52	75.12	7.72	27.71	0.086
BIC0 Mor10u Mor21p	86	77.59	72.18	28.83	22.53	0.094
IC0 Mor10u Mor21p	86.26	77.39	73.3	30.14	23.02	0.093
AAC Mor10u Mor21p	86.26	77.39	71.58	30.14	23.02	0.093
Modelos	R²	Q²	Q²boot	Kx	F	s
SIC0 Mor10u Mor21p	85.66	76.97	63.43	28.82	21.9	0.095
BIC0 Mor21v Mor10e	85.29	75.99	70.1	29.25	21.25	0.096
BIC0 Mor10u Mor21v	85.82	75.64	70.34	27.97	22.19	0.095
X5v BEHv7 DP08	89.5	75.57	72.24	52.44	31.26	0.081
Mor10u Mor21v nCb-	84.53	75.55	67.8	19.85	20.04	0.099
Mor10u Mor21v nCbH	84.53	75.55	67.5	19.85	20.04	0.099
Mor10u Mor21v C- 024	84.53	75.55	68.57	19.85	20.04	0.099
SIC0 Mor10u Mor21v	85.64	75.51	70.24	27.86	21.87	0.095
AAC Mor10e Mor21p	85.01	75.15	69.8	31.22	20.8	0.097
IC0 Mor10e Mor21p	85.01	75.15	70.24	31.22	20.8	0.097
SIC0 Mor21v Mor10e	84.72	74.57	67.92	28.99	20.33	0.098
AAC Mor10u Mor21v	85.31	72.91	68.06	29.75	21.29	0.096
IC0 Mor10u Mor21v	85.31	72.91	68.93	29.75	21.29	0.096
X5v BEHv7 DP11	88.5	71.7	68.14	54.66	28.22	0.085
Mor10u Mor21p C- 024	83.01	71.28	63.85	20.85	17.91	0.104

“Diseño por QSAR de compuestos potencialmente activos contra Trypanosoma cruzi”

X5v BEHv7 DP12	88.06	70.59	66.59	55.16	27.04	0.087
X5v BEHv7 DP13	87.62	69.62	65.83	55.59	25.96	0.088
SPAN Mor19u Mor10e	84.16	69.18	65.26	7.06	19.48	0.1
SPAN Mor10e Mor19e	84.37	69.04	65.39	7.38	19.79	0.099
X5v BEHp7 RDF025m	80.47	68.97	54.95	24.44	15.1	0.111
ATS5m Mor10e Mor21p	80.82	68.26	58.25	27.51	15.45	0.11
GGI4 Mor10u Mor21p	79.36	68.14	59.51	28.79	14.1	0.114
X5v BEHv7 DP15	86.78	68.04	63	56.25	24.07	0.091
ATS5m RDF025m Mor10e	80.41	67.31	62.81	14.99	15.05	0.111
SPAN RDF025m R5m	77.51	66.81	55.23	22.01	12.64	0.119
RDF025m Mor10e R5m	79.9	66.78	60.18	22.44	14.57	0.113
ATS5m Mor21v Mor10e	80.29	66.58	57.8	30.05	14.94	0.112
Mor21v Mor10e nC b H	81.62	65.87	55.88	24.29	16.29	0.108
X5v BEHv7 RDF025m	79.4	65.41	53.49	25.7	14.13	0.114
RDF025m R5m	75.53	65.31	63.43	21.3	18.52	0.119
BEHm1 RDF025m R5m	76.13	65.3	0	27.39	11.69	0.123
RDF025m Mor10u R5m	79.75	65.01	59.74	23.87	14.44	0.113
ATS5m Mor10e	71.26	59.28	57.94	18.27	14.88	0.129
ATS5m	56.48	43.85	47.16	0	16.87	0.153
R5m	42.24	27.62	33.28	0	9.51	0.176

Los valores que más interesan debido a su significado se encuentran resaltados en la Tabla 4. El primer parámetro a considerar es la R^2 donde se toman en cuenta los valores más cercanos a 100 ya que esto indica la variabilidad que una distribución tiene respecto de su media siendo que entre más se acerque a ésta indicará una esperanza matemática o valor esperado. Enseguida, se considera el valor de la Q^2 boot, tomando en cuenta como buenos valores los cercanos a 100, ya que esto expresa la capacidad de predicción promedio y por

último los valores de Kx; donde se toma en cuenta que valores menores de 50 auguran correlación entre descriptores, recordando que variables dependientes (descriptores) correlacionadas en el producto de la regresión producen ambigüedad y por lo tanto mala interpretación estadística. Esto es debido a que los descriptores deben ser independientes y con ello más significativo para determinar la influencia de una propiedad³⁶.

V.2.- Análisis de los modelos sin correlación:

Los 50 modelos se someten a un análisis estadístico independiente del programa que los calculó mediante algún programa de cálculo básico como Excel. Los resultados de correlación entre descriptores por cada modelo se resumen en las tablas 5.1-5.39.

Tabla 5.1 Correlación de los descriptores del modelo 3

Modelo 3			
	Mor21v	Mor10e	Mor24e
Mor21v	1	0.0966044	0.33440164
Mor10e		1	0.02491323
Mor24e			1

Tabla 5.2 Correlación de los descriptores del modelo 5

Modelo 5			
	Mor10u	Mor24e	Mor21p
Mor10u	1	0.02884263	0.08242895
Mor24e		1	0.39095652
Mor21p			1

Tabla 5.3 Correlación de los descriptores del modelo 6

Modelo 6			
	Mor10e	Mor24e	Mor21p
Mor10e	1	0.02491323	0.09576534
Mor24e		1	0.39095652
Mor21p			1

Tabla 5.4 Correlación de los descriptores del modelo 9

Modelo 9			
	SPAN	Mor10u	Mor19e
SPAN	1	0.00011269	0.0157608
Mor10u		1	0.00892581
Mor19e			1

Tabla 5.5 Correlación de los descriptores del modelo 10

Modelo 10			
	IC0	Mor21v	Mor10e
IC0	1	0.07448307	0.13697334
Mor21v		1	0.0966044
Mor10e			1

Tabla 5.6 Correlación de los descriptores del modelo 11

Modelo 11			
	AAC	Mor21v	Mor10e
AAC	1	0.07448307	0.13697334
Mor21v		1	0.0966044
Mor10e			1

Tabla 5.7 Correlación de los descriptores del modelo 12

Modelo 12			
	SPAN	Mor10u	Mor19u
SPAN	1	0.00011269	0.00867982
Mor10u		1	0.01356008
Mor19u			1

Tabla 5.8 Correlación de los descriptores del modelo 13

Modelo 13			
	BIC0	Mor10u	Mor21p
BIC0	1	0.12085245	0.05099488
Mor10u		1	0.08242895
Mor21p			1

Tabla 5.9 Correlación de los descriptores del modelo 14

Modelo 14			
	IC0	Mor10u	Mor21p
IC0	1	0.12988983	0.06416977
Mor10u		1	0.08242895
Mor21p			1

Tabla 5.10 Correlación de los descriptores del modelo 15

Modelo 15			
	AAC	Mor10u	Mor21p
AAC	1	0.12988983	0.06416977
Mor10u		1	0.08242895
Mor21p			1

Tabla 5.11 Correlación de los descriptores del modelo 16

Modelo 16			
	SIC0	Mor10u	Mor21p
SIC0	1	0.12030048	0.05121143
Mor10u		1	0.08242895
Mor21p			1

Tabla 5.12 Correlación de los descriptores del modelo 17

Modelo 17			
	BIC0	Mor21v	Mor10e
BIC0	1	0.05340193	0.11060736
Mor21v		1	0.0966044
Mor10e			1

Tabla 5.13 Correlación de los descriptores del modelo 18

Modelo 18			
	BIC0	Mor10u	Mor21v
BIC0	1	0.12085245	0.05340193
Mor10u		1	0.06525369
Mor21v			1

Tabla 5.14 Correlación de los descriptores del modelo

Modelo 20			
	Mor10u	Mor21v	nCb-
Mor10u	1	0.06525369	0.04361613
Mor21v		1	0.01536414
nCb-			1

Tabla 5.15 Correlación de los descriptores del modelo 21

Modelo 21			
	Mor10u	Mor21v	nCbH
Mor10u	1	0.06525369	0.04361613
Mor21v		1	0.01536414
nCbH			1

Tabla 5.16 Correlación de los descriptores del modelo 22

Modelo 22			
	Mor10u	Mor21v	C-024
Mor10u	1	0.06525369	0.04361613
Mor21v		1	0.01536414
C-024			1

Tabla 5.17 Correlación de los descriptores del modelo 23

Modelo 23			
	SIC0	Mor10u	Mor21v
SIC0	1	0.12030048	0.05221694
Mor10u		1	0.06525369
Mor21v			1

Tabla 5.18 Correlación de los descriptores del modelo 24

Modelo 24			
	AAC	Mor10e	Mor21p
AAC	1	0.13697334	0.06416977
Mor10e		1	0.09576534
Mor21p			1

Tabla 5.19 Correlación de los descriptores del modelo 25

Modelo 25			
	IC0	Mor10e	Mor21p
IC0	1	0.13697334	0.06416977
Mor10e		1	0.09576534
Mor21p			1

Tabla 5.20 Correlación de los descriptores del modelo 26

Modelo 26			
	SIC0	Mor21v	Mor10e
SIC0	1	0.10706544	0.10706544
Mor21v		1	0.0966044
Mor10e			1

Tabla 5.21 Correlación de los descriptores del modelo 27

Modelo 27			
	AAC	Mor10u	Mor21v
AAC	1	0.12988983	0.07448307
Mor10u		1	0.06525369
Mor21v			1

Tabla 5.22 Correlación de los descriptores del modelo 28

Modelo 28			
	IC0	Mor10u	Mor21v
IC0	1	0.12988983	0.07448307
Mor10u		1	0.06525369
Mor21v			1

Tabla 5.23 Correlación de los descriptores del modelo 30

Modelo 30			
	Mor10u	Mor21p	C-024
Mor10u	1	0.08242895	0.04361613
Mor21p		1	0.01390894
C-024			1

Tabla 5.24 Correlación de los descriptores del modelo 33

Modelo 33			
	SPAN	Mor19u	Mor10e
SPAN	1	0.00867982	0.00867982
Mor19u		1	0.00315018
Mor10e			1

Tabla 5.25 Correlación de los descriptores del modelo 34

Modelo 34			
	SPAN	Mor10e	Mor19e
SPAN	1	0.00315018	0.0157608
Mor10e		1	0.00056942
Mor19e			1

Tabla 5.26 Correlación de los descriptores del modelo 35

Modelo 35			
	X5v	BEHp7	RDF025m
X5v	1	0.20337521	0.00776866
BEHp7		1	0.01085244
RDF025m			1

Tabla 5.27 Correlación de los descriptores del modelo 36

Modelo 36			
	ATS5m	Mor10e	Mor21p
ATS5m	1	0.03336955	0.1064573
Mor10e		1	0.09576534
Mor21p			1

Tabla 5.28 Correlación de los descriptores del modelo 37

Modelo 37			
	GGI4	Mor10u	Mor21p
GGI4	1	0.04892092	0.12296178
Mor10u		1	0.08242895
Mor21p			1

Tabla 5.29 Correlación de los descriptores del modelo 39

Modelo 39			
	ATS5m	RDF025m	Mor10e
ATS5m	1	0.00786422	0.03336955
RDF025m		1	0.0299952
Mor10e			1

Tabla 5.30 Correlación de los descriptores del modelo 40

Modelo 40			
	SPAN	RDF025m	R5m
SPAN	1	0.04495896	0.05524915
RDF025m		1	0.04535408
R5m			1

Tabla 5.31 Correlación de los descriptores del modelo 41

Modelo 41			
	RDF025m	Mor10e	R5m
RDF025m	1	0.0299952	0.04535408
Mor10e		1	0.07961541
R5m			1

Tabla 5.32 Correlación de los descriptores del modelo 42

Modelo 42			
	ATS5m	Mor21v	Mor10e
ATS5m	1	0.1565542	0.03336955
Mor21v		1	0.0966044
Mor10e			1

Tabla 5.33 Correlación de los descriptores del modelo 43

Modelo 43			
	Mor21v	Mor10e	nCbH
Mor21v	1	0.0966044	0.01536414
Mor10e		1	0.07942741
nCbH			1

Tabla 5.34 Correlación de los descriptores del modelo 44

Modelo 44			
	X5v	BEHv7	RDF025m
X5v	1	0.22897338	0.00776866
BEHv7		1	0.01061555
RDF025m			1

Tabla 5.35 Correlación de los descriptores del modelo 45

Modelo 45		
	RDF025m	R5m
RDF025m	1	0.04535408
R5m		1

Tabla 5.36 Correlación de los descriptores del modelo 46

Modelo 46			
	BEHm1	RDF025m	R5m
BEHm1	1	0.07219507	0.11247626
RDF025m		1	0.04535408
R5m			1

Tabla 5.38 Correlación de los descriptores del modelo 47

Modelo 47			
	RDF025m	Mor10u	R5m
RDF025m	1	0.02231332	0.04535408
Mor10u		1	0.11495433
R5m			1

Tabla 5.39 Correlación de los descriptores del modelo 48

Modelo 48		
	ATS5m	Mor10e
ATS5m	1	0.03336955
Mor10e		1

Una vez elegidos los modelos que demostraban independencia de los descriptores participantes, se corroboraron, mediante una regresión lineal, los parámetros de R^2 , F de Fisher, buscando los valores de 95% de confianza, constante de la ecuación, así como la desviación estándar. Otro parámetro de elección de modelos fue la determinación de la actividad calculada por cada modelo contra la observada experimentalmente debiendo existir congruencia³⁷. De esta manera, se descartan demasiados modelos resultando útiles únicamente los modelos 3, 5, 6, 9 y 12; los cuales se analizan y detallan a continuación.

Modelo 3:

$$\text{Log } 1/C = -2.428\text{Mor21v } (\pm 0.289) + 0.369\text{Mor10e } (\pm 0.050) + 1.067\text{Mor24e } (\pm 0.185) - 0.027 (\pm 0.181)$$

$$n = 15; R^2 = 90.2; F = 34.05; s = 11; Q^2 = 83.52; Kx = 36.52$$

Tabla 6. Valores de los descriptores de las compuestos, así como la actividad calculada para el modelo 3.

Modelo 3	Mor21v	Mor10e	Mor24e				
Compuesto	Mor21v	Mor10e	Mor24e	Actividad μm	Log 1/C exp.	Log 1/C cal.	Act. Calc. μm
7b	-0.328	0.663	0.361	26.59	1.42	1.39	25.11
7c	-0.432	-0.843	0.502	15.71	1.19	1.24	17.65
7d	-0.546	0.142	0.234	49.58	1.69	1.60	39.93
7f	-0.579	-0.13	0.472	65.44	1.81	1.83	68.40
7h	-0.534	-0.45	0.304	30.06	1.47	1.42	26.81
7j	-0.528	-0.688	0.213	19.62	1.29	1.22	16.93
7l	-0.327	-0.215	0.471	16.45	1.21	1.19	15.51
7m	-0.553	-0.107	0.225	27.21	1.43	1.51	32.87
7n	-0.623	0.167	0.201	45.93	1.66	1.76	57.86
7o	-0.599	-0.375	0.272	49.73	1.69	1.57	37.99
7q	-0.636	-0.996	0.121	22.43	1.35	1.27	19.01
7r	-0.537	-0.521	0.145	15.34	1.18	1.23	17.36
7s	-0.556	-0.347	0.058	17.31	1.23	1.25	18.08
7v	-0.517	-0.77	0.298	14.84	1.17	1.26	18.30
7w	-0.535	-0.501	0.096	14.46	1.16	1.190	15.48

Tabla 7. Valores de correlación entre los descriptores del modelo 3.

Correlación			
	Mor21v	Mor10e	Mor24e
Mor21v	1	0.0966044	0.33440164
Mor10e		1	0.02491323
Mor24e			1

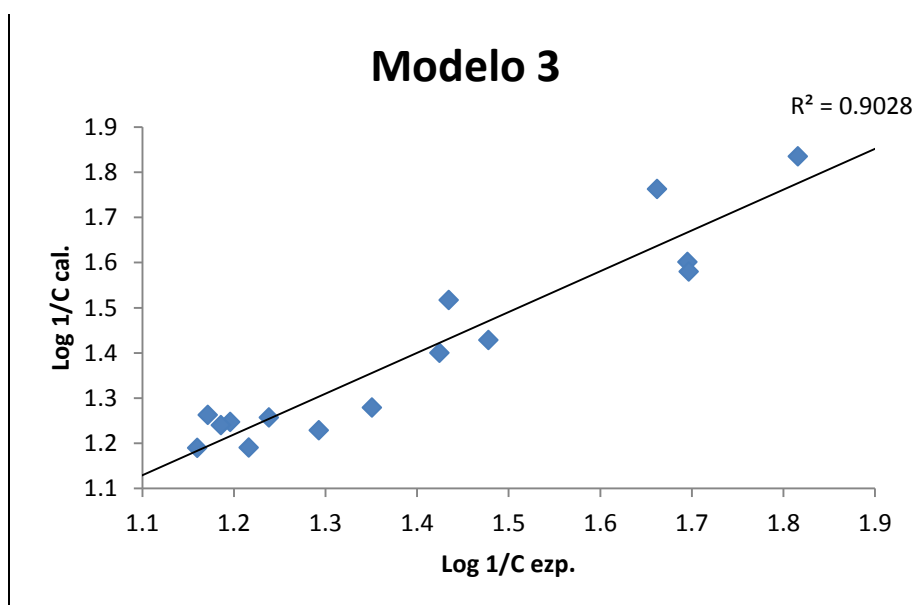


Figura 7. Se representa la relación entre los valores de la actividad experimental y la calculada.

Tabla 8. Valores obtenidos en la estimación lineal para el modelo 3.

Valores de los coeficientes de regresión				
	m2	m1	m	C
	1.067	0.369	-2.428	-0.026
	0.185	0.050	0.289	0.181
R²	90.2	0.078		
F	34.058	11	S	
	0.628	0.067		

Modelo 5:

$$\text{Log } 1/C = 0.423 \text{ Mor10u } (\pm 0.0599) + 1.145 \text{ Mor24e } (\pm 0.198) - 2.325 \text{ Mor21p } (\pm 0.289) + -0.044 (\pm 0.198)$$

$$n = 15; R^2 = 89.8; F = 32.33; s = 11; Q^2 = 83.29; Kx = 38.15$$

Tabla 9. Valores de los descriptores de los compuestos, así como la actividad calculada por el modelo 5.

Modelo 5	Mor10u	Mor24e	Mor21p				
Compuesto	Mor10u	Mor24e	Mor21p	Actividad μm	Log1/C exp	Log1/C cal.	Act cal. μm
7b	0.181	0.361	-0.395	26.59	1.42	1.36	23.10
7c	-0.898	0.502	-0.473	15.71	1.19	1.24	17.77
7d	-0.012	0.234	-0.597	49.58	1.69	1.60	40.38
7f	-0.223	0.472	-0.608	65.44	1.81	1.81	65.31
7h	-0.532	0.304	-0.576	30.06	1.47	1.41	26.14
7j	-0.767	0.213	-0.569	19.62	1.29	1.19	15.75
7l	-0.144	0.471	-0.337	16.45	1.21	1.21	16.48
7m	-0.14	0.225	-0.606	27.21	1.43	1.56	36.53
7n	0.091	0.201	-0.651	45.93	1.66	1.73	54.65
7o	-0.385	0.272	-0.65	49.73	1.69	1.61	41.21
7q	-1.078	0.121	-0.716	22.43	1.35	1.30	20.05
7r	-0.599	0.145	-0.58	15.34	1.18	1.21	16.45
7s	-0.38	0.058	-0.6	17.31	1.23	1.25	18.02
7v	-0.767	0.298	-0.56	14.84	1.17	1.27	18.78
7w	-0.582	0.096	-0.589	14.46	1.16	1.18	15.42

Tabla 10. Valores de correlación entre los descriptores del modelo 5.

Correlación			
	Mor10u	Mor24e	Mor21p
Mor10u	1	0.02884263	0.08242895
Mor24e		1	0.39095652
Mor21p			1

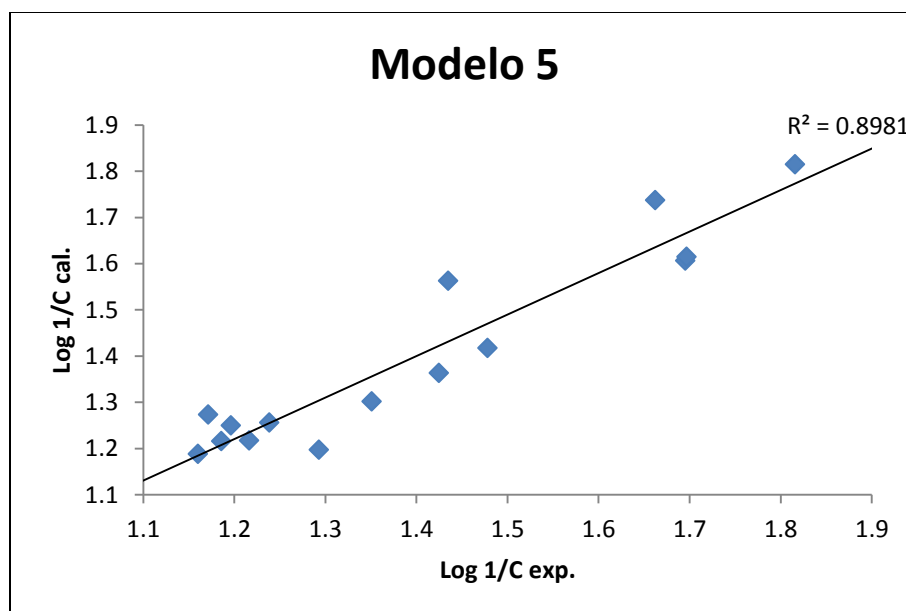


Figura 8. Representa la relación entre los valores de la actividad experimental y la calculada.

Tabla 11. Valores obtenidos en la estimación lineal para el modelo 5.

Valores de los coeficientes de regresión				
	m2	m1	M	C
	-2.32	1.14	0.42	-0.04
	0.28	0.19	0.05	0.19
R^2	89	0.08		
F	32.33	11	S	
	0.62	0.07		

Modelo 6:

$$\text{Log } 1/C = 0.370 \text{ Mor10e } (\pm 0.048) + 1.195 \text{ Mor24e } (\pm 0.186) - 2.416 \text{ Mor21p } (\pm 0.274) - 0.163 (\pm 0.187)$$

$$n = 15; R^2 = 91.0; F = 37.18; s = 11; Q^2 = 83.01; Kx = 38.44$$

Tabla 12. Valores de los descriptores de los compuestos, así como la actividad calculada para el modelo 6

Modelo 6	Mor10e	Mor24e	Mor21p				
Compuesto	Mor10e	Mor24e	Mor21p	termina 20	Log1/C exp	Log1/C cal.	Act. Cal.
7b	0.663	0.361	-0.395	26.59	1.42	1.46	29.40
7c	-0.843	0.502	-0.473	15.71	1.19	1.26	18.52
7d	0.142	0.234	-0.597	49.58	1.69	1.61	40.91
7f	-0.13	0.472	-0.608	65.44	1.81	1.82	66.39
7h	-0.45	0.304	-0.576	30.06	1.47	1.42	26.63
7j	-0.688	0.213	-0.569	19.62	1.29	1.21	16.28
7l	-0.215	0.471	-0.337	16.45	1.21	1.13	13.63
7m	-0.107	0.225	-0.606	27.21	1.43	1.53	33.93
7n	0.167	0.201	-0.651	45.93	1.66	1.71	51.54
7o	-0.375	0.272	-0.65	49.73	1.69	1.59	39.25
7q	-0.996	0.121	-0.716	22.43	1.35	1.34	22.02
7r	-0.521	0.145	-0.58	15.34	1.18	1.21	16.55
7s	-0.347	0.058	-0.6	17.31	1.23	1.22	16.89
7v	-0.77	0.298	-0.56	14.84	1.17	1.26	18.24
7w	-0.501	0.096	-0.589	14.46	1.16	1.19	15.46

Tabla 13. Valores de correlación entre los descriptores del modelo 6.

Correlación			
	Mor10e	Mor24e	Mor21p
Mor10e	1	0.02491323	0.09576534
Mor24e		1	0.39095652
Mor21p			1

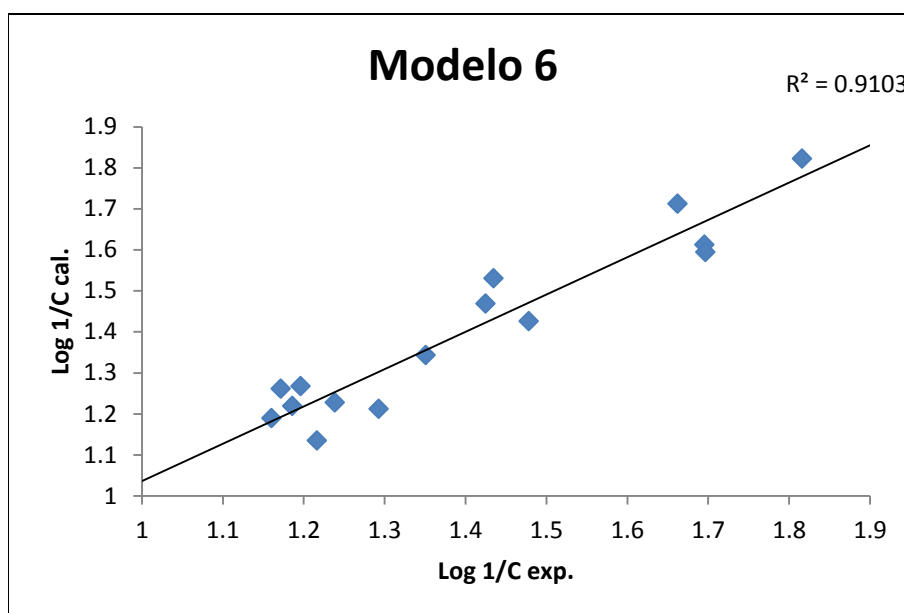


Figura 9. Se representa la relación entre los valores de la actividad experimental y calculada.

Tabla 14. Valores obtenidos en la estimación lineal para el modelo 6.

Valores de los coeficientes de regresión				
	m2	m1	m	C
	-2.41	1.19	0.37	-0.16
	0.27	0.18	0.048	0.187
R ²	91	0.07		
F	37.18	11	S	
	0.63	0.06		

Modelo 9:

$$\text{Log } 1/C = -0.382 \text{ SPAN } (\pm 0.072) + 0.360 \text{ Mor10u } (\pm 0.058) + 0.820 \text{ Mor19e } (\pm 0.126) + 3.094 (\pm 0.459)$$

$$n = 15; R^2 = 89.4; F = 30.91; s = 11; Q^2 = 80.71; Kx = 8.12$$

Tabla 15. Valores de los descriptores de los compuestos, así como la actividad calculada para el modelo 9.

Modelo 9	SPAN	Mor10u	Mor19e				
Compuesto	SPAN	Mor10u	Mor19e	Actividad μm	Log1/C exp	Log1/C cal.	Act.Cal. μm
7b	6.646	0.181	1.02	26.59	1.42	1.45	28.60
7c	6.929	-0.898	1.285	15.71	1.19	1.17	15.01
7d	5.971	-0.012	1.163	49.58	1.69	1.76	57.83
7f	6.318	-0.223	1.441	65.44	1.81	1.78	60.47
7h	6.362	-0.532	1.298	30.06	1.47	1.53	34.34
7j	6.07	-0.767	0.979	19.62	1.29	1.30	19.99
7l	6.508	-0.144	0.754	16.45	1.21	1.17	14.91
7m	6.728	-0.14	1.046	27.21	1.43	1.33	21.40
7n	6.147	0.091	1.083	45.93	1.66	1.66	46.38
7o	5.732	-0.385	0.975	49.73	1.69	1.56	36.69
7q	6.236	-1.078	1.145	22.43	1.35	1.26	18.25
7r	6.397	-0.599	1.048	15.34	1.18	1.29	19.63
7s	6.51	-0.38	0.924	17.31	1.23	1.22	16.87
7v	6.155	-0.767	0.927	14.84	1.17	1.22	16.81
7w	6.257	-0.582	0.939	14.46	1.16	1.26	18.33

Tabla 16. Valores de correlación entre los descriptores del modelo 9.

Correlación			
	SPAN	Mor10u	Mor19e
SPAN	1	0.00011269	0.0157608
Mor10u		1	0.00892581
Mor19e			1

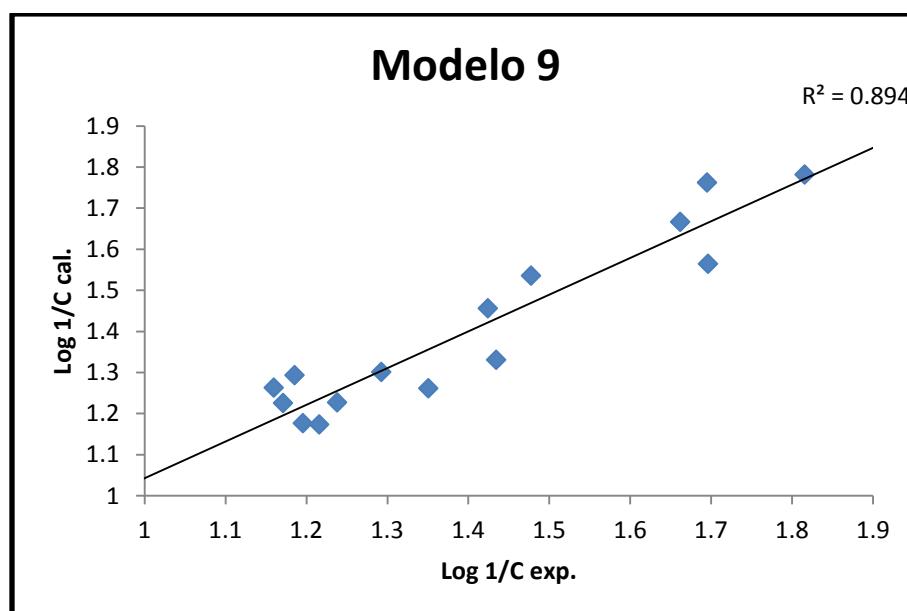


Figura 10. Se representa la relación entre los valores de la actividad experimental y la calculada.

Tabla 17. Valores obtenidos en la estimación lineal para el modelo 9.

Valores de los coeficientes de correlación				
	m2	m1	m	C
	0.82	0.36	-0.38	3.09
	0.12	0.058	0.071	0.45
R²	89	0.08		
F	30.91	11	S	
	0.62	0.073		

Modelo 12:

$$\text{Log } 1/C = -0.366 \text{ SPAN } (\pm 0.075) + 0.368 \text{ Mor10u } (\pm 0.061) + 0.802 \text{ Mor19u } (\pm 0.131) + 2.935 (\pm 0.487)$$

$$n = 15; R^2 = 0.833; F = 27.71; s = 11; Q^2 = 78.52; Kx = 7.72$$

Tabla 18. Valores de los descriptores de los compuestos, así como la actividad calculada para el modelo 12.

Modelo 12	SPAN	Mor10u	Mor19u				
Compuesto	SPAN	Mor10u	Mor19u	Actividad μm	Log 1/C exp	Log 1/C cal.	Act. Cal. μm
7b	6.646	0.181	1.076	26.59	1.42	1.42	26.86
7c	6.929	-0.898	1.374	15.71	1.19	1.16	14.68
7d	5.971	-0.012	1.267	49.58	1.69	1.75	57.37
7f	6.318	-0.223	1.551	65.44	1.81	1.78	60.48
7h	6.362	-0.532	1.391	30.06	1.47	1.52	33.36
7j	6.07	-0.767	1.065	19.62	1.29	1.28	19.15
7l	6.508	-0.144	0.85	16.45	1.21	1.17	15.09
7m	6.728	-0.14	1.139	27.21	1.43	1.33	21.44
7n	6.147	0.091	1.221	45.93	1.66	1.69	49.57
7o	5.732	-0.385	1.064	49.73	1.69	1.54	35.15
7q	6.236	-1.078	1.265	22.43	1.35	1.26	18.50
7r	6.397	-0.599	1.168	15.34	1.18	1.30	20.27
7s	6.51	-0.38	1.045	17.31	1.23	1.24	17.67
7v	6.155	-0.767	1.051	14.84	1.17	1.23	17.36
7w	6.257	-0.582	1.043	14.46	1.16	1.26	18.37

Tabla 19. Valores de correlación entre los descriptores del modelo 12.

Correlación			
	SPAN	Mor10u	Mor19u
SPAN	1	0.00011269	0.00867982
Mor10u		1	0.01356008
Mor19u			1

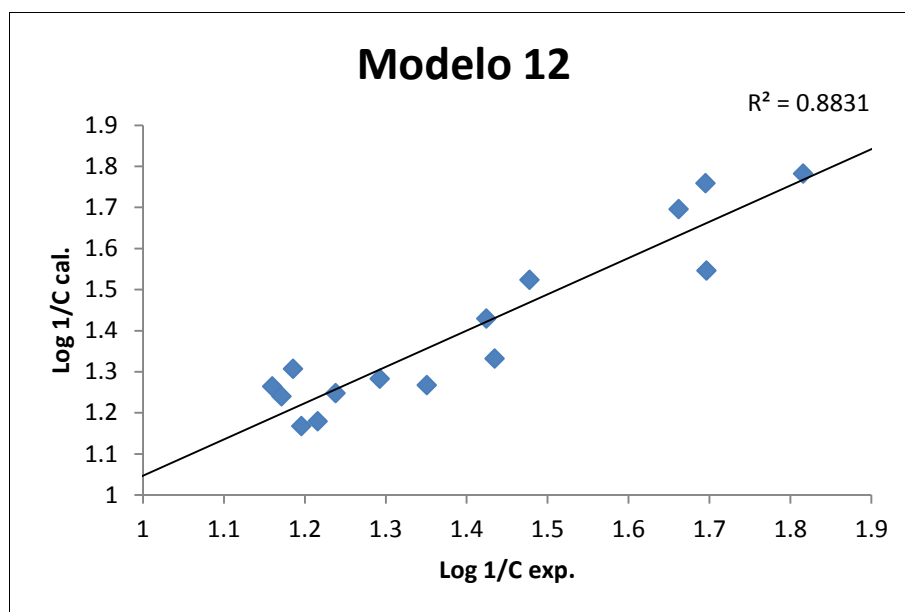


Figura 11. Se representa la relación entre los valores de la actividad calculada y la experimental.

Tabla 20. Valores obtenidos en la estimación lineal para el modelo 12.

Valores de los coeficientes de regresión				
	m2	m1	m	C
	0.80	0.36	-0.36	2.93
	0.13	0.06	0.075	0.48
R²	88	0.085		
F	27.71	11	S	
	0.61	0.08		

V.3.- Obtención de modelos QSAR para las benzacepinas serie 7 a-w citotóxica frente a células vero.

Para el estudio de toxicidad únicamente se emplearon 12 moléculas de la serie de los derivados de amino alcoholes ya que son los que presentaron valores útiles para este estudio. De los 12 compuestos, se reserva una para el proceso de la validación. En la Tabla 21 se resumen los datos de actividad de cada uno de los compuestos utilizados en el estudio.

Tabla 21. Serie de compuestos a utilizar para el QSAR de citotoxicidad.

Compuesto	R	R ₁	R ₂	R ₃	R ₄	CC ₅₀ / Cel. Vero
7b	OCF ₃	H	H	H	H	123.55
7c	Br	H	H	H	H	42.14
7d	Cl	H	H	H	H	190.89
7j	H	H	Cl	H	H	295.36
7l	Cl	Cl	H	H	H	76.6
7m	Cl	Br	H	H	H	57.69
7n	F	Cl	H	H	H	193.8
7q	Me	H	H	Cl	H	82.37
7r	Cl	H	H	Cl	H	106.01
7s	Cl	H	H	F	H	182.39
7v	Cl	H	H	Cl	Cl	49.91
7w	Cl	H	H	Cl	F	92.02
Nifurtimox						61.42

Se llevó a cabo la misma metodología que para el apartado anterior, obteniéndose primeramente los 50 modelos que mejor caben dentro de las restricciones estadísticas impuestas para el estudio en el programa de cómputo. Los resultados se resumen en la Tabla 22.

Tabla 22. Los 50 modelos con sus descriptores y principales parámetros estadísticos.

Modelos	R ²	Q ²	Q ² boot	Kx
MATS8v GATS8v GATS2p	98.57	97.21	47.57	54.02
nCL X2v GATS8m	98.2	96.39	84.45	44.62
RDF060u RDF015m RDF070m	98.12	95.94	86.65	42.5
GATS8v GATS2p RDF070m	98.36	95.7	93.96	33.92
RDF060u RDF070m RDF015e	98.04	95.58	56.84	42.57
MATS8v GATS2m RDF070m	96.9	94.92	0	40.98
ATS8v GATS8m GATS2p	97.25	94.52	71.71	38.85
Ms SP02 RDF060u	98.03	94.3	83.58	47.5
MATS8v GATS1m RDF070m	96.28	93.91	66.75	34.2
GATS2m GATS8v RDF070m	96.79	93.83	89.51	31.9
nAT GATS8m GATS2p	96.63	93.7	80.04	29.38
nBT GATS8m GATS2p	96.63	93.7	85.69	29.38
GATS8m GATS2p RDF070m	96.86	93.43	86.02	37.32
T(N..Cl) X2v GATS8m	96.83	93.38	76.29	42.99
GATS8m GATS2p BELp5	96.61	93.32	0	35.84
GATS8m GATS2p RDF015e	96.34	93.29	71.44	38.71
MATS8v GATS2p RDF070m	97.33	93.22	90.89	42.66
GATS8m GATS2p BELv5	96.63	93.09	0	35.89
GATS8m GATS2p RDF015m	96.15	92.84	0	38.62
GATS8m GATS2p RDF015u	96.25	92.26	0	39.4
GATS8m GATS2v RDF070m	95.9	91.88	81.33	37.6
GATS2p RDF015m RDF070m	95.34	90	79.5	44.92
GATS2p RDF070m RDF015e	95.25	89.73	67.07	45.1
GATS2p RDF015u RDF070m	95.23	89.72	78.22	45.33
ATS3e GATS8m GATS2p	95.7	89.63	33.95	38.53

“Diseño por QSAR de compuestos potencialmente activos contra Trypanosoma cruzi”

GATS2m RDF015m RDF070m	93.84	88.15	65.26	46.48
GATS2m RDF070m RDF015e	93.82	87.99	68.82	46.57
ATS3e GATS2m RDF070m	93.69	87.62	15.24	42.32
GATS2m RDF070m	92.95	87.37	86.98	54.21
GATS8m GATS2p	92.26	87.06	83.92	31.36
GATS2p RDF070m	93.22	86.56	86.26	61.07
GATS2m GATS8m RDF070m	93.72	86.53	59.54	31.88
GATS8m GATS2p RDF060u	92.67	86.32	0	40.64
GATS2v RDF070m	93.1	86.28	85.35	64.19
GATS2m BELp5 RDF070m	93.46	86.26	0	41.18
GATS8m GATS2v BELv5	94.12	86.14	0	33.69
GATS2m RDF070m PCWTe	93.54	86.05	75.69	40.62
GATS8m GATS2v RDF015m	93.84	85.47	56.18	36.65
GATS2m RDF070m Mor15e	93.03	85.4	0	29.65
GATS2m GATS2p RDF070m	94.21	85.28	0	70.67
nAT GATS2m RDF070m	92.97	84.44	77.39	43.91
MATS8v GATS8m GATS2p	92.28	84.34	65.36	34.03
GATS8m GATS2v	92.03	84.2	82.11	27.22
GATS2p RDF070m PCWTe	93.5	84.16	0	42.47
X2v GATS2m RDF070m	92.99	83.84	48.68	62.37
GATS8m GATS2p PCWTe	95.88	83.84	72.67	24.12
X2v GATS8m GATS2p	92.29	83.67	0	62.27
GATS8m RDF015m RDF070m	91.01	83.57	0	15.13
GATS2v	84.5	74	77.13	0
GATS2p	82.39	71.43	73.06	0

Así mismo, se eliminan modelos que podrían contener ambigüedad o dependencia entre “variables independientes” se llevó a cabo para cada modelo mediante la determinación de coeficiente de correlación de Pearson (R^2)

rescatando únicamente los siguientes modelos con valores absolutos menores a 0.5^{36} :

Tabla 23.1 Correlación de los descriptores del modelo 2

Modelo 2			
	nCL	X2	GATS8m
nCL	1	0.00976257	0.03761892
X2		1	0.01184291
GATS8m			1

Tabla 23.2 Correlación de los descriptores del modelo 3

Modelo 3			
	RDF060u	RDF015m	RDF070m
RDF060u	1	0.2959032	0.34542497
RDF015m		1	0.00915868
RDF070m			1

Tabla 23.3 Correlación de los descriptores del modelo 4

Modelo 4			
	GATS8v	GATS2p	RDF070m
GATS8v	1	0.02429046	0.02165301
GATS2p		1	0.37292748
RDF070m			1

Tabla 23.4 Correlación de los descriptores del modelo 6

Modelo 6			
	MATS8v	GATS2m	RDF070m
MATS8v	1	0.26433622	0.01920691
GATS2m		1	0.29386571
RDF070m			1

Tabla 23.5 Correlación de los descriptores del modelo 7

Modelo 7			
	ATS8v	GATS8m	GATS2p
ATS8v	1	0.49673148	0.00022109
GATS8m		1	0.0983323
GATS2p			1

Tabla 23.6 Correlación de los descriptores del modelo 9

Modelo 9			
	MATS8v	GATS1m	RDF070m
MATS8v	1	0.23253008	0.01920691
GATS1m		1	0.14227433
RDF070m			1

Tabla 23.7 Correlación de los descriptores del modelo 10

Modelo 10			
	GATS2m	GATS8v	RDF070m
GATS2m	1	0.04086849	0.29386571
GATS8v		1	0.02165301
RDF070m			1

Tabla 23.8 Correlación de los descriptores del modelo 11

Modelo 11			
	nAT	GATS8m	GATS2p
nAT	1	0.01345609	0.18084274
GATS8m		1	0.0983323
GATS2p			1

Tabla 23.9 Correlación de los descriptores del modelo 12

Modelo 12			
	nBT	GATS8m	GATS2p
nBT	1	0.01345609	0.18084274
GATS8m		1	0.0983323
GATS2p			1

Tabla 23.10 Correlación de los descriptores del modelo 13

Modelo 13			
	GATS8m	GATS2p	RDF070m
GATS8m	1	0.0983323	0.01766125
GATS2p		1	0.37292748
RDF070m			1

Tabla 23.11 Correlación de los descriptores del modelo 14

Modelo 14			
	T(N..CI)	X2v	GATS8m
T(N..CI)	1	0.23869132	0.04068341
X2v		1	0.32853453
GATS8m			1

Tabla 23.13 Correlación de los descriptores del modelo 16

Modelo 16			
	GATS8m	GATS2p	RDF015e
GATS8m	1	0.0983323	0.0453614
GATS2p		1	0.35310993
RDF015e			1

Tabla 23.15 Correlación de los descriptores del modelo 18

Modelo 18			
	GATS8m	GATS2p	BELv5
GATS8m	1	0.0983323	0.0690769
GATS2p		1	0.23633093
BELv5			1

Tabla 23.17 Correlación de los descriptores del modelo 20

Modelo 20			
	GATS8m	GATS2p	RDF015u
GATS8m	1	0.0983323	0.06654793
GATS2p		1	0.33679515
RDF015u			1

Tabla 23.19 Correlación de los descriptores del modelo 22

Modelo 22			
	GATS2p	RDF015m	RDF070m
GATS2p	1	0.34811336	0.37292748
RDF015m		1	0.00915868
RDF070m			1

Tabla 23.24 Correlación de los descriptores del modelo 24

Modelo 24			
	GATS2p	RDF015u	RDF070m
GATS2p	1	0.33679515	0.37292748
RDF015u		1	0.01532462
RDF070m			1

Tabla 23.12 Correlación de los descriptores del modelo 15

Modelo 15			
	GATS8m	GATS2p	BELp5
GATS8m	1	0.0983323	0.06599513
GATS2p		1	0.23958136
BELp5			1

Tabla 23.14 Correlación de los descriptores del modelo 17

Modelo 17			
	MATS8v	GATS2p	RDF070m
MATS8v	1	0.23891367	0.01920691
GATS2p		1	0.37292748
RDF070m			1

Tabla 23.16 Correlación de los descriptores del modelo 19

Modelo 19			
	GATS8m	GATS2p	RDF015m
GATS8m	1	0.0983323	0.04664718
GATS2p		1	0.34811336
RDF015m			1

Tabla 23.18 Correlación de los descriptores del modelo 21

Modelo 21			
	GATS8m	GATS2v	RDF070m
GATS8m	1	0.0741087	0.01766125
GATS2v		1	0.41198213
RDF070m			1

Tabla 23.20 Correlación de los descriptores del modelo 23

Modelo 23			
	GATS2p	RDF070m	RDF015e
GATS2p	1	0.37292748	0.35310993
RDF070m		1	0.00945002
RDF015e			1

Tabla 23.25 Correlación de los descriptores del modelo 25

Modelo 25			
	ATS3e	GATS8m	GATS2p
ATS3e	1	0.13680332	0.21817621
GATS8m		1	0.0983323
GATS2p			1

Tabla 23.26 Correlación de los descriptores del modelo 16

Modelo 26			
	GATS2m	RDF015m	RDF070m
GATS2m	1	0.48361692	0.29386571
RDF015m		1	0.00915868
RDF070m			1

Tabla 23.28 Correlación de los descriptores del modelo 28

Modelo 28			
	ATS3e	GATS2m	RDF070m
ATS3e	1	0.36976588	0.00392908
GATS2m		1	0.29386571
RDF070m			1

Tabla 23.30 Correlación de los descriptores del modelo 30

Modelo 30		
	GATS8m	GATS2p
GATS8m	1	0.0983323
GATS2p		1

Tabla 23.32 Correlación de los descriptores del modelo 32

Modelo 32			
	GATS2m	GATS8m	RDF070m
GATS2m	1	0.04635148	0.29386571
GATS8m		1	0.01766125
RDF070m			1

Tabla 23.34 Correlación de los descriptores del modelo 34

Modelo 34		
	GATS2v	RDF070m
GATS2v	1	0.41198213
RDF070m		1

Tabla 23.36 Correlación de los descriptores del modelo 36

Modelo 36			
	GATS8m	GATS2v	BELv5
GATS8m	1	0.0741087	0.0690769
GATS2v		1	0.21286045
BELv5			1

Tabla 23.27 Correlación de los descriptores del modelo 15

Modelo 27			
	GATS2m	RDF070m	RDF015e
GATS2m	1	0.29386571	0.48532939
RDF070m		1	0.00945002
RDF015e			1

Tabla 23.29 Correlación de los descriptores del modelo 29

Modelo 29		
	GATS2m	RDF070m
GATS2m	1	0.29386571
RDF070m		1

Tabla 23.31 Correlación de los descriptores del modelo 31

Modelo 31		
	GATS2p	RDF070m
GATS2p	1	0.37292748
RDF070m		1

Tabla 23.33 Correlación de los descriptores del modelo 33

Modelo 33			
	GATS8m	GATS2p	RDF060u
GATS8m	1	0.0983323	0.29614232
GATS2p		1	0.12051574
RDF060u			1

Tabla 23.35 Correlación de los descriptores del modelo 35

Modelo 35			
	GATS2m	BELp5	RDF070m
GATS2m	1	0.37846746	0.29386571
BELp5		1	5.111E-05
RDF070m			1

Tabla 23.37 Correlación de los descriptores del modelo 37

Modelo 37			
	GATS2m	RDF070m	PCWTe
GATS2m	1	0.29386571	0.10882965
RDF070m		1	0.11070707
PCWTe			1

Tabla 23.38 Correlación de los descriptores del modelo 38

Modelo 38			
	GATS8m	GATS2v	RDF015m
GATS8m	1	0.0741087	0.04664718
GATS2v		1	0.32519888
RDF015m			1

Tabla 23.40 Correlación de los descriptores del modelo 41

Modelo 41			
	nAT	GATS2m	RDF070m
nAT	1	0.33781578	0.02535251
GATS2m		1	0.29386571
RDF070m			1

Tabla 23.42 Correlación de los descriptores del modelo 43

Modelo 43		
	GATS8m	GATS2v
GATS8m	1	0.0741087
GATS2v		1

Tabla 23.44 Correlación de los descriptores del modelo 46

Modelo 46			
	GATS8m	GATS2p	PCWTe
GATS8m	1	0.0983323	0.00758582
GATS2p		1	0.0923364
PCWTe			1

Tabla 23.39 Correlación de los descriptores del modelo 39

Modelo 39			
	GATS2m	RDF070m	Mor15e
GATS2m	1	0.29386571	7.2091E-05
RDF070m		1	0.05417728
Mor15e			1

Tabla 23.41 Correlación de los descriptores del modelo 42

Modelo 42			
	MATS8v	GATS8m	GATS2p
MATS8v	1	0.03811629	0.23891367
GATS8m		1	0.0983323
GATS2p			1

Tabla 23.43 Correlación de los descriptores del modelo 44

Modelo 44			
	GATS2p	RDF070m	PCWTe
GATS2p	1	0.37292748	0.0923364
RDF070m		1	0.11070707
PCWTe			1

Tabla 23.45 Correlación de los descriptores del modelo 48

Modelo 48			
	GATS8m	RDF015m	RDF070m
GATS8m	1	0.04664718	0.00915868
RDF015m		1	0.00915868
RDF070m			1

V.4.- Modelos seleccionados para el análisis de Toxicidad:

Se analizaron los datos de los 44 modelos que no presentaron correlación entre sus descriptores y se hizo a cada uno la estimación lineal para obtener los valores para la R^2 , F de Fisher, los valores de los coeficientes de regresión, la constante de la ecuación, los valores de 95% de confianza, constante de la ecuación, así como la desviación estándar. Se obtuvieron además los valores de toxicidad calculada por cada ecuación. Después del análisis se llegó a los siguientes modelos como los “mejores”:

“Diseño por QSAR de compuestos potencialmente activos contra Trypanosoma cruzi”

Modelo 3 para determinación de toxicidad:

$$\text{Log } 1/C = -0.077 \text{ RDF060u } (\pm 0.012) + 0.287 \text{ RDF70} (\pm 0.026) - 0.142 \text{ RDF070m} (\pm 0.009) + 0.352 (\pm 0.256)$$

$$n = 11; R^2 = 98.1; F = 121.62; s = 7; Q^2 = 95.94; Kx = 42.5$$

Tabla 24. Valores de los descriptores de los compuestos, se muestra la citotoxicidad calculada y experimental.

Modelo 3	RDF060u	RDF015m	RDF070m				
Compuesto	RDF060u	RDF015m	RDF070m	Toxicidad μm	Log 1/C exp	Log 1/C cal.	Toxi.Cal. μm
mol 7b	13.991	13.043	6.872	123.55	2.09	2.04	111.76
mol 7c	10.531	11.292	8.177	42.14	1.62	1.62	42.27
mol 7d	11.998	11.442	3.163	190.89	2.28	2.26	185.55
mol 7j	15.735	13.186	3.055	295.36	2.47	2.49	314.25
mol 7l	14.374	11.43	4.417	76.6	1.88	1.90	80.15
mol 7m	14.146	11.311	5.454	57.69	1.76	1.73	54.94
mol 7n	14.594	11.813	2.721	193.8	2.28	2.23	172.96
mol 7q	12.575	12.353	6.956	82.37	1.915	1.94	88.53
mol 7r	12.792	11.636	4.274	106.01	2.02	2.10	127.42
mol 7s	11.065	11.618	4.289	182.39	2.26	2.23	170.20
mol 7u	10.092	11.461	8.27	49.91	1.69	1.69	49.57

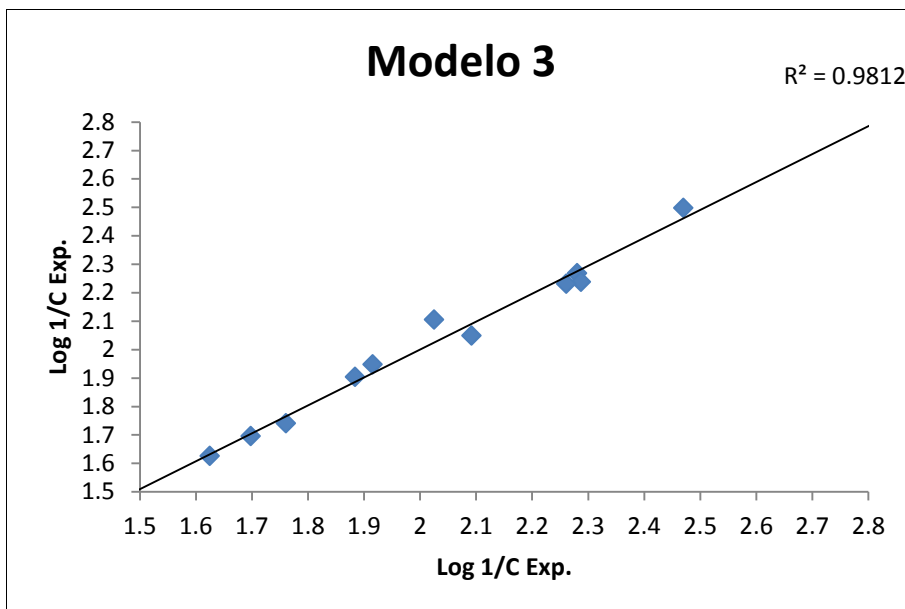


Figura 12. . Se representa la relación entre los valores de la citotoxicidad experimental y calculada.

Tabla 25. Valores de correlación entre los descriptores del modelo 3.

Correlación			
	RDF060u	RDF015m	RDF070m
RDF060u	1	0.2959032	0.34542497
RDF015m		1	0.00915868
RDF070m			1

Tabla 26. Valores obtenidos en la estimación lineal para el modelo 3.

Valores de los coeficientes de regresión				
	m2	m1	m	constante
	-0.14	0.28	-0.07	0.35
	0.009	0.026	0.012	0.256
R ²	98.1	0.045		
F	121.62	7		
	0.746	0.014		

Modelo 4:

$$\text{Log } 1/C = 0.698 \text{ GATS8v } (\pm 0.149) + 2.258 \text{ GATS2p } (\pm 0.219) - 0.0537 \text{ RDF070m } - 0.512 (\pm 0.008)$$

$$n = 11; R^2 = 98.3; F = 140.31; s = 7; Q^2 = 95.7; Kx = 33.92$$

Tabla 27. Valores de los descriptores de los compuestos, se describe la citotoxicidad calculada y la experimental obtenida a partir del modelo 4.

Modelo 4	GATS8v	GATS2p	RDF070m	Toxicidad μm	Log 1/C exp	Log 1/C cal	Toxic.Cal. μm
mol 7b	1.003	1.009	6.872	123.55	2.091	2.097	125.04
mol 7c	0.868	0.897	8.177	42.14	1.624	1.679	47.83
mol 7d	0.923	1.025	3.163	190.89	2.280	2.276	189.10
mol 7j	1.123	1.05	3.055	295.36	2.470	2.478	301.05
mol 7l	0.841	0.928	4.417	76.6	1.884	1.932	85.70
mol 7m	0.933	0.827	5.454	57.69	1.761	1.713	51.69
mol 7n	0.928	1.025	2.721	193.8	2.287	2.303	201.34
mol 7q	0.814	0.968	6.956	82.37	1.915	1.867	73.78
mol 7r	0.982	0.928	4.274	106.01	2.025	2.039	109.43
mol 7s	0.925	1.025	4.289	182.39	2.261	2.217	165.02
mol 7u	1.05	0.849	8.27	49.91	1.698	1.693	49.37

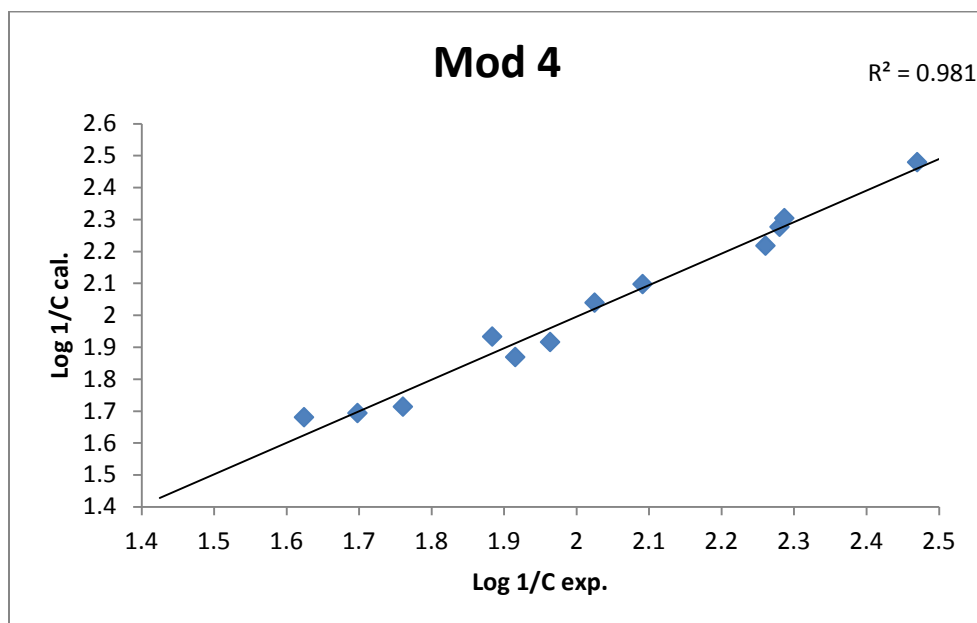


Figura 13. Representación de la relación entre los valores de la citotoxicidad experimental y la calculada.

Tabla 28. Valores de correlación entre los descriptores del modelo 4.

Correlación			
	GATS8v	GATS2p	RDF070m
GATS8v	1	0.02429046	0.02165301
GATS2p		1	0.37292748
RDF070m			1

Tabla 29. Valores obtenidos en la estimación lineal para el modelo 4.

Valores de los coeficientes de regresión				
	m2	m1	m	constante
	-0.053	2.257	0.698	-0.512
	0.008	0.219	0.148	0.269
R²	98.3	0.042		
F	140.311	7	s	
	0.748	0.012		

Modelo 10:

$$\text{Log1/C} = 2.167 \text{ GATS2m } (\pm 0.305) + 0.607 \text{ GATS8v } (\pm 0.210) - 0.0639 \text{ RDF070m } (\pm 0.0109) + 0.348 (\pm 0.292)$$

$$n = 11; R^2 = 96.7; F = 70.40; s = 7; Q^2 = 93.83; Kx = 31.9$$

Tabla 30. Valores de los descriptores de los compuestos, se muestra la citotoxicidad calculada y experimental obtenida a partir del modelo 10.

Modelo 10	GATS2m	GATS8v	RDF070m	Toxicidad μm	Log 1/C exp	Log 1/C cal.	Toxic.cal. μm
mol 7b	0.721	1.003	6.872	123.55	2.091	2.081	120.63
mol 7c	0.568	0.868	8.177	42.14	1.624	1.584	38.39
mol 7d	0.743	0.923	3.163	190.89	2.280	2.317	207.88
mol 7j	0.763	1.123	3.055	295.36	2.470	2.489	308.77
mol 7l	0.634	0.841	4.417	76.6	1.884	1.951	89.42
mol 7m	0.552	0.933	5.454	57.69	1.761	1.763	57.98
mol 7n	0.697	0.928	2.721	193.8	2.287	2.249	177.59
mol 7q	0.715	0.814	6.956	82.37	1.915	1.948	88.75
mol 7r	0.634	0.982	4.274	106.01	2.025	2.046	111.25
mol 7s	0.697	0.925	4.289	182.39	2.261	2.147	140.37
mol 7u	0.583	1.05	8.27	49.91	1.698	1.721	52.65

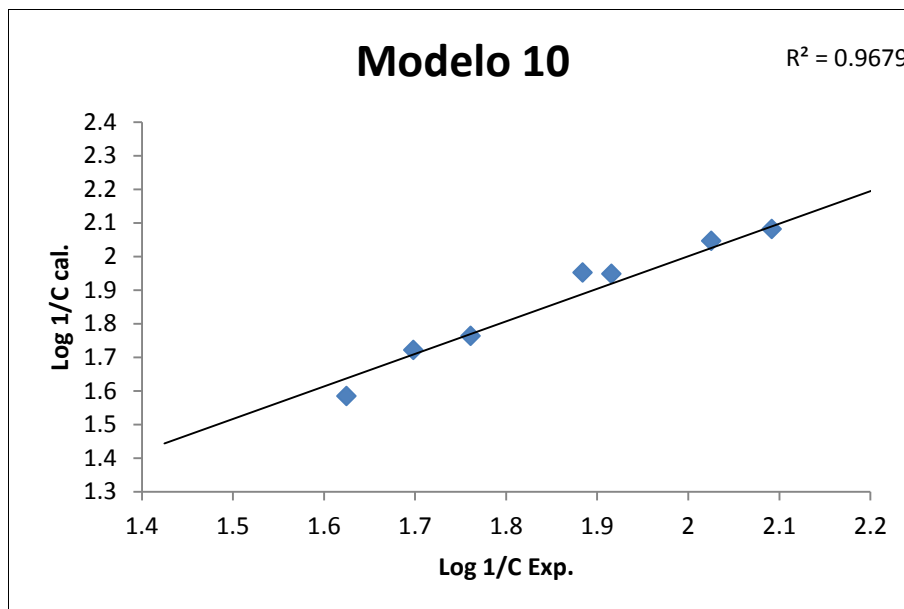


Figura 14. . Representación de la relación entre los valores de la citotoxicidad experimental y la calculada.

Tabla 31. Valores de correlación entre los descriptores del modelo 10.

Correlación			
	GATS2m	GATS8v	RDF070m
GATS2m	1	0.04086849	0.29386571
GATS8v		1	0.02165301
RDF070m			1

Tabla 32. Valores obtenidos en la estimación lineal para el modelo 10.

Valores de los coeficientes de regresión				
	m2	m1	m	constante
	-0.063	0.607	2.167	0.348
	0.010	0.209	0.305	0.292
R²	96.7	0.059		
F	70.402	7	s	
	0.736	0.024		

Modelo 17:

$$\text{Log1/C} = -0.598 \text{ MATS8v } (\pm 0.182) + 1.792 \text{ GATS2p } (\pm 0.32619079) - 0.064 \text{ RDF070m } (\pm 0.010) + 0.706 (\pm 0.360)$$

$$n = 11; R^2 = 97.3; F = 85.14; s = 7; Q^2 = 93.22; Kx = 42.66$$

Tabla 33. Valores de los descriptores de los compuestos, se muestra la citotoxicidad calculada y experimental obtenida a partir del modelo 17.

Modelo 17 MATS8v GATS2p RDF070m							
Compuesto	MATS8v	GATS2p	RDF070m	Toxicidad μm	Log1/C exp	Log1/C cal.	Toxic.Cal. μm
mol 7b	-0.054	1.009	6.872	123.55	2.091	2.102	126.68
mol 7c	0.164	0.897	8.177	42.14	1.624	1.687	48.67
mol 7d	0.105	1.025	3.163	190.89	2.280	2.275	188.77
mol 7j	-0.121	1.05	3.055	295.36	2.470	2.462	290.32
mol 7l	0.207	0.928	4.417	76.6	1.884	1.960	91.21
mol 7m	0.23	0.827	5.454	57.69	1.761	1.698	49.92
mol 7n	0.102	1.025	2.721	193.8	2.287	2.306	202.44
mol 7q	0.201	0.968	6.956	82.37	1.915	1.871	74.35
mol 7r	0.07	0.928	4.274	106.01	2.025	2.051	112.52
mol 7s	0.102	1.025	4.289	182.39	2.261	2.204	160.32
mol 7u	0.023	0.849	8.27	49.91	1.698	1.679	47.81

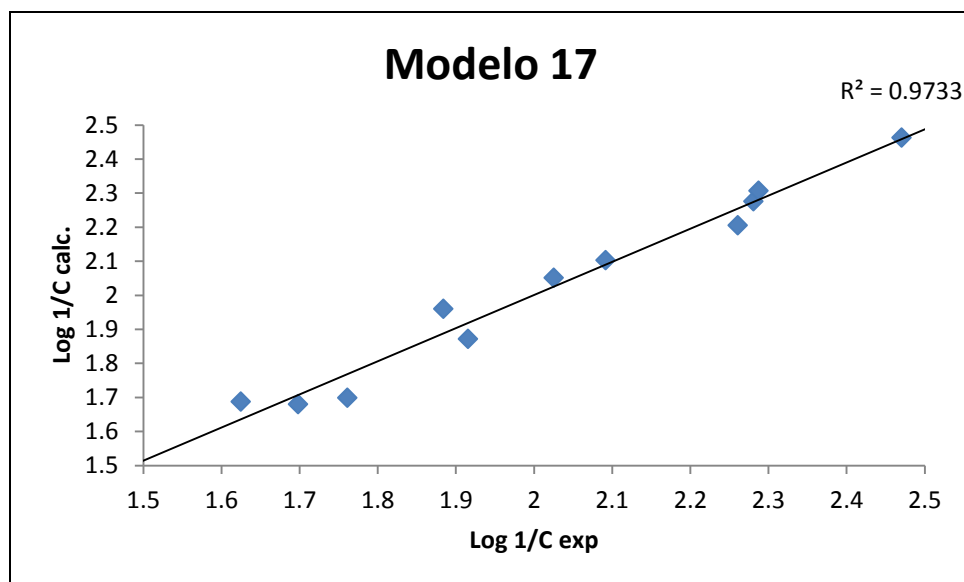


Figura 15. Representación de la relación entre los valores de la citotoxicidad experimental y la calculada.

Tabla 34. Valores de correlación entre los descriptores del modelo 17.

correlación			
	MATS8v	GATS2p	RDF070m
MATS8v	1	0.23891367	0.01920691
GATS2p		1	0.37292748
RDF070m			1

Tabla 35. Valores obtenidos en la estimación lineal para el modelo 17.

Valores de los coeficientes de regresión				
	m2	m1	m	constante
	-0.064	1.792	-0.598	0.705
	0.010	0.326	0.181	0.359
R²	97.3	0.053		
F	85.142	7	s	
	0.740	0.020		

VI.- DISCUSIÓN DE RESULTADOS

VI.1 QSAR de compuestos activos contra Trypanosoma cruzi

Los resultados de los mejores modelos obtenidos se resumen en la Tabla 36, donde se aprecia que todos ellos poseen valores de descriptores estadísticos significativos, considerando que para un estudio QSAR se toman como aceptable una R^2 superior a 60. Así mismo, los valores de K_x indican que existe baja correlación entre descriptores lo cual fue corroborado por el análisis de regresión entre variables independientes.

Por su parte, los valores de Q^2 boot indican una alta capacidad predictiva del modelo. En este contexto cabe resaltar que Q^2 es un tipo de validación interna que se fundamenta en extraer un dato de la serie explorada y realizar su predicción de actividad mediante un modelo obtenido con el resto de la serie y con los descriptores del modelo. Valores cercanos a 100 significan alta capacidad predictiva del modelo. Así, aunque Q^2 es un buen indicador de capacidad predictiva, en este estudio se consideró como restricción valores de Q^2 Boot, considerablemente más confiable, con la finalidad de obtener modelos de forma más estricta. Q^2 Bootstrap (Q^2 Boot), técnica bootstrap debe ser considerada como un plan de remuestreo, posibilitando estudiar el error estadístico, ya sea en cuanto un sesgo, error estándar o tasa de error en una predicción³⁸. Esta dada por la siguiente ecuación²³:

$$PRESS = \sum_{i=1}^n (y_i - \widehat{y}_{i/i})^2$$

PRESS es la suma de las desviaciones cuadradas entre la actividad prevista y valores reales para cada molécula en el conjunto de prueba.³⁹

Tabla 36. Resumen de resultados de los mejores modelos encontrados en el análisis QSAR de benzacepinas activas contra *Trypanosoma cruzi*.

Número de modelo	Descriptores	R ²	Q ² boot	Kx	R ² Estimación lineal
Modelo 3	Mor21v Mor10e Mor24e	90.28	76.5	36.52	0.90
Modelo 5	Mor10u Mor24e Mor21p	89.81	77.9	38.15	0.89
Modelo 6	Mor10e Mor24e Mor21p	91.03	79.38	38.44	0.91
Modelo 9	SPAN Mor10u Mor19e	89.4	77.83	8.12	0.89
Modelo 12	SPAN Mor10u Mor19u	88.31	75.12	7.72	0.88

QSAR se basa en un principio conocido como Navaja de Ockham que, de manera general, indica que cuando existen varias posibles respuestas en igualdad de condiciones, la respuesta más sencilla suele ser la correcta. De este principio de parsimonia, se deriva que es preferible considerar ecuaciones con el menor número posible de descriptores. En la Tabla 36, se aprecia además que los mejores modelos contienen tres descriptores y aunque no se logran encontrar mejores ecuaciones con dos o un descriptor, estas son bastante aceptables dada su validación estadística. Así mismo, es notoria, de manera global, la participación de los descriptores Mor.

Dado que los modelos obtenidos son semejantes, en el sentido estadístico, se decidió llevar a cabo una restricción más que fue la validación externa. Cabe recordar que un modelo QSAR no es más que un modelo matemático sustentado estadísticamente y que su objetivo final es que sea funcional, para lo cual la validación externa determinara esta característica. Esto es que si el modelo pretende ser predictivo debe entonces tener la capacidad de, predecir actividad de compuestos no incluidos en el sistema. Idealmente, no se debe conocer la

actividad del (o de los) compuesto(s) externo(s) hasta su comparación una vez obtenida su actividad teórica de acuerdo al modelo elegido. La actividad teórica debe aproximarse a la experimental aunque no necesariamente ser la misma.

Las Tablas 37, 38, 39, 40 y 41 muestran los resultados de la predicción de los compuestos que fueron excluidos del estudio y que se someten a análisis con uno de los modelos seleccionados.

Tabla 37. Valores de actividad calculada y experimental de los compuestos 7a-t utilizando el modelo 3.

Modelo 3					
Compuesto	Mor21v	Mor10e	Mor24e	Actividad Exp.	Actividad cal.
mol 7a	-0.529	-0.579	0.362	31.82	26.94
mol 7e	-0.507	-0.46	0.281	12.37	21.60
mol 7g	-0.578	-0.125	0.167	26.37	32.28
mol 7p	-0.505	-0.172	0.356	62.97	32.81
mol 7t	-0.516	-0.802	0.522	48.88	30.71

Tabla 38. Valores de actividad calculada y experimental de los compuestos 7a-t utilizando el modelo 5.

Modelo 5					
Compuesto	Mor10u	Mor24e	Mor21p	Actividad experimental	Actividad cal.
mol 7a	-0.635	0.362	-0.588	31.82	29.38
mol 7e	-0.521	0.281	-0.543	12.37	20.84
mol 7g	-0.237	0.167	-0.626	26.37	31.74
mol 7p	-0.336	0.356	-0.55	62.97	31.58
mol 7t	-0.804	0.522	-0.567	48.88	33.95

Tabla 39. Valores de actividad calculada y experimental de los compuestos 7a-t utilizando el modelo 6.

Modelo 6					
Compuesto	Mor10e	Mor24e	Mor21p	Actividad	Actividad cal.
mol 7a	-0.579	0.362	-0.588	31.82	29.4050693
mol 7e	-0.46	0.281	-0.543	12.37	18.5224099
mol 7g	-0.125	0.167	-0.626	26.37	40.9138468
mol 7p	-0.172	0.356	-0.55	62.97	16.2824873
mol 7t	-0.802	0.522	-0.567	48.88	13.6317932

Tabla 40. Valores de actividad calculada y experimental de los compuestos 7a-t utilizando el modelo 9.

Modelo 9					
Compuesto	SPAN	Mor10u	Mor19e	Actividad experimental	Actividad calculada
mol 7a	5.916	-0.635	1.294	31.82	28.60
mol 7e	6.217	-0.521	1.127	12.37	15.01
mol 7g	5.78	-0.237	1.002	26.37	57.83
mol 7p	5.746	-0.336	1.091	62.97	19.99
mol 7t	5.958	-0.804	0.989	48.88	14.91

Tabla 41. Valores de actividad calculada y experimental de los compuestos 7a-t utilizando el modelo 12.

Modelo 12					
Compuesto	SPAN	Mor10u	Mor19u	Actividad experimental	Actividad calculada
mol 7a	5.916	-0.635	1.401	31.82	26.8669885
mol 7e	6.217	-0.521	1.278	12.37	14.6872709
mol 7g	5.78	-0.237	1.118	26.37	57.376627
mol 7p	5.746	-0.336	1.196	62.97	19.1504551
mol 7t	5.958	-0.804	1.116	48.88	15.0915516

Es apreciable que el compuesto mol 7p es el que se mantiene más alejado de su actividad calculada, independientemente del modelo explorado. Esto puede deberse a varias razones entre las que se encuentran, como sobresalientes, que el modelo no sea el adecuado, que las características del compuesto no compartan mecanismo de acción con el resto de la serie o que la actividad experimental no haya sido medida correctamente por cuestión de error experimental o baja solubilidad entre otros. En estos casos resulta preferible considerarlo como “outlier” es decir un punto que dista del resto de los datos en un análisis estadístico.^{40,41} En este sentido es importante ser sumamente cuidadosos y no caer en vicios o malas prácticas estadísticas analizando con rigor la decisión de considerar un dato como “outlier”. En este caso, si se decide eliminar el dato del compuesto mol7p, la estadística mejora en la mayoría de los datos de predicción; sin embargo, químicamente la actividad experimental tiene sentido o explicación y no justifica actividad anómala cuando se compara con sus análogos, lo que hace descartar un probable error de medición experimental y si se incluyeron en el análisis.

La comparación de los modelos se llevó a cabo por dos vías. La primera por la diferencia de la sumatoria de las actividades experimentales; restando las calculadas para cada uno de los modelos y la segunda por la diferencia de R^2 del comportamiento de los datos experimentales con respecto a la R^2 del comportamiento de los datos calculados considerando todos los compuestos. Los resultados se muestran en la Tablas 42 y Figura 17 respectivamente.

Tabla 42. Residuos de la diferencia entre la sumatoria de la actividad experimental y la calculada de los mejores modelos.

Modelo	Actividad Experimental	Actividad calculada	Residuos
3	204.15	223.47	19.32
5	204.15	229.34	25.19
6	204.15	211.79	7.64
9	204.15	231.16	27.0162138
12	204.15	227.02	22.86

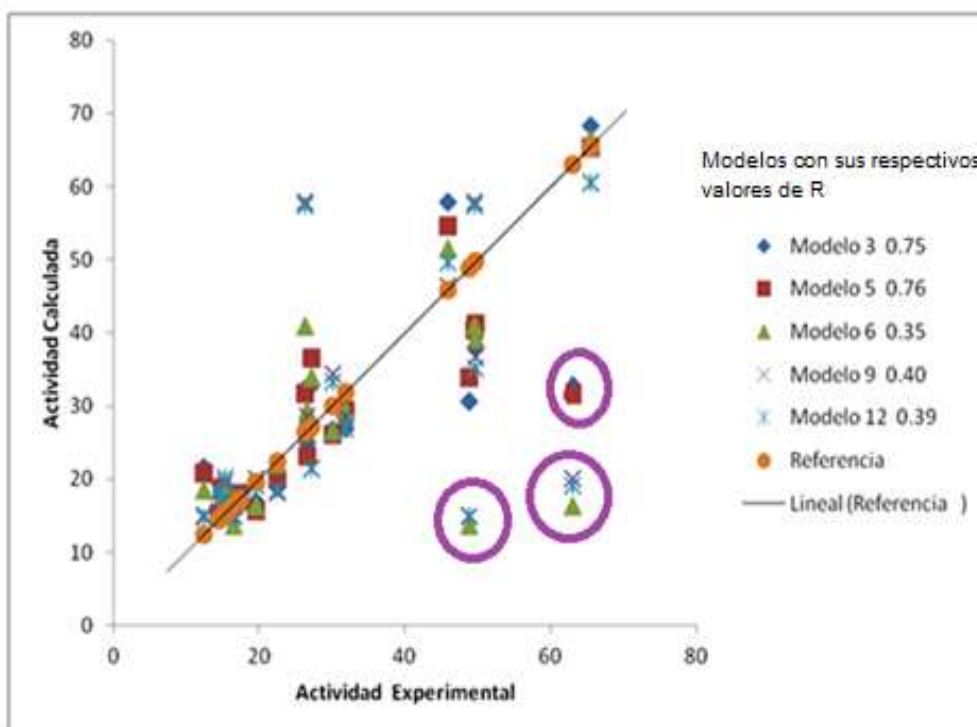
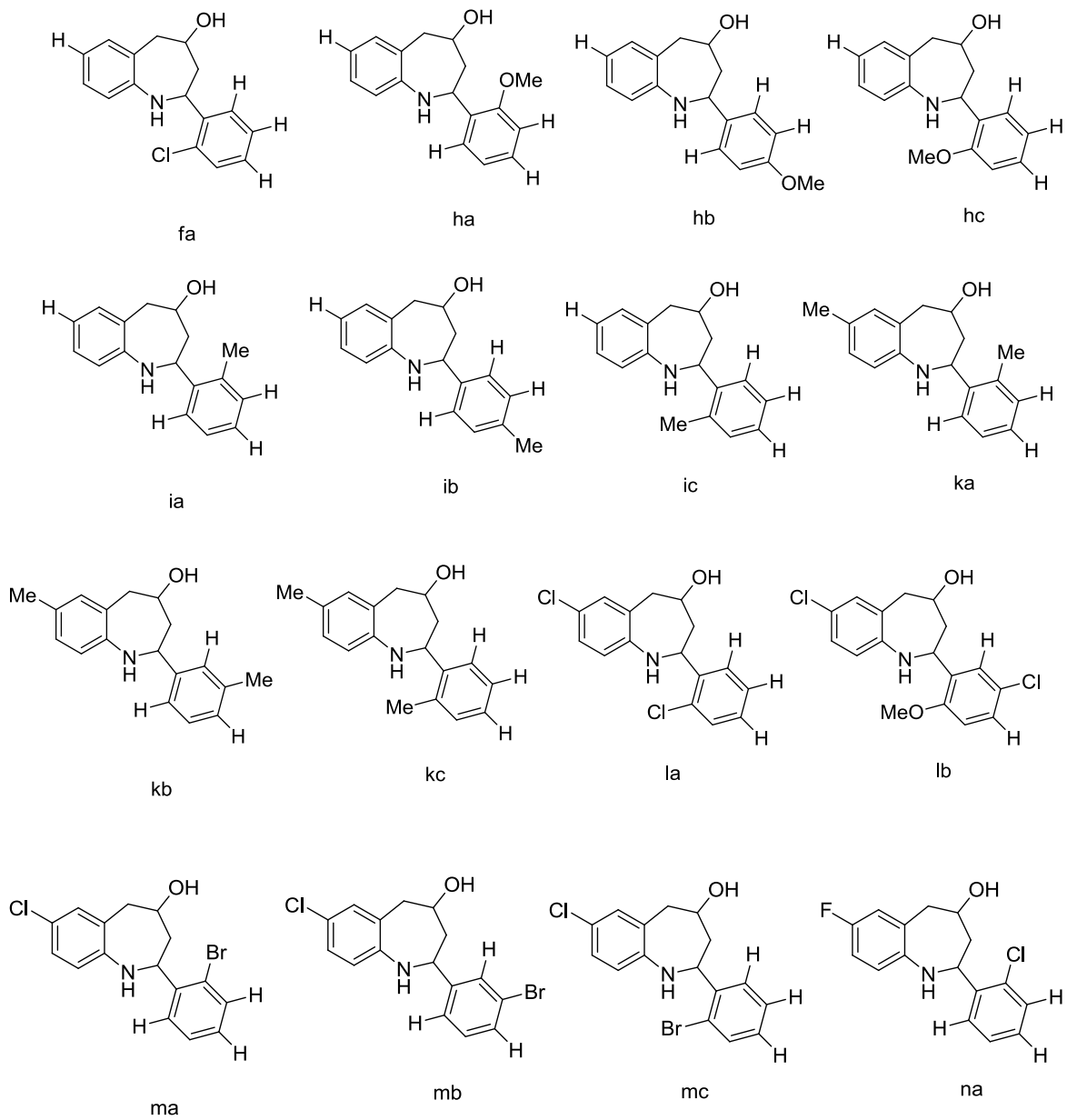


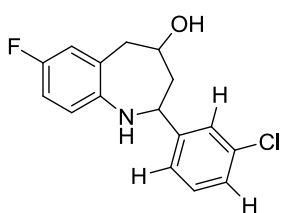
Figura 16. Actividad experimental contra observada de los 5 modelos.

El parecido de los modelos seleccionados en su estadística y en su composición de variables independientes es similar aunque en lo que refiere a su capacidad de predicción, de acuerdo a los resultados mostrados en la Figura 18, los modelos 3 y 5 son de alta capacidad predictiva y distan demasiado de los modelos 6, 9 y 12, esto se observa con los valores de R^2 mostrados. Aunque estos últimos puede ser de índole explicativa, la capacidad predictiva permite tamizar aun más el procedimiento, obteniendo dos mejores modelos. Los modelos 3 y 5 pueden ser confiables en predecir actividad en compuestos desconocidos y mantienen un claro parentesco en sus (descriptores Mor). Este resultado es interesante pues se puede anticipar que la actividad de este tipo de compuestos estará relacionada a esta familia de descriptores.

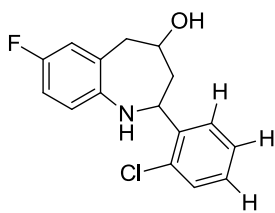
VI.2.-Propuesta de nuevos análogos potencialmente activos contra Trypanosoma cruzi.

Como uno de los objetivos de este estudio es la propuesta de estructuras novedosas de tipo cis-2-aryl-4-hidroxitetrahydro-1-benzacepinas, que de acuerdo a sus características estructurales y fisicoquímicas obtenidas a partir de descriptores teóricos, puedan presentar actividad. Se llevó a cabo una propuesta inicial de 39 compuestos (figura 17) derivados de cis-2-aryl-4-hidroxitetrahydro-1-benzacepinas con variaciones en los sustituyentes en el grupo fenilo. La predicción de la actividad se llevó a cabo con los modelos 3 y 5 y los resultados se resumen en la Tabla 43.

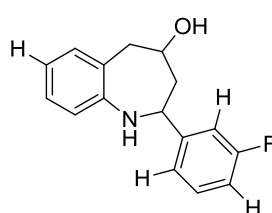




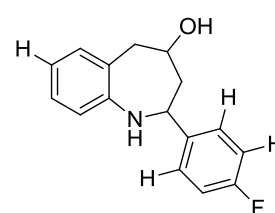
nb



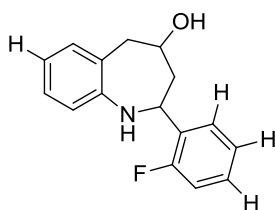
nc



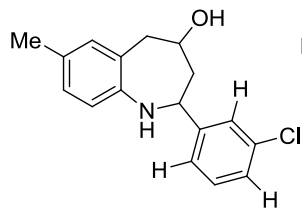
pa



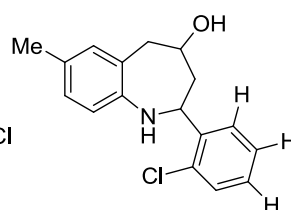
pb



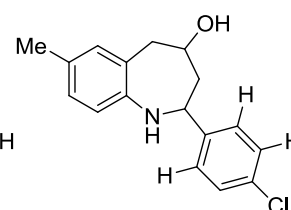
pc



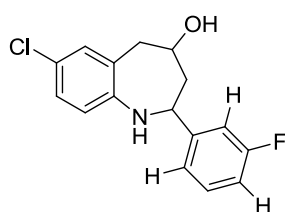
qa



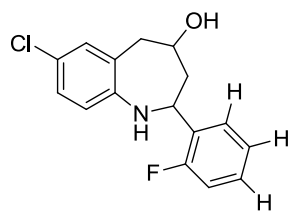
qb



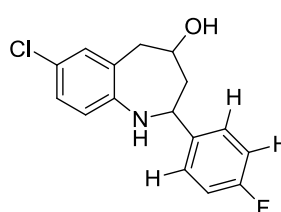
qc



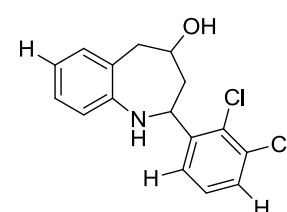
sa



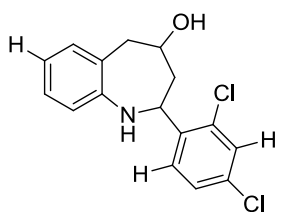
sb



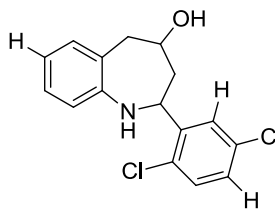
sc



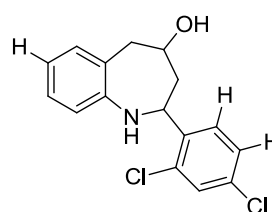
ta



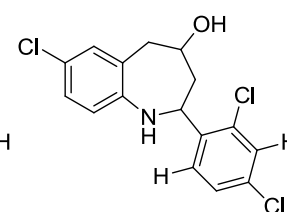
tb



tc



td



va

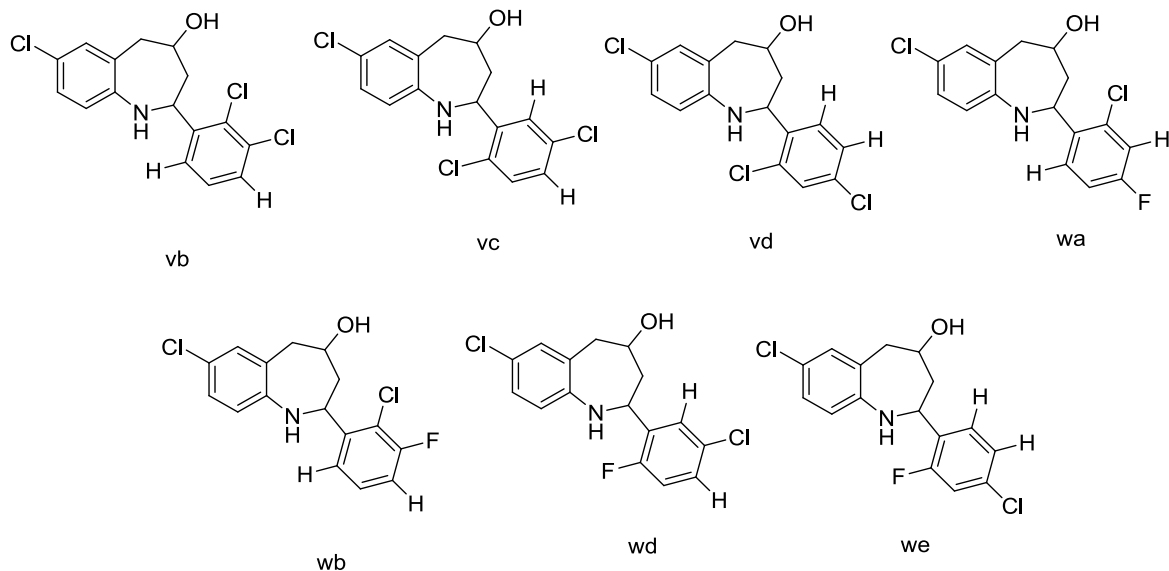


Figura 17. Serie propuesta de derivados de cis-2-aril-4-hidroxitetrahydro-1-benzacepinas.

Tabla 43. Resultados de actividad calculada por los modelos 3 y 5 para la serie de compuestos propuesto.

Compuesto	R	R ₁	R ₂	R ₃	R ₄	Modelo 3	Modelo 5
7 a	H	H	H	H	H	26.94	29.38
7b	OCF ₃	H	H	H	H	25.11	23.10
7c	Br	H	H	H	H	17.65	17.77
7d	Cl	H	H	H	H	39.93	40.38
7e	F	H	H	H	H	21.60	20.84
7f	H	Cl	H	H	H	68.40	65.31
7fa	H	H	H	H	Cl	18.72	17.46
7g	H	F	H	H	H	32.28	31.74
7h	H	H	OMe	H	H	26.81	26.14
7ha	H	H	H	OMe	H	46.51	46.22
7hb	H	OMe	H	H	H	30.22	28.66
7hc	H	H	H	H	OMe	43.01	43.49
7i	H	H	Me	H	H	25.29	26.82
7ia	H	H	H	Me	H	45.60	47.16

7ib	H	Me	H	H	H	23.73	22.17
7ic	H	H	H	H	Me	34.64	35.64
7j	H	H	Cl	H	H	16.93	15.75
7k	Me	Me	H	H	H	43.07	45.43
7ka	Me	H	H	Me	H	46.17	50.16
7kb	Me	H	Me	H	H	30.85	31.58
7kc	Me	H	H	H	Me	59.75	56.42
7l	Cl	Cl	H	H	H	15.51	16.48
7la	Cl	H	H	H	Cl	13.66	13.27
7lb	Cl	H	Cl	H	H	14.24	13.48
7m	Cl	Br	H	H	H	32.82	36.53
7ma	Cl	H	H	Br	H	10.21	8.18
7mb	Cl	H	Br	H	H	9.44	9.80
7mc	Cl	H	H	H	Br	9.54	8.42
7n	F	Cl	H	H	H	57.86	54.65
7na	F	H	H	Cl	H	28.05	27.16
7nb	F	H	Cl	H	H	14.33	13.22
7nc	F	H	H	H	Cl	30.98	31.19
7o	H	H	H	Cl	H	37.99	41.21
7p	H	H	H	F	H	32.81	31.58
7pa	H	H	F	H	H	25.29	25.04
7pb	H	F	H	H	H	32.70	31.22
7pc	H	H	H	H	F	33.21	33.11
7q	Me	H	H	Cl	H	19.01	20.05
7qa	Me	H	Cl	H	H	23.60	25.27
7qb	Me	H	H	H	Cl	37.35	33.83
7qc	Me	Cl	H	H	H	60.40	61.05
7r	Cl	H	H	Cl	H	17.36	16.45
7s	Cl	H	H	F	H	18.08	18.02
7sa	Cl	H	F	H	H	25.90	27.17
7sb	Cl	H	H	H	F	13.88	14.91
7sc	Cl	F	H	H	H	25.12	23.99
7t	H	H	H	Cl	Cl	30.71	33.95
7ta	H	H	Cl	Cl	H	43.96	49.24
7tb	H	Cl	H	Cl	H	32.69	35.32
7tc	H	H	Cl	H	Cl	23.62	18.45
7td	H	Cl	H	H	Cl	33.28	36.73
7v	Cl	H	H	Cl	Cl	28.30	18.78
7va	Cl	Cl	H	Cl	H	56.84	61.99
7vb	Cl	H	Cl	Cl	H	20.29	23.94
7vc	Cl	H	Cl	H	Cl	8.05	7.21

7vd	Cl	Cl	H	H	Cl	9.66	10.32
7w	Cl	H	H	Cl	F	15.48	15.42
7wa	Cl	F	H	Cl	H	50.53	55.84
7wb	Cl	H	F	Cl	H	54.09	54.38
7wd	Cl	H	Cl	H	F	9.51	10.80
7we	Cl	Cl	H	H	F	23.35	20.44

VI.3.- Descriptores del modelo para predicción de actividad:

Los modelos 3 y 5 que fueron seleccionados como los “mejores modelos” consideran los descriptores Mor21v, Mor10e, Mor24e, Mor10u Mor24e y Mor21p en sus ecuaciones. Estos pertenecen a la Familia 3D MoRSE , integrada a su vez por 160 descriptores. Los descriptores 3D Morse (Representación de la molécula 3D de estructuras basadas en la difracción de electrones) se derivan de simulaciones de espectros de infrarrojo utilizando una función de dispersión generalizada (Soltzberg & Wilkins, 1977). Un típico descriptor Morse es denotado por Morsw donde s y w toman los valores $1 \leq s \leq 32$ y w (u, m, v, e y p)⁴² .

Los descriptores 3D Morse se calculan para 5 diferentes propiedades atómicas: caso no ponderado (u), masa atómica (m), volumen de Van der Waals (v), la electronegatividad atómica de Sanderson (e) y la polarizabilidad atómica (p).

Los 3D-MoRSE (3D-Molecule Representation of Structures based on Electron diffraction); convierten la estructura tridimensional a un número de variables fijas, el método usa una versión modificada de la transformación molecular usando estudios en difracción de electrón. Varias propiedades atómicas tales como masa atómica, cargas atómicas parciales, residual de electronegatividades atómicas y polarizabilidades atómicas que pueden ser codificado a un set de valores 3D- MoRSE que son independientes. La siguiente expresión se usa para calcular los descriptores³⁰.

$$Morsw = \sum_{i=1}^{NAT-1} \sum_{j=1+1}^{NAT} w_i \cdot w_j \cdot \frac{\sin(s - r_{ji})}{s - r_j}$$

Donde:

Morsw: es la intensidad del electrón dispersado

w: propiedad atómica

r_{ij} : son las distancias interatómicas entre los átomos i,j

NAT: es el número de átomos

S: es una medida del ángulo de dispersión dado por $s = 4\pi \sin(\theta/2) / \lambda$, θ es el ángulo de dispersión y λ es la longitud de onda.

Electronegatividad atómica de Sanderson.

R. T. Sanderson propone que la electronegatividad de un átomo es una medida de la “densidad” de su nube electrónica, comparada con la de un átomo inerte hipotético con el mismo número de electrones, ya que si un átomo es muy electronegativo, es decir, atrae fuertemente a otros electrones, los suyos propios los mantendrá muy próximos entre sí. Sanderson define la densidad electrónica media (DE) mediante la relación:

$$DE = \frac{3Z}{4\pi r^2}$$

Donde r es el radio covalente en angstrom⁴³.

De acuerdo a los modelos seleccionados los descriptores son:

Mor21v: 3D-MoRSE – s 21/ ponderado por volumen atómico de van der Waals.

Mor10e: 3D-MoRSE – s 10 / ponderado por electronegatividad atómica Sanderson.

Mor 24e 3D-MoRSE - s 24 / ponderado por electronegatividad atómica de Sanderson.

Mor21p: 3D-MoRSE - s 21 / ponderado por polarizabilidad atómica.

Mor10u: 3D-MoRSE - s 10 u / sin ponderar.

Que basan sus medidas en la electronegatividad de Sanderson, que se refiere al tamaño y carga nuclear, mostrándonos que si un átomo atrae fuertemente su propia atmosfera electrónica también atraerá con fuerza un electrón de otro átomo si posee un orbital vacante para acomodar ese electrón. Si por otra parte las repulsiones interelectrónicas son suficientes como para que disminuya la atracción de un átomo sobre sus propios electrones, tal átomo no atraerá fuertemente a un electrón externo aunque disponga de orbitales vacíos⁴⁴. En cuanto a la polarizabilidad, habla de la respuesta de la molécula a un campo externo.

Para explicar el papel de cada uno de los descriptores del modelo se toman tres compuestos, una de la serie original y 2 de la serie propuesta con los mejores valores de actividad.

Siendo su estructura la mostrada en la Figura 18.

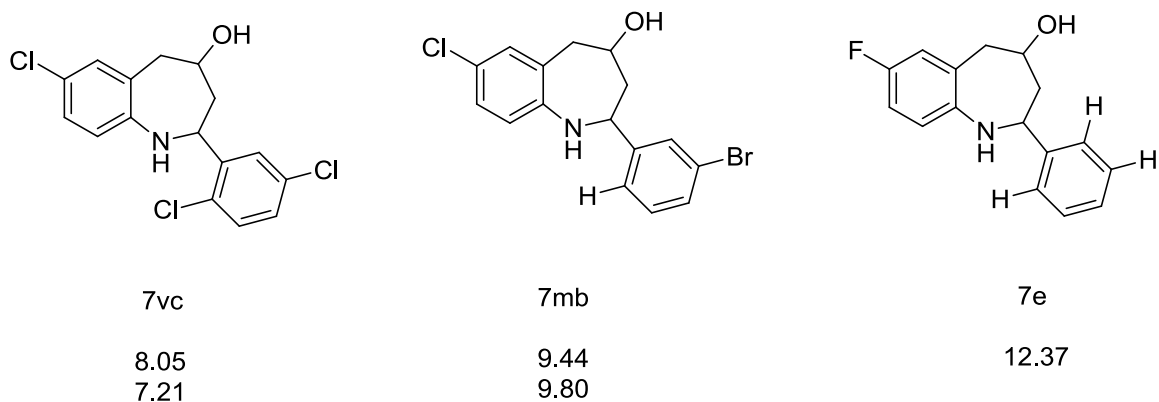
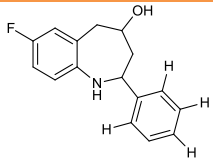
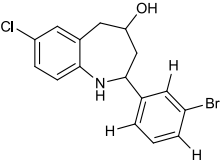
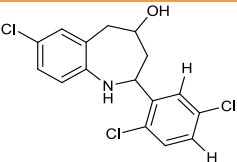


Figura 18. Estructuras de las moléculas más activas.

Es importante partir de la molécula de la cual se tiene una referencia, es decir la molécula 7e; donde se nota que la actividad dada es cuando los descriptores toman los valores de; Mor10e: -0.460, Mor24e: 0.281 y Mor21p: -0.543, observando que deben ser cifras negativas para el caso de Mor10e y Mor21p (bajas), en cuanto a Mor24e deben ser cifras positivas (altas), haciendo una confrontación con los descriptores Mor10e y Mor24e ya que ambos se refieren a la electronegatividad de Sanderson, el valor de s, que es representado por 10 y 24 nos indican la dispersión en varias direcciones, por lo tanto cuando los valores de “s” sean ascendentes deben ser valores altos en comparación con los valores de “s” descendentes; en el caso de la polarizabilidad para que se tenga una buena actividad deben ser valores bajos. Lo mismo se observa en las moléculas 7mb y 7vc y esta relación la podemos visualizar en la tabla 44.

Tabla 44. Valores de los descriptores del modelo 5 de las moléculas 7e, 7mb y 7vc.

Moléculas	Mor21v	Mor10e	Mor24e	Actividad calculada	Actividad experimental
 <p>7e</p>	-0.507	-0.460	0.281	20.84	12.37
 <p>7mb</p>	-0.304	0.206	0.176	9.80	
 <p>7vc</p>	-0.414	-0.716	0.184	7.21	

Para apreciar más los valores de los descriptores, estos se representan por medio de las siguientes gráficas.

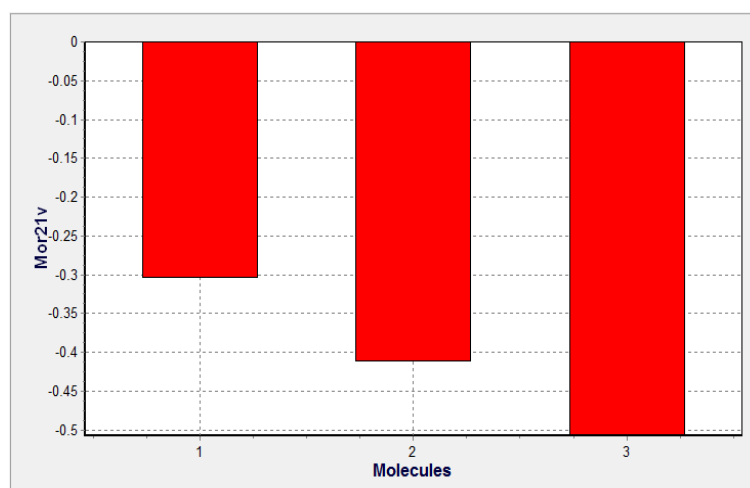


Figura 19. Representación de barras de los valores del descriptor Mor21v, para las moléculas 7mb: 1, 7vc: 2 y 7e: 3, respectivamente.

La coordenada 1 corresponde al compuesto 7vc, siendo la que tiene valores más bajos del descriptor Mor21v y es la que tiene una mejor actividad de las 3.

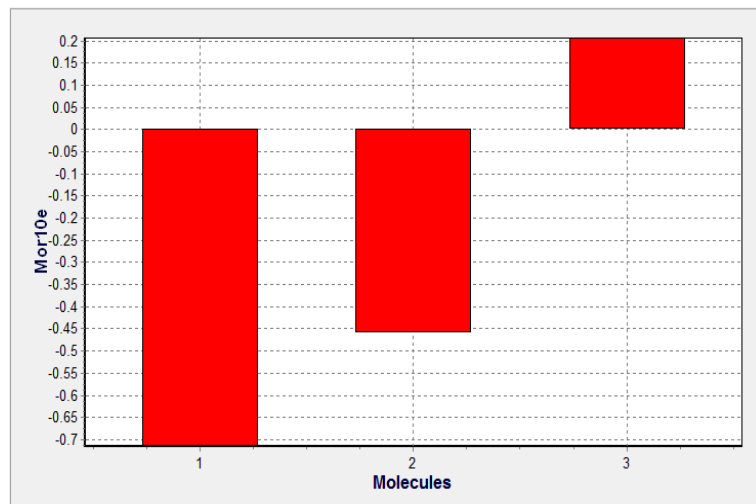


Figura 20. Valores del descriptor Mor10e para los compuesto 7vc: 1, 7e:2 y 7mb: 3.

En la Figura 21 se representa los valores del descriptor Mor10e que como se menciona arriba, para que los compuestos presenten una buena actividad deben tener valores altos y se ve con el compuesto 7mb representada con la coordenada 1 y el compuesto 7vc asignada en la coordenada 2, estas son moléculas de la serie propuesta que comparando sus valores para el descriptor es notable que cuando este asciende mejora la actividad.

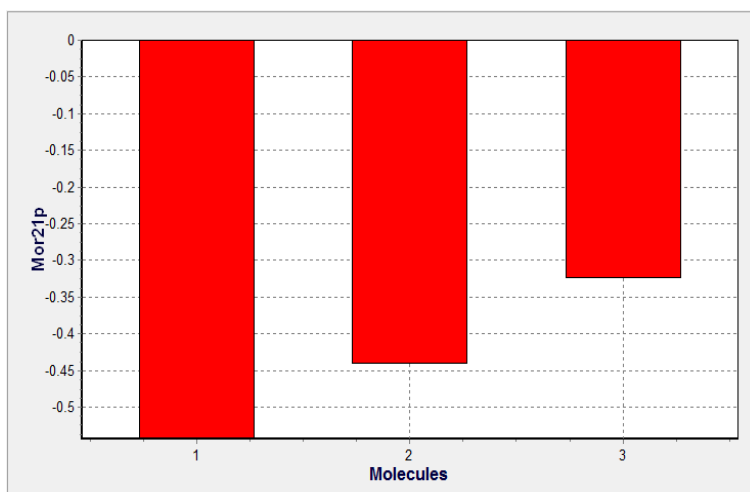


Figura 21. Representación de los valores del descriptor Mor21p para los compuestos 7e: 1, 7vc: 2vc y 7mb: 3.

Enseguida en la Figura 21 se ilustra los valores del descriptor Mor21p, donde se observa que la molécula con mejor actividad es la que presenta un valor bajo para el descriptor, este análisis se hace comparando la molécula 7vc y 7mb representadas con las coordenadas 2 y 3 respectivamente, donde la molécula 7vc es la que presenta mejor actividad con el valor más bajo para el descriptor.

Finalmente, se presenta la figura 22 que muestra los valores del descriptor Mor10u.

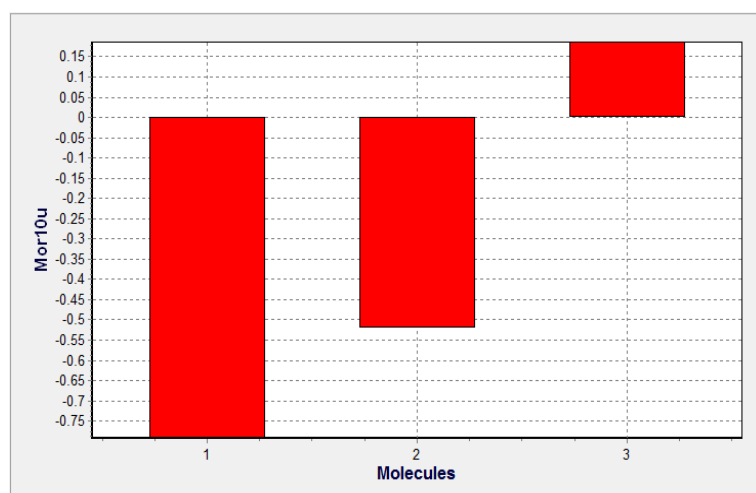


Figura 22. Valores del descriptor Mor10u para los compuestos, 7vc: 1, 7e: 2 y 7mb: 3.

Tabla 45. Tabla de electronegatividades de diferentes fuentes.

Electronegatividades	H	C	N	O	F	Cl	Br
Pauling	2.1	2.5	3.0	3.5	4.0	3.0	2.8
Pauling	2.20	2.55	3.04	3.44	3.98	3.16	2.76
Sanderson	2.31	2.47	2.93	3.46	3.92	3.28	3.96
Sanderson	2.502	2.745	3.194	3.654	4.0	3.475	3.219

Por lo que se deduce que las moléculas que presentarán mejor actividad son las que están sustituidas en R y R₃ con los átomos menos electronegativos.

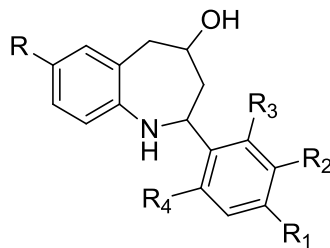


Figura 23. Representación del núcleo común

VI. 4 Citotoxicidad:

Una ventaja que presenta la serie de compuestos explorada es que se cuenta con sus datos de citotoxicidad. Esto no suele ser común y resulta bastante interesante pues no únicamente se puede determinar la actividad contra *Trypanosoma* sino además el origen molecular de su citotoxicidad. Contando con ambos datos, es posible entonces, al menos en teoría, proponer estructuras que cumplan con los requisitos deseables para un fármaco: selectividad y lo menos posible de efectos secundarios indeseables.

Dado que los fármacos que actúan sobre diversos los parásitos, pueden presentar citotoxicidad por la semejanza en rutas bioquímicas o estructura de proteínas putativas a organismos pluricelulares. Este es el caso de los compuestos aquí explorados.

Surgió entonces el interés y la necesidad de llevar a cabo un estudio QSAR de estos compuestos utilizando sus datos de citotoxicidad sobre células vero.

A diferencia de los resultados obtenidos para actividad contra *Trypanosoma*, la citotoxicidad se explica principalmente por los descriptores GAT y de Coeficientes de distribución radial (RDF) y aunque también hay cierto parecido estadístico entre los diferentes modelos, se optó por llevar a cabo el mismo tamiz que se utilizó en el QSAR del apartado anterior, es decir se realiza una validación adicional para contar con el “mejor modelo” posible. Infortunadamente, en este caso no se contó con una cantidad de compuestos lo suficientemente grande (únicamente disponen de 12 compuestos que presentaron valores menores a 300) como para excluir varios compuestos de la población. Así, se excluyó únicamente un compuesto de la serie.

Los resultados de validación se resumen en la Tabla 46 en donde se aprecia que los modelos que más distan en predecir la actividad del compuesto externo son los modelos 3 y 4 con una diferencia de más del 20% mientras que los otros dos modelos, modelo 10 y 17, se alejan de la predicción en aproximadamente 10

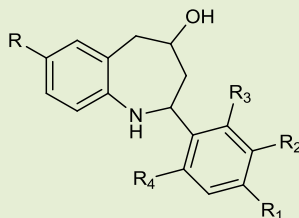
%. Ambos modelos son estadísticamente significativos y pueden considerarse como aceptables.

Tabla 46. Los mejores modelos y sus principales parámetros estadísticos.

Número de modelo	Descriptores	R ²	Q ² boot	Kx	R ² Estimación lineal
Modelo 3	RDF060u RDF015m RDF070m	98.12	86.65	42.5	0.981
Modelo 4	GATS8v GATS2p RDF070m	98.36	93.96	33.92	0.983
Modelo 10	GATS2m GATS8v RDF070m	96.79	89.51	31.9	0.967
Modelo 17	MATS8v GATS2p RDF070m	97.33	90.89	42.66	0.973

Así, se predijo entonces la actividad citotóxica de los compuestos que fueron propuestos como buenos candidatos contra *Trypanosoma* en el apartado anterior. Los resultados de actividad calculada se resumen en la Tabla 47.

Tabla 47. Presenta los valores de citotoxicidad calculados con el modelo 3,4, 10 y 17 para la serie propuesta.



Compuesto	R	R ₁	R ₂	R ₃	R ₄	Mod 3	Mod 4	Mod 10	Mod 17
7fa-nva (1)	H	H	H	H	Cl	240.92	272.10	295.07	326.99
7f-nva (2)	H	Cl	H	H	H	212.80	227.28	255.96	268.55
7ha-nva (3)	H	H	H	OMe	H	44.73	272.06	919.11	402.82
7hb-nva (4)	H	OMe	H	H	H	408.98	295.43	1095.75	339.12
7hc-nva (5)	H	H	H	H	OMe	71.66	287.14	980.08	402.82
7h-nva (6)	H	H	OMe	H	H	321.27	275.90	1036.93	308.38
7ia-nva (7)	H	H	Me	H	H	151.84	238.11	608.78	361.52
7ib-nva (8)	H	Me	H	H	H	113.90	247.46	652.95	334.23
7ic-nva (9)	H	H	H	Me	H	366.53	256.28	664.45	361.52
7i-nva (10)	H	H	Me	H	H	108.34	326.99	792.95	496.23
7j-nva (11)	H	H	Cl	H	H	170.48	232.98	245.30	326.99
7ka-nva (12)	Me	H	H	Me	H	76.34	124.03	342.29	233.07
7kb-nva (13)	Me	H	Me	H	H	307.27	162.93	435.49	293.33
7kc-nva (14)	Me	H	H	H	Me	131.99	121.67	334.57	233.07
7k-nva (15)	Me	Me	H	H	H	889.84	143.41	416.41	216.37
7la-nva (16)	Cl	H	H	H	Cl	199.96	122.89	127.72	185.10
7lb-nva (17)	Cl	H	Cl	H	H	326.40	342.53	377.42	354.67
7l-nva (18)	Cl	Cl	H	H	H	105.60	100.50	108.09	153.28
7ma-nva (19)	Cl	H	H	Br	H	58.04	54.80	58.55	128.72
7mb-nva (20)	Cl	H	Br	H	H	65.45	82.54	85.92	148.75
7mc-nva (21)	Cl	H	H	H	Br	208.93	82.08	94.69	128.72
7m-nva (22)	Cl	Br	H	H	H	44.52	48.55	53.81	99.36
7na-nva (23)	F	H	H	Cl	H	249.69	221.62	184.94	317.67
7nb-nva (24)	F	H	Cl	H	H	117.42	222.65	180.78	345.99
7nc-nva (25)	F	H	H	H	Cl	224.95	241.83	205.18	317.67
7n-nva (26)	F	Cl	H	H	H	173.59	200.75	176.96	260.53
7o-nva (27)	H	H	H	Cl	H	166.25	288.21	307.31	354.67
7pa-nva (28)	H	H	F	H	H	210.66	372.80	494.82	519.07
7pb-nva (29)	H	F	H	H	H	395.32	429.78	593.37	496.69
7pc-nva (30)	H	H	H	H	F	262.85	392.54	522.66	513.38
7p-nva (31)	H	H	H	F	H	266.10	410.22	550.80	513.38
7qa-nva (32)	Me	H	Cl	H	H	165.72	104.31	131.36	191.21

7qb-nva (33)	Me	H	H	H	Cl	54.45	79.90	97.59	181.21
7qc-nva (34)	Me	Cl	H	H	H	370.70	97.70	130.52	157.03
7q-nva (35)	Me	H	H	Cl	H	88.98	73.78	88.75	181.21
7r-nva (36)	Cl	H	H	Cl	H	112.13	107.33	108.71	185.10
7sa-nva (37)	Cl	H	F	H	H	39.710	140.84	117.03	264.15
7sb-nva (38)	Cl	H	H	H	F	132.22	170.74	146.17	260.53
7sc-nva (39)	Cl	F	H	H	H	409.04	185.05	163.96	252.41
7s-nva (40)	Cl	H	H	F	H	171.76	180.32	155.98	260.53
7ta-nva (41)	H	Cl	Cl	H	H	388.01	177.49	176.82	257.95
7tb-nva (42)	H	Cl	H	Cl	H	529.36	153.07	163.99	194.24
7tc-nva (43)	H	H	Cl	H	Cl	124.97	211.80	218.21	257.95
7td-nva (44)	H	Cl	H	H	Cl	387.96	151.61	162.14	194.24
7t-nva (45)	H	H	H	Cl	Cl	201.21	160.28	161.11	234.57
7va-nva (46)	Cl	H	H	Cl	Cl	214.99	74.53	92.53	120.25
7vb-nva (47)	Cl	H	Cl	Cl	H	174.18	92.64	108.25	160.79
7vc-nva (48)	Cl	H	Cl	H	Cl	91.89	87.25	100.80	160.79
7vd-nva (49)	Cl	H	H	H	H	269.53	86.48	110.44	120.25
7v-nva (50)	Cl	H	H	Cl	Cl	46.97	49.65	53.01	144.22
7wa-nva (51)	Cl	F	H	Cl	H	365.30	110.77	102.35	180.82
7wb-nva (52)	Cl	H	F	Cl	H	240.41	113.04	103.57	189.49
7wd-nva (53)	Cl	H	Cl	H	F	72.26	87.04	73.23	205.81
7we-nva (54)	Cl	Cl	H	H	F	155.51	88.36	82.59	154.12
7w-nva (55)	Cl	H	H	H	H	101.93	83.02	71.22	186.38

La familia 6 que consta de noventa y seis descriptores de tipo autocorrelaciones 2D, calculados por DRAGON son autocorrelaciones espaciales calculadas sobre un gráfico molecular medido de H-reducido por propiedades físico-químicas del átomo (w) que incluyen:

- Autocorrelaciones ATS (*Autocorrelation of a Topological Structure*).
- Autocorrelaciones MATS calculado por el coeficiente de Moran.
- Autocorrelaciones GATS calculado por el coeficiente de Geary.

Estos descriptores moleculares describen cómo una propiedad considerada se distribuye a lo largo de una estructura molecular topológica

El GATSkw autocorrelación de Geary, siendo W la propiedad atómica para ponderar el gráfico molecular y el k =lag, se calcula aplicando el coeficiente de Geary a la gráfica molecular:

$$GATSkw = \frac{\frac{1}{2\Delta} - \sum_{l=1}^{nSk} \sum_{j=1}^{nSk} \delta_{ij} - (w_i - w_j)^2}{\frac{1}{(nSk - 1)} - \sum_{l=1}^{nSk} (w_l - \bar{w})^2}$$

Donde:

W_i : es alguna propiedad atómica

\bar{w} : es un valor promedio en la molécula

nSk: es el numero non de átomos de hidrogeno

δ_{ij} : 1

k: dij

k: es la distancia topológica entre 2 átomos considerados

Δ : es la suma de los deltas de Kronecker.

El símbolo para cada uno de los descriptores de autocorrelación es seguido por dos índices d y w en donde d representa el lag y w representa la propiedad. El lag se define como la distancia d topológica entre los pares de átomos. La distancia topológica entre un par de átomos (ij) se da en la entrada "ij" de la matriz nivel topológico. El lag puede tener un valor entre 0 y 8. La (w) puede ser m (masa atómica relativa), p (polarizabilidad), e (electronegatividad de Sanderson) y v (volumen de Van der Waals). Masa relativa se define como la relación de masa de un átomo respecto al carbono. Del mismo, p, e y v son escalas por los valores correspondientes para el carbono. Sea n el número de átomos en la molécula³⁰.

El coeficiente de Geary, es una función de tipo distancia que varía de cero a infinito. Es una autocorrelación espacial fuerte que produce valores bajos de este índice; por otra parte se traduce a autocorrelación positiva en valores entre 0 y 1, mientras que la autocorrelación negativa produce valores superiores a 1, además la referencia de “no correlación” es el valor del coeficiente igual a 1. Autocorrelación en lag 0 no siempre es una suma de los cuadrados de las propiedades atómicas. Cabe mencionar que las propiedades atómicas w , que usan gráficos de pesos moleculares son: masa atómica (m), volumen atómico de Van Der Waals (v), electronegatividad atómica de Sanderson (e) y polarizabilidad atómica (p).

Finalmente para evitar números grandes que dificultan su manejo, se transforman a logaritmos de valores de autocorrelación Moreau-Broto (ATS) como $\ln(1+\text{valor})$.

Autocorrelación Geary

Índice general de la autocorrelación espacial que, si se aplica a una gráfica molecular, puede ser definida como:

$$c(d) = \frac{\frac{1}{2\Delta} \cdot \sum_{i=1}^A \sum_{j=1}^A \delta_{ij} \cdot (w_i - w_j)^2}{\frac{1}{(A-1)} \cdot \sum_{i=1}^A (w_i - \bar{w})^2}$$

Donde w_i es una propiedad atómica, \bar{w} es el valor promedio en la molécula, A es el número atómico, d es la distancia topológica considerada, δ_{ij} es un valor delta Kronecker y Δ es la suma de los deltas Kronecker (número de pares de vértices en una distancia igual a d).

El coeficiente Geary es una función de distintos tipos de distancias de cero a infinito. Una fuerte autocorrelación produce altos valores de este índice, por otra parte, autocorrelaciones positivas se traducen en valores entre 0 a 1, mientras que autocorrelaciones negativas producen valores grandes donde 1, por lo tanto, hace referencia a “no correlación” cuando c es igual a 1³⁷.

Los 150 descriptores de la familia 13, descriptores de distribución radial RDF por sus siglas en inglés (*Radial Distribution Function*) recientemente se han propuesto sobre la base de una función de distribución radial que se puede interpretar como la probabilidad de encontrar un átomo en un volumen esférico de radio R ⁴⁵.

La forma general de la función de distribución radial está representado por:

$$RDFRw = f - \sum_{l=1}^{nAT-1} \sum_{j=l+1}^{nAT} w_l - w_j \cdot e^{-\beta(R-r_{ij})^2}$$

Donde:

f: factor de escala

w: propiedades características de los átomos i y j

r_{ij} : es la distancia interatómica

nAT: es el número de átomos en la molécula

El término exponencial contiene la distancia interatómica r_{ij} y el parámetro suavizable β (Å^{-2}).

Los cuales definen la distribución probable de la distancia interatómica individual, β puede ser interpretado como un factor de temperatura, que define el movimiento de los átomos³⁰.

RDFRw se calcula generalmente en un número de puntos discretos con intervalos definidos. En DRAGON, se utilizan un valor β igual a 100 Å^{-2} y un tamaño igual a $0,5 \text{ Å}$, por lo tanto, 30 descriptores RDF son calculados por cinco diferentes esquemas de ponderación: el caso no ponderado (u), de masa atómica (m), el volumen de Van der Waals (v), la electronegatividad atómica de Sanderson (e) y, la polarizabilidad atómica (p).

Además de la información sobre las distancias interatómicas en la molécula completa, los descriptores RDF proporcionan más información, sobre distancias de enlace, tipos de llamada, sistemas planos y no planos y tipos de átomos. Este hecho es una consideración importante para una elucidación código asistida por ordenador.

Un típico descriptor RDF se denota por RDF_{sw} donde $1,0 \leq s \leq 15,5$ en unidades de 0,5 y w como se ha definido anteriormente.

El modelo 3 incluye los descriptores:

RDF060u: Función de distribución radial - 6.0 / ponderado por masas atómicas.

RDF015m: Función de distribución radial - 1.5 / ponderado por masas atómicas.

RDF070m: Función de distribución radial - 7.0 / ponderado por masas atómicas.

De acuerdo al modelo 4, los descriptores son:

GATS8v: Autocorrelación de Geary - lag 8 / ponderado por volumen atómico de van der Waals.

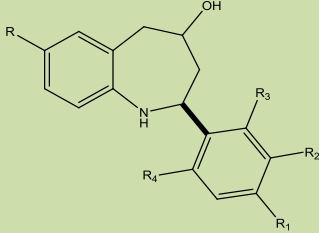
GATS2p: Autocorrelación de Geary - lag 2 / ponderado por la polarizabilidad atómica.

RDF070m: Función de distribución radial - 7.0 / ponderado por masas atómicas.

En el caso del modelo 4 donde se tienen dos descriptores de autocorrelación, GATS8v ponderado para volumen atómico de van der Waals o radio, que se determinan a partir de las medidas de las distancias internucleares mediante una distribución correcta de tales distancias. Tal distribución debe tener el entorno de los átomos vecinos, la naturaleza de las interacciones entre los átomos y la presencia o ausencia de enlace y GATS2p que mide la polarizabilidad que nos indicara sobre la respuesta de la molécula a un campo externo. Por su

parte el descriptor RDF para sus valores toma en cuenta la masa atómica. Para el análisis de los descriptores se tomarán como ejemplo los compuestos 7c, 7j, 7ma y 7pb (Tabla 48). Con la finalidad de evaluar los compuestos que presentan toxicidad y los que no la presentan, se consideraron cuatro compuestos con valores más bajos, así como dos compuestos con valores más altos, que el Nifurtimox (61.42).

Tabla 48. Compuestos para utilizados para la validación de los modelos.



Compuesto	R	R ₁	R ₂	R ₃	R ₄	Mod. 3	Mod. 4	Actividad experimental
7c	Br	H	H	H	H	42.27	47.83	42.14
7j	H	H	Cl	H	H	314.25	301.05	295.36
7ma	Cl	H	H	Br	H	58.04	54.80	
7pb	H	F	H	H	H	395.32	429.78	
Nifurtimox								61.42

El uso de los valores de citotoxicidad se formula mediante la siguiente lógica: daño celular a bajas concentraciones indica bajas dosis para presentar toxicidad, en cambio una alta concentración refleja daño celular significará que para que cause toxicidad tendrían que manejarse dosis altas.

Primeramente se llevó a cabo la evaluación de los descriptores GATS8v y GATS2p, ya que son de la misma familia y únicamente difieren en el parámetro de ponderación. Las moléculas con alta toxicidad son las que presenta para el caso del descriptor GATS8v valores altos, esto se hizo comparando los valores del descriptor en las moléculas 7c y 7pb, lo que podemos sospechar que el radio de

Van der Waals para los sustituyentes deben ser altos para que disminuya la toxicidad. (Tabla 49).

Tabla 49. Valores de masa, volumen de van der Waals, electronegatividad y polarizabilidad.

Simbolo	Masa	V d W Volume	Electronegatividad	Polarizabilidad
C	1.000	1.000	1.000	1.000
H	0.084	0.299	0.942	0.381
N	1.167	0.695	1.160	0.625
O	1.332	0.512	1.327	0.455
P	2.579	1.181	0.916	2.063
S	2.670	1.088	1.076	1.648
F	1.582	0.410	1.455	0.318
Cl	2.952	1.000	1.265	1.239
Br	6.653	1.384	1.171	1.733
I	10.56	1.728	1.011	3.040
Si	2.339	1.424	0.778	3.057

Esto resulta del todo lógico al analizar los compuestos 7j y 7ma, ya que los valores para el descriptor GATS8v son bajos para el compuesto 7ma y presenta una toxicidad alta y en el caso del compuesto 7j es alto el valor del descriptor y su citotoxicidad es baja. (ver Figura 25)

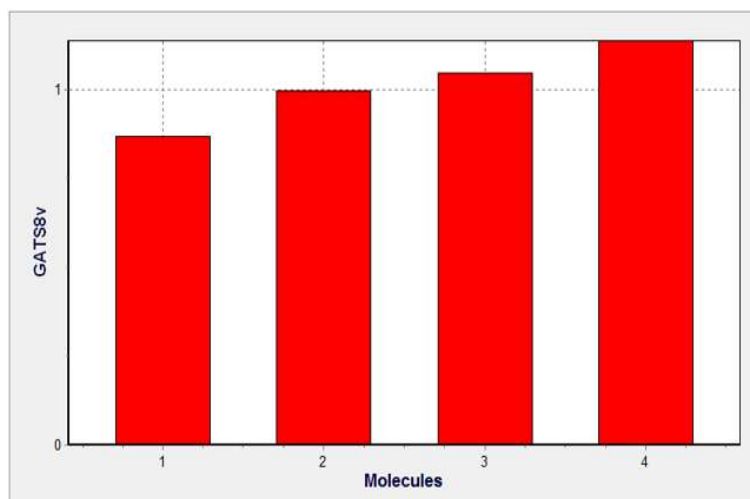


Figura 24. Valores del descriptor GATS8v para los compuestos 7c: 1, 7pb: 2, 7ma: 3 y 7j: 4.

En cuanto al descriptor GATS2p (Figura 25), los valores indican que cuando su valor sea más alto será menos toxico, esto se concluye al comparar el compuesto 7c contra 7pb y 7j. Lo que puede se interpreta a que entre más alto es el valor de la polarizabilidad menos toxico es el compuesto, además de que se puede marcar que las menos toxicas son las que no tienen sustituyente en la posición R.

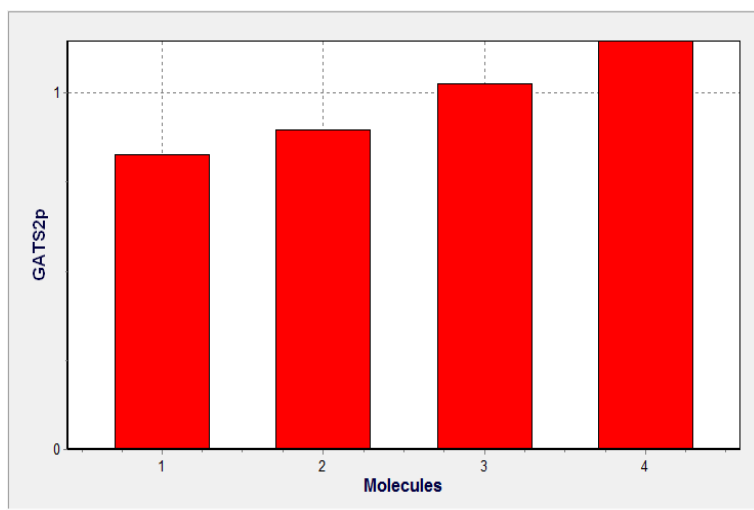


Figura 25. Valores del descriptor GATS2p para los compuestos 7ma: 1, 7c: 2, 7j: 3 y 7pb: 4.

Por último se aprecia que para el descriptor RDF070m (Figura 26), entre más bajos sean los valores menos tóxicos serán los compuestos, dado que el descriptor esta medido en base a la masa atómica, por lo tanto entre más alta sea la masa atómica más tóxico será el compuesto. Esto se deduce al comparar 7pb con 7c.

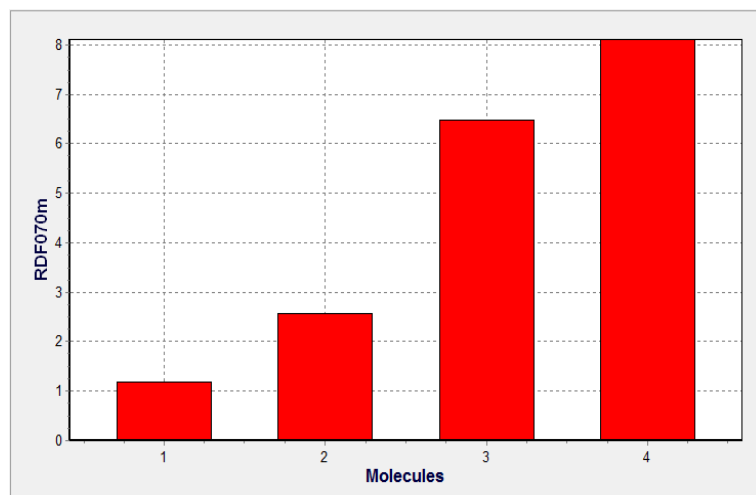


Figura 26. Valores del descriptor RDF070m para los compuestos 7j: 1, 7pb: 2, 7ma: 3 y 7c: 4.

Con la información de los resultados del QSAR de actividad y toxicidad se analizan los compuestos potencialmente activos contra *Trypanosoma cruzi* y menos tóxicos que el Nifurtimox. Para esto se grafican los compuestos de la serie propuesta con valores menores de citotoxicidad calculada de 300 μ M. Se emplearon los modelos 3 y 5 para la actividad calculada y los modelos 3 y 4 para la citotoxicidad.

La primera comparación es con el modelo 3 de actividad y el modelo 3 de citotoxicidad representada en la figura 30, donde se etiqueta a la molécula que de acuerdo a la representación tiene un valor de 269.53 concentración para citotoxicidad (lo que indica que es menos tóxica) y un valor de coordenada para actividad de 9.66, con esto se visualiza que la molécula 7vd, que corresponde en la gráfica a estos valores es una buena propuesta. Ahora tomando el valor más bajo de concentración para actividad calculada, se tiene a la molécula 7vc donde presenta un valor calculado de actividad de 8.05 y 91.89 para citotoxicidad, indicador de que se tiene una buena propuesta de diseño. (ver Figura 27)

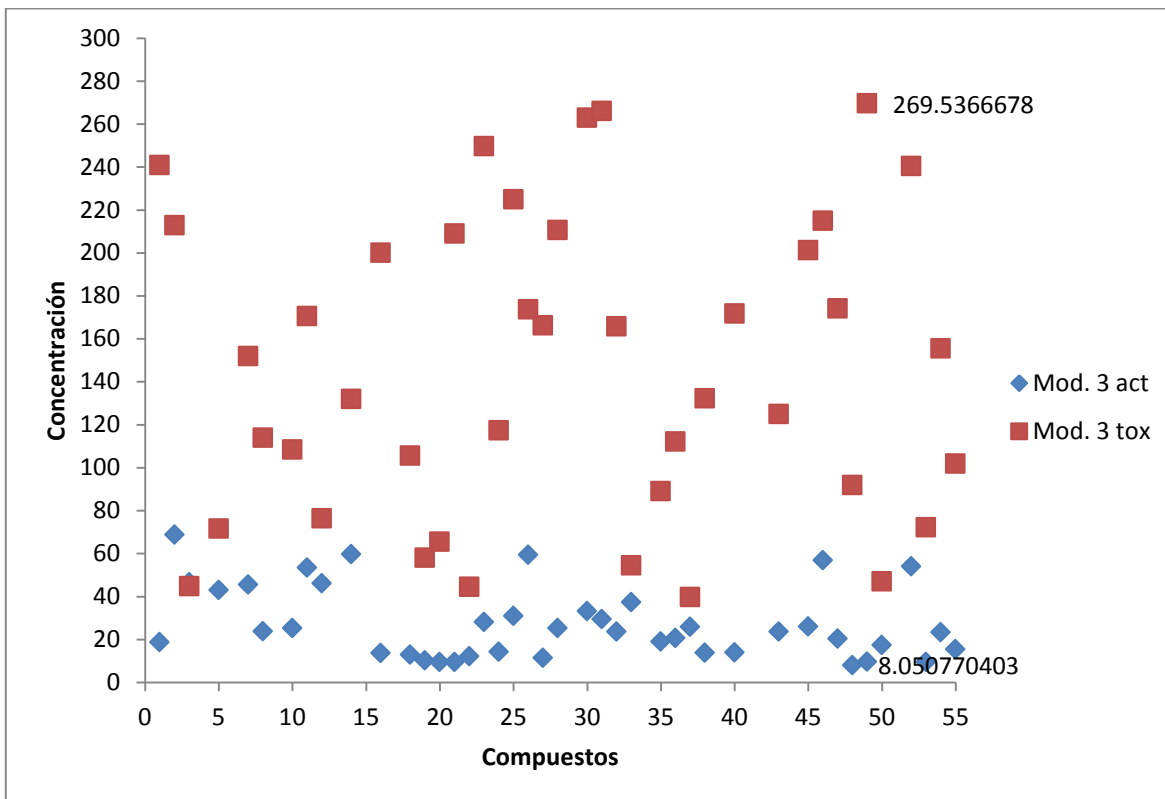


Figura 27. Comparación de la actividad calculada con el modelo 3 y la citotoxicidad calculada con el modelo 3.

En la Figura 28 se compara a la actividad de los compuestos calculada con el modelo 5 y la citotoxicidad calculada con el modelo 3, donde el punto más elevado tiene un valor de 269.53 para citotoxicidad y un valor de 10.32 de actividad, el compuesto que corresponde a estos valores es el 7vd. Ahora para el valor de actividad más bajo se tiene al compuesto 7vc con 7.21 y para toxicidad 91.89.

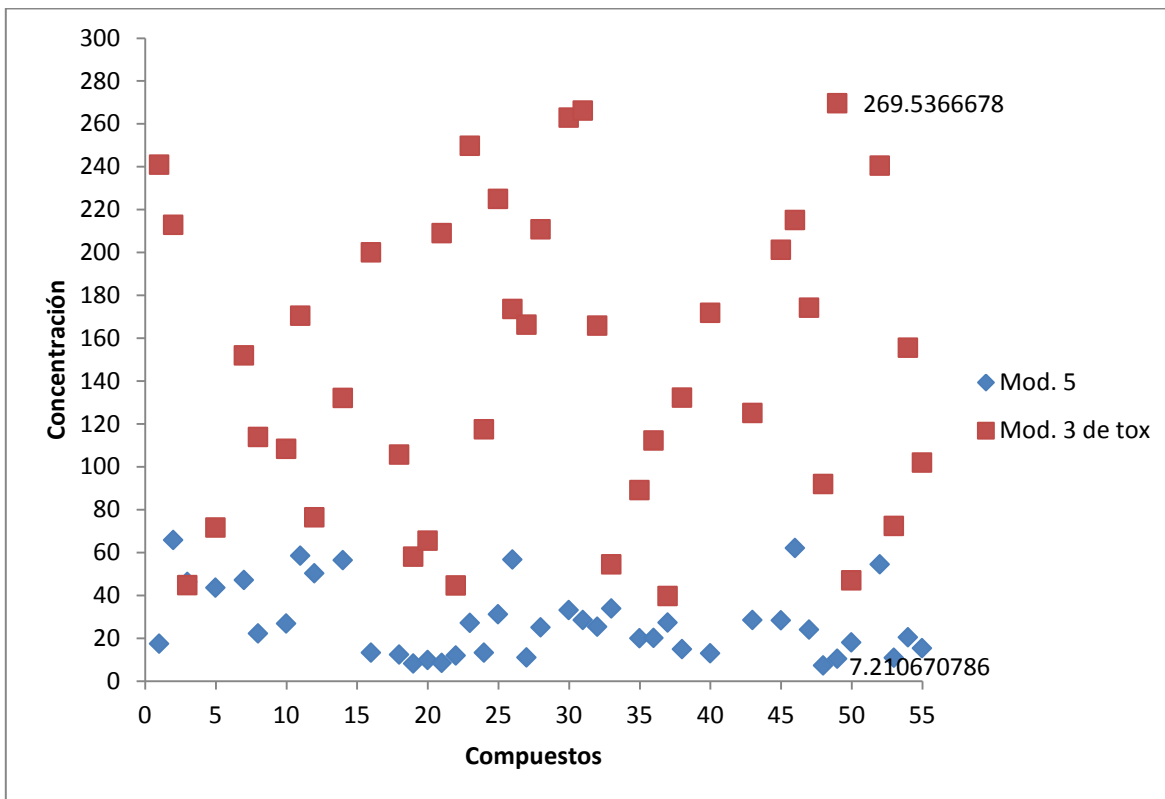


Figura 28. Comparación de la actividad calculada con el modelo 5 y la citotoxicidad calculada con el modelo 3.

Se hizo la comparación entre los valores de actividad del modelo 3 y los valores de toxicidad del modelo 4 (Figura 29), resultando prometedor el compuesto 7hb con un valor de 30.22 para actividad y 295.43 de toxicidad. El compuesto 7vc presenta un valor de actividad de 8.05 y de toxicidad de 87.25 siendo un compuesto adecuado contra *Trypanosoma cruzi*.

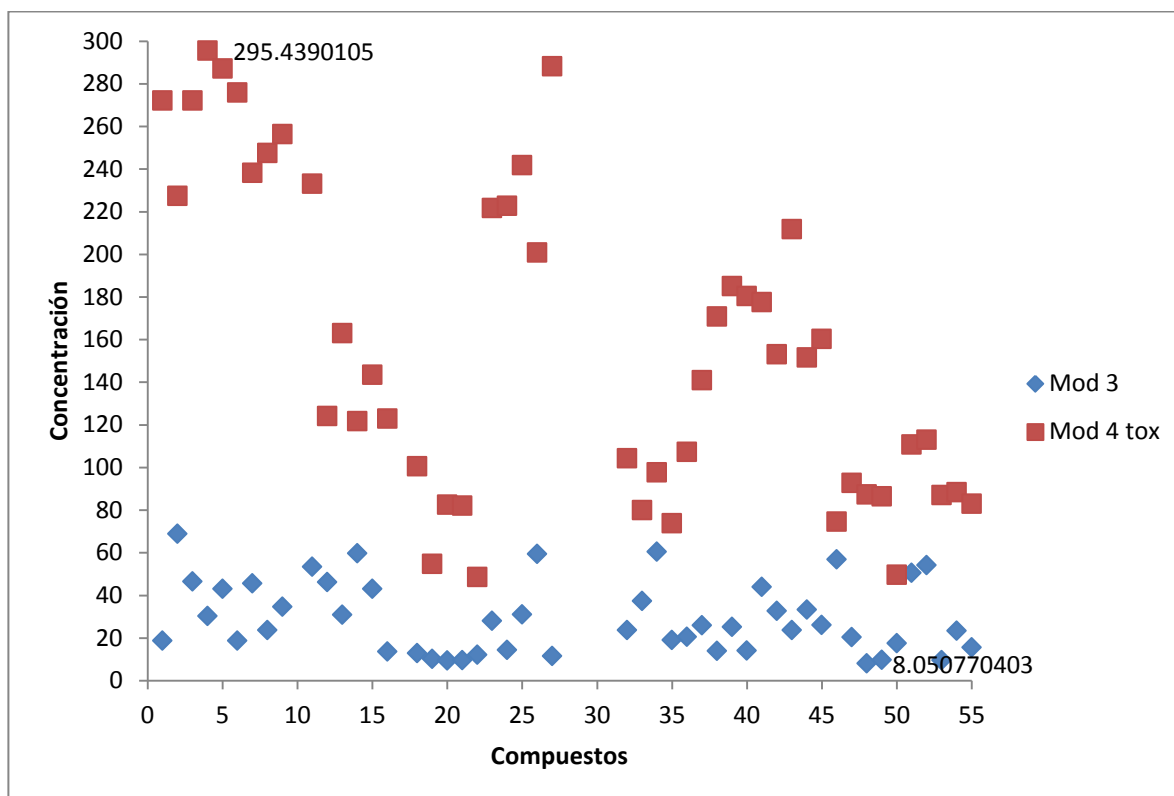


Figura 29. Comparación de la actividad calculada con el modelo 3 y la citotoxicidad calculada con el modelo 4.

Por último en la Figura 30 se gráfica la actividad calculada con el modelo 5 y la citotoxicidad calculada con el modelo 4, donde se tiene que el compuesto 7hb tiene valores de 295.43 para toxicidad y 28.66 para actividad. El compuesto 7vc tiene una actividad calculada de 7.21 y un valor de citotoxicidad de 87.25, que comparado con los valores reportados para el Nifurtimox es considerado un buen compuesto.

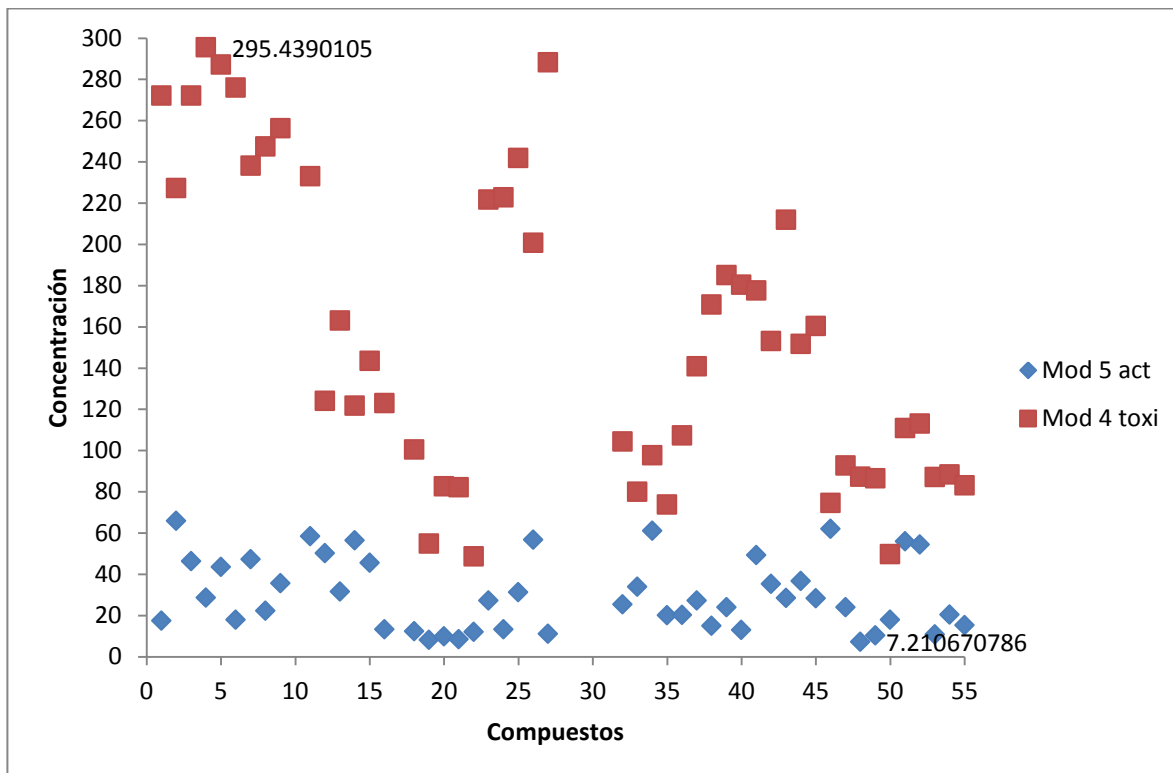


Figura 30. Comparación de la actividad calculada con el modelo 3 y la citotoxicidad calculada con el modelo 4.

VII.- CONCLUSIONES:

- Se lleva a cabo un estudio QSAR por técnica de algoritmos genéticos, del cual se logran modelos matemáticos que permiten determinar la actividad de la serie 7a-w y de nuevos derivados de cis-2-aryl-4-hidroxitetrahydro-1-benzacepina, con ello concluir que la actividad de este tipo de compuestos está relacionada con la Familia 3DMoRSE que incluye a los descriptores Mor

- Se aplicó el estudio QSAR para la obtención de modelos que permiten determinar la citotoxicidad de los derivados de cis-2-aryl-4-hidroxitetrahydro-1-benzacepina, donde los modelos nos expresan el origen molecular de su citotoxicidad que está relacionada con las familias de autocorrelaciones 2D y descriptores RDF, determinada por los descriptores GATS y RDF.

- Los modelos matemáticos obtenidos por la técnica QSAR, permitirán predecir la actividad de otros compuestos que tengan como estructura base el sistema de anillo tetrahydro-1-benzacepina.

- Los modelos no solo predicen la actividad y citotoxicidad de compuestos con un sistema de anillo tetrahydro-1-benzacepina, sino que también permiten explicar la actividad y citotoxicidad de dichos compuestos.

VIII.- BIBLIOGRAFIA:

¹<http://www.paho.org/Portuguese/AD/DPC/CD/res-descentralizacion.pdf>; Descentralización y gestión del control de las enfermedades transmisibles en América latina.

²http://medicina.ufm.edu/index.php/Enfermedad_de_Chagas,_Chagas-Mazza_o_Tripanosomiasis_Americana, Enfermedad de Chagas, Chagas-Mazza o Tripanosomiasis Americana. Facultad de medicina. Universidad Francisco Marroquín.

³ <http://www.who.int/mediacentre/factsheets/fs340/es/>. La enfermedad de Chagas (tripanosomiasis americana) Nota descriptiva N°340, Agosto de 2012.

⁴ <http://www.intramed.net/contenidover.asp?contenidoID=71751>; Tratamiento para la enfermedad de Chagas crónica.

⁵ Velasco-Castrejón, O., Valdespino, J.L., Tapia-Conyer, R., Salvatierra, B., Guzmán-Bracho, C., Magos, C., Llausás, A., Gutiérrez, G., Sepúlveda, J. (1992). Seroepidemiología de la enfermedad de Chagas en México. *Salud Pública de México*. 34, 186-196.

⁶ Biagi Francisco, 2004, *Enfermedades Parasitarias*, 3ª edición., editorial Manual Moderno, 135, 137, 138, 142, 143.p

⁷http://www.cfsph.iastate.edu/Factsheets/es/tripanosomiasis_american_a_chagas.pdf.

⁸ Salazar Schettino Paz M., De Haro Arteaga Irene, Cabrera Bravo Margarita, Tres especies de triatóminos y su importancia como vectores de *Trypanosoma cruzi* en México., *Medicina Buenos Aires*, 2005, 63-65-69.

⁹ Guhl Felipe 1999, *Medicina (Buenos Aires)*, Estado actual del control de la enfermedad de Chagas en Colombia, 59 (suplemento II):103-116

¹⁰ Brooks Geo F., Butel Janet S., Morse Stephen A., 2005, *Microbiología Médica* de Jawetz, Melnick y Adelberg; 18ª Edición; Editorial Manual Moderno.

¹¹ Tay Zavala Jorge, Lara Aguilera Ramón, 2002, *Parasitología médica*. 7ª Edición. Méndez Editores, 117 p.

¹² <http://www.facmed.unam.mx/deptos/salud/periodico/chagas/cruzi.html>.

¹³ Laurence L. Bruton, John S. Lazo, Keith L. Parker; 2006, Las bases farmacológicas de la terapéutica L. S. Goodman, A. Gilman; 11^a Edición; Editorial McGraw-Hill Interamericana, 1061-1063.p.

¹⁴ Brunton L. Laurence; Parker L. Keith; 2006, Manual de farmacología y terapéutica; 11^a Edición; Editorial McGraw-Hill Interamericana, 683, 689 p.

¹⁵ 9.-Gomez-Ayala S., Castrillón A. J., Alirio Palma, Leal M. S., Escobar P., Bahsas A., Synthesis, structural elucidation and in vitro antiparasitic activity against *Trypanosoma cruzi* and *Leishmania chagasi* parasites of novel tetrahydro-1-benzazepine derivatives, *Bioorganic & Medicinal Chemistry*, 2010, 18, 4721-47394.

¹⁶ Carrasco R., Escalona J.C., Padrón J.A.; Introducción al diseño de Fármacos, Folleto para la docencia de la asignatura de Farmacia, Universidad del Oriente.

¹⁷ Yurkanis Bruice P.2008, Química Orgánica, 5^a Edición. Universidad de California, Santa Clara. Pearson,ed., 1299 p.

¹⁸ Galbis Pérez J. A.; *Panorama actual de la química farmacéutica*; Universidad de Sevilla: Sevilla, **2004**.

¹⁹ Diseño molecular de nuevos fármacos, farmacología para todos.com.

²⁰ Nasser Goudarzi, Mohammad Goodarzi, Tao Chen; QSAR prediction of HIV inhibition activity of styrylquinoline derivatives by genetic algorithm coupled with multiple linear, *Medicinal Chemistry Research*; **2012**, 21, 437-443.

²¹ Melián Batista Belén, Moreno Pérez José A., Moreno Vega J. Marcos. Algoritmos Genéticos; *Números*, revista didáctica de las matemáticas; **2009**, Volumen 71, 29–47p.

²² Fernández Michael, Caballero Julio, Morales Helguera Aliuska, Eduardo A. Castro y Pérez González Maykel. Quantitative structure–activity relationship to predict differential inhibition of aldose reductase by flavonoid compounds. *Bioorganic & Medicinal Chemistry* (**2005**), 13. 3269–3277

²³ Todeschini R., Ballabio D., Consonni V., Mauri A., y Pavan M. 2004. **Mobydigs Computer Software**, 1.0; TALETE srl: Milano.

²⁴ Selassie, C. D.; Hitory of Quanttitative Structure-Activity Relationships; *Burger's Medicinal Chemistry and Drug Discovery*: January 15, **2003**.

²⁵ Arteaga Tovar Samuel, Espinosa Piña José Alfredo; **2008**, *Probabilidad y estadística*; 1ra. Edición, Editorial DGETI. 102 p.

²⁶ Murray R. Spiegel, Larry J. Stephens; *Estadística*; 3ra. Edición; Editorial McGrall-Hill. 317 p.

²⁷ López Jáuregui, A. y Elosua Oliden, P., Facultad de Psicología. Universidad del País Vasco. **2004**, Estimaciones bootstrap para el coeficiente de determinación: un estudio de simulación, *Revista Electrónica de Metodología Aplicada.*, Vol. 9 nº 2, 1-14 p.

²⁸ Nikola Minovski, Aneta Jezierska-Mazzarello, Marjan Vračko, Tom Šolmajer; **2011**, Investigation of 6-fluoroquinolones activity against Mycobacterium tuberculosis using theoretical molecular descriptors: a case study; *Central European Journal of Chemistry*;,855-866.

²⁹ Zavala Franco A., González Campos J. B. y Chacó García L., Estudio QSAR por algoritmos genéticos de reconocedores de surco del DNA, *Biológicas*, Diciembre **2010**, 12(2): 108 – 115.

³⁰ Todeschini R., Consonni V., Mauri A. and Pavan M. 2005. **DRAGON—Software for the calculation of molecular descriptors**. Ver. 5.3, Talete srl, Milano, Italy.

³¹ Scotti Marcus T., Fernandez Mariane B., Ferreira Marcelo J. P. and Emerenciano Vicente P. **2007**, Quantitative structure–activity relationship of sesquiterpene lactones with cytotoxic activity. *Bioorganic & Medicinal Chemistry* 15 2927–2934

³² Liane Saíz-Urra, Maykel Pérez González and Marta Teijeira, **2006**, QSAR studies about cytotoxicity of benzophenazines with dual inhibition toward both topoisomerases I and II: 3D-MoRSE descriptors and statistical considerations about variable selection, *Bioorganic & Medicinal Chemistry*, 14, 7347–7358

-
- ³³ Press W, H., et. al., **1968**. Numerical Recipes: The Art of Scientific Computing. *Cambridge University Press*. Chapter 10.
- ³⁴ Microsoft Corporation. (2010). Microsoft® Excel® Ayuda de Excel.
- ³⁵ Paliakov Ekaterina, Henary Maged, Say Martial, Patterson Steven E., Parker Alesia, Manzel Lori, Macfarlane Donald E., Bojarskid Andrzej J. and Strekowskia Lucjan, **2007**. Fujita-Ban QSAR analysis and CoMFA study of quinoline antagonists of immunostimulatory CpG-oligodeoxynucleotides. *Bioorganic & Medicinal Chemistry*, *15*, 324–332.
- ³⁶ Qing-Li Wei, Shu-Sheng Zhang, Jun Gao, Wei-hua Li, Liang-Zhong Xua and Zhi-Gang Yub: **2006**, Synthesis and QSAR studies of novel triazole compounds containing thioamide as potential antifungal agents. *Bioorganic & Medicinal Chemistry*, *14*, 7146–7153.
- ³⁷ Rudolf Kiralj, Márcia M. C. Ferreira; *J. Braz.* **2009**, Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application; *Chem. Soc.*, *20*, 770-787
- ³⁸ Solanas Antonio, Sierra Vicenta; **1992**, *Bootstrap: fundamentos e introducción a sus aplicaciones*, Universidad de Barcelona, *Anuario de Psicología*, *55*, 143-154.
- ³⁹ Ping Lan, Wan-Na Chen, Zhi-Jian Huang, Ping-Hua Sun, Wei-Min Chen. **2011** Understanding the structure-activity relationship of betulinic acid derivatives as anti-HIV-1 agents by using 3D-QSAR and docking. *J Mol Model*, *17*:1643–1659
- ⁴⁰ Black, Anand rangarajan Michael J.; On the Unification of Line Processes, Outlier Rejection, and Robust Statistics with Applications in Early Vision; *International Journal of Computer Vision* 19(1), 57-91 (1996).
- ⁴¹ Delgado de la Torre Rosario; 2008, *Probabilidad y estadística para ciencias e ingenierías*, Primera edición, Publicaciones Delta, 211 p.
- ⁴² Sarchitect Designer, version 2.5 User Guide. Copyright © 2005-2008. Strand Life Sciences Pvt. Ltd.
- ⁴³ Zavala Franco Anai; **2012**, *Estudio QSAR por Algoritmos Genéticos de Reconocedores de Surco del DNA*. Tesis de licenciatura. Escuela de Químico Farmacobiología, Universidad Michoacana de San Nicolás de Hidalgo. Morelia Mich. 66, 68 p.
-

⁴⁴ Moeller T., **1994**; *Química Inorgánica*, nva. versión, Reverte Editorial, 71,89 p.

⁴⁵ Vicente Esther, Duchowicz Pablo R., Benítez Diego, Castro Eduardo A., Hugo Cerecetto, Mercedes González, Antonio Monge; 2010, Anti-T. cruzi activities and QSAR studies of 3-arylquinoxaline-2-carbonitrile di-N-oxides; *Bioorganic & Medicinal Chemistry Letters*, 20, 4831–4835.