

TRABAJOS ORIGINALES ORIGINALS WORKS

Aplicación de la topología molecular en la predicción de la inhibición de *Trypanosoma cruzi* Hexokinasa y un grupo de derivados bifosfonatos

*Application of molecular topology to the prediction of inhibition
of Trypanosoma cruzi Hexokinase by bisphosphonates*

GARCÍA-DOMENECH R*, ESPINOZA N, GALARZA RF, MORENO-PADILLA MJ, ROJAS-RUIZ B,
ROLDAN-ARROYO LL, SANCHEZ-LAVADO MI, GÁLVEZ J.

Dept. Química Física, Facultad de Farmacia, Universitat de Valencia, Avd. V.A. Estellés, s/n,
46100-Burjassot, Valencia, Spain

* Correspondencia al autor: Tel. +34 963544291. E-mail: ramon.garcia@uv.es

RESUMEN

Se ha desarrollado un modelo topológico-matemático para la búsqueda de nuevos derivados bifosfonatos activos frente a la hexokinasa de *Trypanosoma cruzi*. Utilizando el análisis lineal discriminante se ha seleccionado una función con cuatro variables capaz de predecir adecuadamente la CI_{50} para cada compuesto de las series de entrenamiento y test. El modelo propuesto se ha aplicado a una librería molecular y se han propuesto nuevas estructuras potencialmente activas frente a *T. cruzi*.

PALABRAS CLAVE: Topología molecular. *Trypanosoma cruzi*. Hexokinasa. Análisis lineal discriminante. Librería molecular.

ABSTRACT

A topological-mathematical model has been arranged to search for new derivatives of bisphosphonate compounds acting as inhibitors against *Trypanosoma cruzi* hexokinase. By using linear discriminant analysis, a four-variable function was achieved allowing an accurate prediction of the IC_{50} for each compound of the training and test series. After carrying out a virtual screening based upon such a model, new structures potentially actives against *T. cruzi* are proposed.

KEYWORDS: Molecular topology. *Trypanosoma cruzi* hexokinase. Linear discriminant analysis. Virtual screening.

Fecha de recepción: 20-09-2008

Fecha aceptación: 30-09-2008

INTRODUCCIÓN

La tripanosomiasis americana o enfermedad de Chagas es una afección producida por el protozoo flagelado *Trypanosoma cruzi*. La infección es transmitida por insectos hemípteros hematófagos de la subfamilia triatominae y afecta tanto al hombre como a animales domésticos y salvajes.

INTRODUCTION

The American Trypanosomiasis or Chagas disease is a pathology whose causative agent is the protozoan parasite *Trypanosoma cruzi*. The infection is transmitted by hemiptera bloodfeeding insect vector, subfamily Triatominae which affects human and domestic-wild animals.

Esta dolencia tiene tres formas clínicas: la aguda, la inaparente y la crónica. Alrededor del 30% de los pacientes infectados desarrollan una enfermedad progresiva y grave, con compromiso severo del corazón o de las vísceras del aparato digestivo.

Se estima que hay 18 millones de personas infectadas y 100 millones más con riesgo de infectarse en 15 países latinoamericanos desde México a Argentina¹. Es la tercera enfermedad tropical más importante a nivel mundial en cuanto a morbi-mortalidad y es imposible separarla de la pobreza por lo que genera un problema social y sanitario importante para la población afectada².

Para el tratamiento de la patología se utilizan en la actualidad dos medicamentos el nifurtimox y el benznidazol que presentan una alta toxicidad y sólo muestran eficacia en la fase aguda de la infección.

A partir de esta realidad se hace necesario el desarrollo de nuevos fármacos que permitan combatir la enfermedad de manera más eficaz.

La hexokinasa es la primera enzima involucrada en la glicólisis de muchos organismos, entre los que se incluyen los agentes etiológicos de la enfermedad de Chagas (*Trypanosoma cruzi*) y la enfermedad del sueño africana (*Trypanosoma brucei*). Recientes estudios han mostrado que análogos de los bisfosfonatos son potentes inhibidores de la hexoquinasa de *T. cruzi*, y pueden representar una nueva diana en la búsqueda de compuestos activos frente a *T. cruzi*³.

La topología molecular a través de las relaciones cuantitativas estructura – actividad, RCEA, es una herramienta que sería útil para el desarrollo de nuevas moléculas con actividad tripanocida. El método de conectividad molecular fue iniciado por Randic⁴ y desarrollado más tarde por Kier y Hall,⁵ a partir del cual es posible identificar diferentes niveles estructurales de una molécula a través de unos valores numéricos conocidos como índices de conectividad o topológicos. De esta manera, cada molécula se asimila a un grafo en donde cada vértice representa un átomo y cada eje las conexiones existentes entre ellos.

A partir del grafo, se construye la matriz topológica cuyos elementos adquieren valores nulos o no en función de las conexiones existentes entre dichos elementos del grafo. El tratamiento posterior de la matriz topológica conduce a los índices de conectividad que caracterizan a ese grafo de

Chagas disease shows three clinical stages, namely acute, silent and chronic stage. Around 30% of the infected patients develop a progressive and serious illness, compromising the heart and digestive system.

An estimate 18 million people are infected and one hundred million have high risk to get the infection in fifteen south-American countries from Mexico to Argentina¹. It's the third tropical disease more important worldwide in morbidity and mortality and it's impossible to tear it apart from poverty so it's an important social and health problem for the infected population².

Nowadays the therapy for this pathology is based on two drugs: nifurtimox and benznidazol, which have a hazardous level of toxicity and are only effective on the acute stage of the disease.

That is the reason why it's necessary to find new drugs against this disease.

Hexokinase is the first enzyme involved in glycolysis in most organisms including the etiological agents of Chagas disease (*Trypanosoma cruzi*) and African sleeping sickness (*Trypanosoma brucei*). Recent studies have shown that bisphosphonates analogues, are potent inhibitors of *T. cruzi* hexokinase, which can represent a novel target to find new compounds actives against *T. cruzi*³.

Molecular topology based QSAR (quantitative structure–activity relationships) is an useful tool which would be able to develop new Trypanosomaticide activity molecules. Molecular connectivity method's pioneer was Randic⁴ and later this method was further developed by Kier y Hall⁵ so that nowadays it is possible to identify different molecular structures by some numeric values known as connectivity (or topological) indices. So each molecule is represented by a graph, where each vertex represents an atom and each axis the connection between atoms (bonds).

From the graph, it is built up the topological matrix whose elements acquire integer values or zero depending on the existence or not of a link or connection between the graph elements, respectively. Thereafter the treatment of the topological matrix drives the connectivity indices of this graph to an easy and fast way. These descriptors have been used successfully to seek topological-mathematical models to select and to design new active compounds in different therapeutic areas and particularly antiparasitic area^{4,5,6,7,8,9}.

una manera rápida y sencilla. Estos descriptores moleculares se han utilizado con éxito en la búsqueda de modelos topológico-matemáticos para la selección y diseño de nuevos compuestos activos en diferentes áreas terapéuticas y en particular en el antiparasitario^{4,5,6,7,8,9}.

En este trabajo, la inhibición de la hexokinasa de *T. cruzi*, TcHK, por un grupo de bisfosfonatos ha sido investigada para obtener un modelo de predicción QSAR usando la topología molecular y el análisis lineal discriminante. Adicionalmente, se ha diseñado una librería molecular a la que se ha aplicado el modelo propuesto para la selección de nuevos compuestos teóricamente bioactivos.

MATERIAL Y MÉTODOS

Compuestos analizados

Para este estudio hemos seleccionado un grupo de 42 bisfosfonatos inhibidores de la TcHK. Las Tablas 1 y 2 muestran las estructuras y la actividad inhibitoria, $CI_{50}(\mu M)$, para cada compuesto a partir de los ensayos *in vitro*, según recoge la bibliografía^{3,10}.

In this work, the inhibition of *T. cruzi* hexokinase, TcHK, of a group of bisphosphonate derivatives was investigated to obtain a QSAR model of prediction using molecular topology and linear discriminant analysis. In addition, a screening molecular was realized to select new compounds with theoretical bioactivity.

MATERIALS AND METHODS

Analysed compounds

In this study we have selected a group of 42 bisphosphonate derivatives inhibitors of TcHK. Tables 1 and 2 show the structures and inhibitory activity expressed as values of $IC_{50}(\mu M)$ obtained for the *in vitro* assays for each compound reported through the papers^{3,10}.

TABLA 1. Estructuras químicas de los compuestos bifosfonatos estudiados en orden decreciente a su actividad inhibitoria frente a TcHK.

TABLE 1. Structures of bisphosphonates studied in order of decreasing potency in TcHK inhibition.

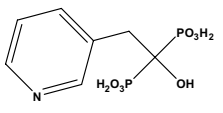
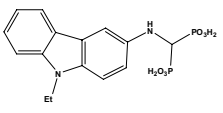
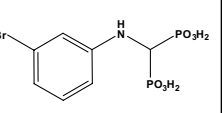
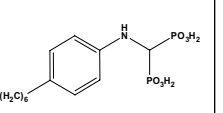
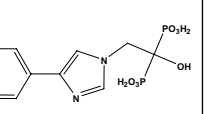
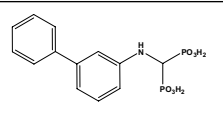
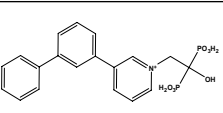
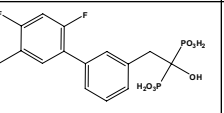
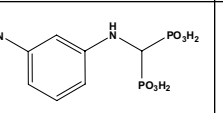
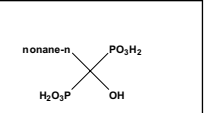
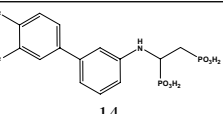
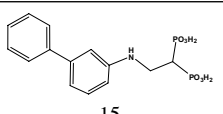
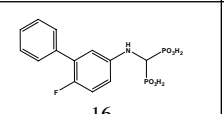
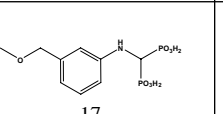
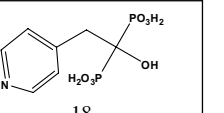
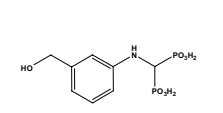
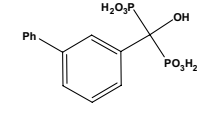
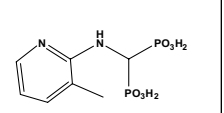
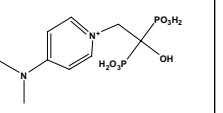
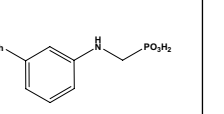
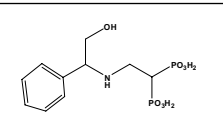
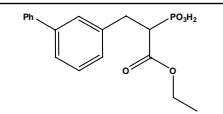
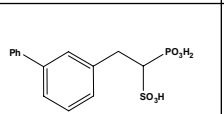
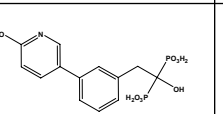
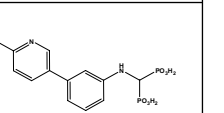
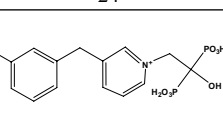
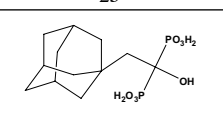
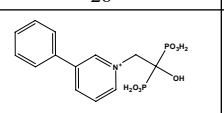
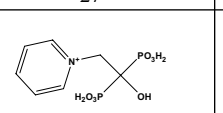
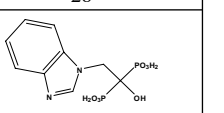
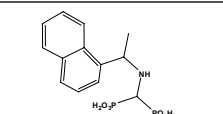
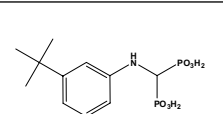
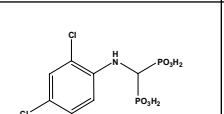
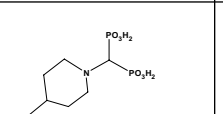
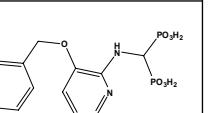
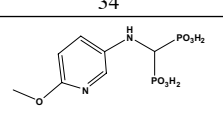
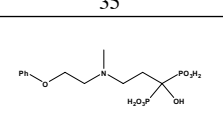
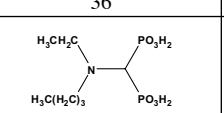
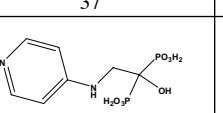
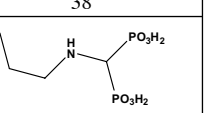
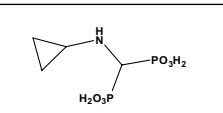
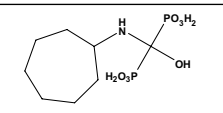
 4	 5	 6	 7	 8
 9	 10	 11	 12	 13
 14	 15	 16	 17	 18
 19	 20	 21	 22	 23
 24	 25	 26	 27	 28
 29	 30	 31	 32	 33
 34	 35	 36	 37	 38
 39	 40	 41	 42	 43
 44	 45			

TABLA 2. Resultados de predicción obtenidos con el análisis lineal discriminante, CI_{50} para TcHK y los derivados bifosfonatos analizados.

TABLE 2. Results of prediction obtained by lineal discriminant analysis with IC_{50} for TcHK and the bisphosphonate derivatives analysed.

compuesto compound	CI_{50} (μM) IC_{50} (μM)	$^1\chi$ $^1\chi$	G_4^v G_4^v	J_2^v J_2^v	J_5^v J_5^v	DF DF	Clas. Class.
Grupo de entrenamiento activo ($CI_{50} < 20 \mu M$) Active group training ($IC_{50} < 20 \mu M$)							
5	0.81	9.847	1.869	0.252	0.055	1.19	A
6	0.95	6.661	0.782	0.187	0.029	2.25	A
7	1.45	9.721	1.278	0.193	0.045	3.28	A
8	1.82	8.500	2.024	0.383	0.045	-1.85	I
9	1.95	8.321	0.959	0.197	0.038	2.23	A
10	2.29	10.761	2.222	0.321	0.045	3.07	A
11	2.29	9.933	1.884	0.326	0.046	0.88	A
12	2.4	7.604	0.776	0.171	0.035	2.41	A
13	2.75	8.561	1.126	0.191	0.040	2.81	A
14	2.75	9.149	1.286	0.206	0.040	3.31	A
15	3.47	8.821	0.967	0.254	0.038	1.04	A
16	3.63	8.738	1.267	0.212	0.039	2.91	A
17	4.07	7.721	1.248	0.230	0.037	1.67	A
18	12.6	6.618	0.813	0.338	0.036	-4.08	I
19	14.8	7.221	0.976	0.230	0.031	1.56	A
Grupo de entrenamiento inactivo ($CI_{50} > 20 \mu M$) Inactive group training ($IC_{50} > 20 \mu M$)							
4	300	6.618	0.804	0.324	0.028	-1.87	I
22	300	7.972	1.906	0.360	0.058	-4.59	I
23	300	6.689	0.808	0.208	0.039	-0.45	I
24	300	8.209	1.149	0.313	0.043	-2.13	I
26	300	8.321	1.149	0.295	0.040	-0.82	I
27	300	9.100	1.495	0.334	0.042	-0.52	I
28	300	8.732	1.365	0.265	0.044	0.55	A
29	300	9.557	2.080	0.403	0.041	-0.43	I
30	300	9.442	1.718	0.172	0.086	-3.47	I
31	300	8.689	1.858	0.338	0.043	-0.34	I
32	300	6.618	1.222	0.368	0.031	-2.83	I
33	300	7.833	1.343	0.321	0.038	-1.14	I
35	300	7.911	1.590	0.224	0.061	-1.93	I
36	300	7.077	0.833	0.218	0.039	-0.30	I
37	300	7.121	2.162	0.255	0.058	-1.26	I
38	300	9.284	1.276	0.268	0.050	-0.55	I
39	300	7.221	1.071	0.258	0.034	0.29	A
40	300	9.512	1.594	0.261	0.068	-3.00	I
42	300	7.118	0.677	0.290	0.030	-1.11	I
43	300	5.693	0.990	0.234	0.053	-4.55	I
Grupo test Test group							
20	34.7	8.232	0.957	0.233	0.030	2.69	A
21	45.7	6.667	0.692	0.349	0.067	-11.10	I
25	300	8.735	2.027	0.392	0.059	-4.83	I
34	300	8.581	2.115	0.305	0.059	-1.83	I
41	300	7.141	2.180	0.241	0.081	-5.51	I
44	300	5.710	0.913	0.233	0.047	-3.53	I
45	300	8.078	1.551	0.180	0.080	-4.33	I

Descriptores moleculares

En este trabajo se han utilizado un grupo de descriptores bien conocidos tales como: índices de conectividad tipo Randić-Kier-Hall hasta orden cuarto (${}^m\chi_t$, ${}^m\chi_t^v$) (11),(12), índices de carga topológica, TCI, hasta orden quinto, (J_m , G_m , J_m^v , G_m^v)¹³, cocientes y diferencias entre índices de conectividad de valencia y de no valencia (${}^mC_t = {}^m\chi_t / {}^m\chi_t^v$ and ${}^mD_t = {}^m\chi_t - {}^m\chi_t^v$). En total, cada compuesto fue caracterizado con una serie de 52 índices. Todos los descriptores usados en este trabajo han sido obtenidos gracias al programa Desmol11 (disponible vía e-mail)¹⁴.

Algoritmo QSAR: Análisis lineal discriminante

El objetivo del análisis lineal discriminante, ALD, que es considerado como algoritmo heurístico capaz de distinguir entre dos o más categorías u objetos, es el de encontrar una función lineal capaz de diferenciar entre compuestos activos e inactivos por los diferentes valores de sus descriptores. Dos series de compuestos han sido considerados para el análisis: La primera formada por compuestos con probada actividad inhibitoria (en nuestro caso, todos los componentes con $CI_{50} < 20\mu M$) y la segunda por compuestos inactivos ($CI_{50} > 20\mu M$). La capacidad discriminante fue evaluada por el porcentaje de clasificación correcta de cada grupo. El análisis ALD se realizó con la ayuda del paquete estadístico BMDP 7M¹⁵. La selección de descriptores se hizo en base al parámetro de F-Snedecor, y el criterio de clasificación a través de la distancia más corta de Mahalanobis (distancia de cada caso a la media de todos los casos utilizados en la ecuación de regresión).

A partir de la función discriminante seleccionada, se puede construir el correspondiente diagrama de distribución de la actividad inhibitoria, PDD. Este diagrama es útil para establecer los intervalos de la función discriminante en donde la expectancia (probabilidad), E, de encontrar compuestos activos es máxima. Los diagramas PDD son histogramas en los que la expectancia E se sitúa en el eje de ordenadas y los valores de la función discriminante, FD, en abscisas. Para un intervalo arbitrario de FD, se define la expectancia de actividad como: $E_a = a/(i+1)$; "a" representa el número de compuestos activos en

Molecular descriptors

A set of well-known topological descriptors was used in this work: Subgraph Randić-Kier-Hall like indices up to the fourth order (${}^m\chi_t$, ${}^m\chi_t^v$) (11),(12), topological charge indices, TCI, up to the fifth order, (J_m , G_m , J_m^v , G_m^v)¹³, quotients and differences between valence and non-valence connectivity indices (${}^mC_t = {}^m\chi_t / {}^m\chi_t^v$ and ${}^mD_t = {}^m\chi_t - {}^m\chi_t^v$). Each compound was characterized by a set of 52 descriptors. All descriptors used in this work were obtained with the aid of the Desmol11 program (available by e-mail request)¹⁴.

QSAR algorithms: Linear discriminant analysis

The objective of the linear discriminant analysis (LDA), which is considered as a heuristic algorithm able to distinguish between two or more categories or objects, is to find a linear function capable to discriminate between the active and inactive compounds as for their different descriptor values. Two sets of compounds, the first with a proven inhibitory activity (in our case, all the compounds with $IC_{50} < 20\mu M$) and the second comprised of inactive compounds ($IC_{50} > 20\mu M$) were considered for the analysis. The discriminant ability was tested by the percentage of correct classifications into each group. LDA was performed by using the BMDP 7M package¹⁵. The selection of the descriptors was based on the F-Snedecor parameter, and the classification criteria was the shortest Mahalanobis distance (distance of each case to the mean of all cases used in the regression equation).

From the selected discriminant function, the corresponding distribution diagram of inhibitory TeHK activity, PDD, was drawn. This diagram was pictured just to establish the intervals of the discriminant function in which the expectancy, E, of finding active compounds is maximum. PDDs are histogram-like plots of connectivity functions in which expectancies appear on the ordinate axis. For an arbitrary interval of values of a given function, we define the expectancy of activity as: $E_a = a/(i+1)$; where "a" is the number of active compounds in the interval divided by the total number of active compounds, and "i" is the number of inactive compounds in the interval divided by the total number of inactive compounds. The expectancy of inactivity is defined in a sym-

el intervalo dividido por el número total de compuestos activos, e "i" el número de compuestos inactivos en el intervalo dividido por el número total de compuestos inactivos. La expectancia de inactividad viene definida de una manera simétrica como $E_i = i/(a+1)$. Esta representación nos da una buena visualización de las regiones de mínimo solapamiento, y permite la selección de intervalos de FD en los que la probabilidad de encontrar compuestos activos es máxima¹⁶.

RESULTADOS Y DISCUSIÓN

Para obtener la función discriminante, FD, aplicamos el análisis ALD a un grupo de entrenamiento formado por 35 compuestos y, para validarla, a un grupo test de 7 compuestos. La serie de entrenamiento está formada por dos subgrupos: grupo activo (compuestos con valores de $IC_{50} < 20 \mu M$) y grupo inactivo (compuestos con $IC_{50} > 20 \mu M$). La serie test la forman compuestos con valores de $IC_{50} > 20 \mu M$ y seleccionados aleatoriamente del grupo de inactivos.

La función discriminante seleccionada fue:
 $DF = 5.261 + 1.013 \chi + 3.011 G_4^v - 32.49 J_2^v - 207.4 J_5^v$ Eq. 1
 $N=35 \quad F=5.90 \quad \lambda(\text{parámetro de Wilks}) = 0.556$

A partir de aquí, un determinado compuesto será seleccionado como un potencial inhibidor de TcHK si $DF > 0$, en caso contrario será clasificado como "inactivo". La matriz de clasificación es muy significativa para la serie de entrenamiento (86.7% de predicción correcta para el grupo activo, 13 de los 15 compuestos son clasificados correctamente, y un 90% para el grupo inactivo, 18 de un total de 20 (ver Tabla 2)).

En la Eq. 1 aparecen descriptores topológicos que evalúan, por una parte, aspectos topológicos de cada compuesto tales como la ramificación molecular, χ , y, por otra, las distribuciones de la carga intramolecular a través de índices de carga G_4^v , J_2^v y J_5^v (ver tabla 2).

La Figura 1 muestra el diagrama de distribución de la actividad inhibitoria obtenido con la función FD (las columnas blancas y negras representan a los grupos inactivo y activo, respectivamente). La región de FD con el mínimo solapamiento para los compuestos con teórica actividad inhibitoria aparece cuando $DF > 0$ y $DF < 4$, por lo que la expectancia o probabilidad de mayor actividad estará en ese intervalo.

metrical way, as $E_i = i/(a+1)$. This representation provides a good visualization of the regions of minimum overlap, and allows the selection of regions in which the probability of finding active compounds reaches a maximum¹⁶.

RESULTS AND DISCUSSION

To obtain the discriminant function, we apply the LDA to a training group comprised of 35 compounds and to validate it, to a 7 compounds test group. The training series is comprised by two subgroups: an active group (compounds with values of $IC_{50} < 20 \mu M$) and an inactive group (compounds with $IC_{50} > 20 \mu M$). The test series is arranged by compounds with values of $IC_{50} > 20 \mu M$ and randomly selected from the inactive group.

The discriminant function selected was:
 $DF = 5.261 + 1.013 \chi + 3.011 G_4^v - 32.49 J_2^v - 207.4 J_5^v$ Eq. 1
 $N=35 \quad F=5.90 \quad \lambda(\text{Wilks' lambda}) = 0.556$

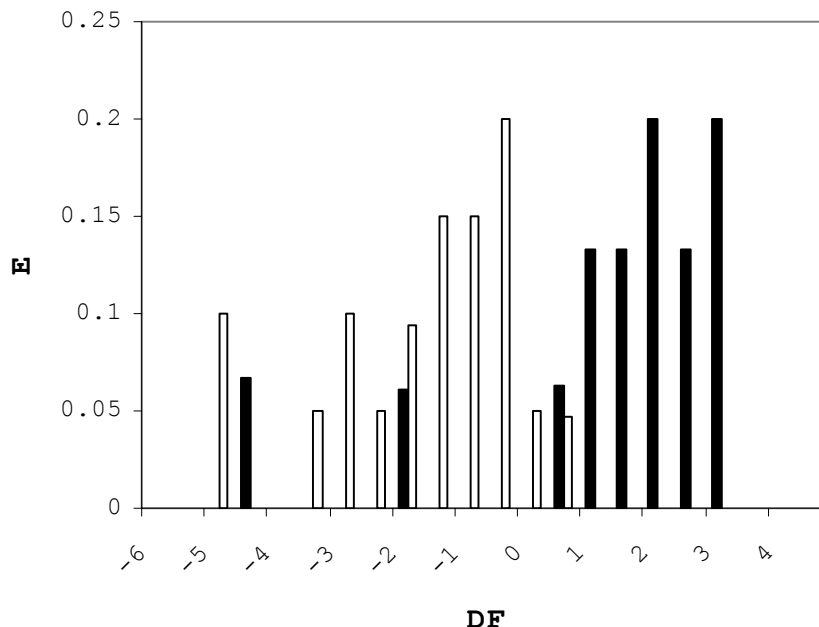
From here, a given compound will be selected as a potential inhibitor TcHK if $DF > 0$, otherwise it is classified as "inactive". The classification matrix is very significantly for the training set (86.7% of correct prediction for the active group, 13 out of 15 correctly classified, and 90.0% for the inactive group, 18 correct out of 20 (see Table 2)).

In Eq. 1 appear topological descriptors that evaluate, on the one hand, the topological aspects of each compound as for instance the molecular branching, χ , and, on the other, the distributions of the intramolecular charge through the charge indices G_4^v , J_2^v and J_5^v (see Table 2).

Figure 1 depicts the inhibitory activity distribution diagram obtained with the function DF (white and black bars represent inactive and active groups, respectively). It is apparent that the regions with minimum overlap for the compounds with theoretical inhibitory activity occur for $DF > 0$ and $DF < 4$, so the highest expectation of activity occurs in these intervals.

FIGURA 1. Diagrama de distribución farmacológica para la actividad inhibitoria frente a TCHK al graficar la expectancia (E) frente a la función FD (Eq. (1)) (las barras negras representan los compuestos con $IC_{50} < 20\mu M$ y la blancas, los compuestos con $IC_{50} > 20\mu M$).

FIGURE 1. Pharmacological distribution diagram for inhibitory activity against TCHK by plotting expectancy (E) versus DF (Eq. (1)) function (the black bars represent the compounds with $IC_{50} < 20\mu M$ and the white bars, the compounds with $IC_{50} > 20\mu M$).



Una forma sencilla de evaluar la calidad de la función discriminante seleccionada es aplicarla a un grupo test externo. En nuestro caso utilizamos 7 compuestos no usados en el análisis discriminante y todos ellos con valores de $IC_{50} > 20\mu M$. La tabla 2 recoge los resultados obtenidos para cada compuesto. Como puede observarse, todos los compuestos excepto el n° 20, son clasificados correctamente como inactivos, $DF < 0$.

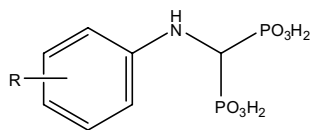
Basándonos en el modelo topológico seleccionado, lo aplicamos a una librería molecular de derivados bisfosfonatos construida con la ayuda de la base Scifinder Scholar, con la intención de seleccionar nuevos compuestos bisfosfonatos activos. La tabla 3 muestra la estructura base utilizada y algunos de los compuestos seleccionados por el modelo.

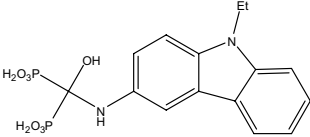
An easy way to evaluate the quality of the selected discriminant function is to apply it into an external test group. In our case we used 7 compounds which haven't been included in the discriminant analysis, all of them with values of $IC_{50} > 20\mu M$. Table 2 shows the results obtained for each compound. As shown, all the compounds, except the number 20, are correctly classified as inactive, $DF < 0$.

Based upon the selected topological model, we carried out a virtual screening by the Scifinder Scholar database, with the aim of selecting new bisphosphonate derivatives showing a better activity. Table 3 shows the base structure (scaffold) used and some of the compounds selected by the model.

TABLA 3. Rastreo computacional aplicado a derivados bifosfonatos sacados de la base Scifinder Scholar y aplicando el modelo topológico obtenido con la actividad inhibitoria frente a TcHK.

TABLE 3. Computational screening applied to bisphosphonate analogues obtained of Scifinder Scholar base and using the inhibitory TcHK model selected by molecular topology.



CAS N°	R	FD
866108-65-8	3-amino	2.25
866108-36-3	2,4-difluoro	3.96
866108-30-7	2-dimethylamino	0.56
851915-98-5	4-O-methyl	0.92
851915-96-3	4-NO ₂	1.39
847738-17-4	4-CH ₂ -Ph	2.79
690232-12-3	3,5-dimethyl	1.00
652989-94-1	4-Ph	1.52
652989-91-8	4-O-nbutyl	0.35
918500-41-1		3.09

De los resultados obtenidos en la tabla 3 se pueden realizar algunos comentarios interesantes:

Some interesting comments can be stated from results in Table 3:

- La actividad inhibitoria del TcHK se ve reforzada al introducir fragmentos moleculares como 4-O-methyl con FD(N° 851915-98-5)=0.92, 4-NO₂ con FD(N° 851915-96-3)=1.39, 4-CH₂-Ph con FD(N° 847738-17-4)=2.79, 4-Ph con FD(N° 652989-94-1)=1.52 o 4-O-nbutyl con FD(N° 652989-91-8)=0.35 en la posición *para* del anillo aromático.
- El compuesto n° 35 incrementa significativamente la actividad al sustituir el fragmento molecular t-butilo por amino FD(N° 866108-65-8)=2.25. Resultados similares se observa con el compuesto n° 36 al intercambiar el Cl por F FD(N° 866108-36-3)=3.96.

- The inhibitory TcHK activity is enhanced by the addition of molecular fragments such as 4-O-methyl with DF(N°851915-98-5)=0.92, 4-NO₂ with DF(N°851915-96-3)=1.39, 4-CH₂-Ph with DF(N°847738-17-4)=2.79, 4-Ph with DF(N°652989-94-1)=1.52 or 4-O-nbutyl with DF(N°652989-91-8)=0.35 on position *para* of the aromatic ring.
- Compound n° 35 increases significantly the activity if molecular fragment t-butyl is substituted by amine DF(N°866108-65-8)=2.25. Similar results are obtained with the compound n° 36 on changing Cl by F DF(N°866108-36-3)=3.96.

Algunos de ellos, por ejemplo N°918500-41-1 y N°652989-94-1 han mostrado diferente actividad antimicrobiana¹⁷, otros como el N°866108-65-8, N°866108-36-3, N°866108-30-7 y N°847738-17-4 tienen actividad antiprotozoaria frente a *T. brucei*¹⁸.

Some of them, for example N°918500-41-1 and N°652989-94-1, showed different antimicrobial activity¹⁷, others such as N°866108-65-8, N°866108-36-3, N°866108-30-7 and N°847738-17-4 have antiprotozoal activity against *T. brucei*¹⁸. The compound N°690232-12-3 has herbicidal activity¹⁹. The incorporation of activity against *T. cruzi* to these compounds would be really

El compuesto N°690232-12-3 presenta actividad herbicida¹⁹. La incorporación de actividad frente a *T. cruzi* a estos compuestos sería muy interesante ya que se ampliaría el campo terapéutico de posible utilización de los mismos.

Los resultados obtenidos en este trabajo necesitan, lógicamente, ser corroborados con los correspondientes ensayos de inhibición de la actividad TCHK, lo que nos permitiría la validación o evaluación del modelo propuesto y servir como herramienta útil para la búsqueda de nuevos compuestos activos frente a *T. cruzi*.

CONCLUSIONES

La topología molecular se ha utilizado satisfactoriamente en la búsqueda de un modelo QSAR para la predicción de la actividad inhibitoria frente a TCHK de un grupo de compuestos bisfosfonatos. Todos los descriptores moleculares usados son grafo-teóricos. El modelo matemático seleccionado en este trabajo retiene características estructurales que identifica la propiedad correlacionada, CI_{50} , y, por tanto, puede ser utilizado en la búsqueda de nuevos compuestos activos a través de bases moleculares. Se ha construido una librería molecular formada por varios centenares de derivados bisfosfonatos para la optimización de la actividad inhibitoria frente a la hexokinasa de *T. cruzi*. Se han obtenido interesantes mejoras de esta actividad.

AGRADECIMIENTOS

Agradecemos al Fondo de Investigación Sanitaria, Ministerio de Sanidad, España (proyecto SAF2005-PI052128) y al Master Internacional en Enfermedades Parasitarias Tropicales, Universitat de Valencia, por la financiación de este estudio.

interesting as it would enlarge their therapeutic field and use.

These suggestive results need to be corroborated with the corresponding inhibitory TCHK activity assays, which should allow the validation or evaluation of the model proposed and serve as an useful tool for the search of novel compounds with a higher activity against *T. cruzi*.

CONCLUSIONS

Molecular topology has been successfully used to find a QSAR model predicting the inhibitory activity against TCHK of a group of bisphosphonate compounds. All the molecular descriptors used are graph-theoretical ones. The mathematical model employed in this work retains the main structural features that involve the correlated property, IC_{50} , and, therefore, can be applicable to the search of new active compounds by virtual screening throughout databases. We have built up a virtual library with several hundreds of bisphosphonate derivatives for virtually seeking and optimizing the inhibitory activity against *T. cruzi* hexokinase. Interesting improvements in the activity have been obtained.

ACKNOWLEDGEMENTS

We thank the Fondo de Investigación Sanitaria, Ministerio de Sanidad, Spain (project SAF2005-PI052128) and Master Internacional en Enfermedades Parasitarias Tropicales, Universitat de Valencia, for support of this work.

BIBLIOGRAFÍA/BIBLIOGRAPHY

1. Organization WH. Burdens and Trends in Chagas disease. [cited]; Available from: www.who.int/ctd/chagas/burdens
2. Montes G, Ponce C, Ponce E. Insecticidal Paintand Fumigant Canisters for Chagas Disease Control: Community acceptance in Honduras Pan American Journal of Public Health. 1999;6:311 – 20.
3. Hudock MP, Sanz-Rodriguez CE, Song Y, Chan JM, Zhang Y, Odeh S, et al. Inhibition of Trypanosoma cruzi hexokinase by bisphosphonates. J Med Chem. 2006 Jan 12;49(1):215-23.
4. Galvez J, Garcia-Domenech R, de Julian-Ortiz JV, Soler R. Topological Approach to Drug Design. [Erratum to document cited in CA122:177672]. Journal of Chemical Information and Computer Sciences. 1995;35(5):938.

5. Gozalbes R, Brun-Pascaud M, Garcia-Domenech R, Galvez J, Girard P-M, Doucet J-P, et al. Anti-Toxoplasma activities of 24 quinolones and fluoroquinolones in vitro: prediction of activity by molecular topology and virtual computational techniques. *Antimicrobial Agents and Chemotherapy*. 2000;44(10):2771-6.
6. Galvez J, de Julian-Ortiz JV, Garcia-Domenech R. Application of molecular topology to the prediction of potency and selection of novel insecticides active against malaria vectors. *Theochem*. 2005;727(1-3):107-13.
7. Mahmoudi N, de Julian-Ortiz J-V, Ciceron L, Galvez J, Mazier D, Danis M, et al. Identification of new antimalarial drugs by linear discriminant analysis and topological virtual screening. *Journal of Antimicrobial Chemotherapy*. 2006;57(3):489-97.
8. Garcia-Domenech R, Galvez J, de Julian-Ortiz JV, Pogliani L. Some New Trends in Chemical Graph Theory. *Chem. Rev*. 2008;108(3):1127-69.
9. Mahmoudi N, Garcia-Domenech R, Galvez J, Farhati K, Franetich J-F, Sauerwein R, et al. New active drugs against liver stages of Plasmodium predicted by molecular topology. *Antimicrobial Agents and Chemotherapy* 2008;52(4):1215-20.
10. Racagni GE, Machado de Domenech EE. Characterization of Trypanosoma cruzi hexokinase. *Mol Biochem Parasitol*. 1983 Oct;9(2):181-8.
11. Kier LB, Murray WJ, Randic M, Hall LH. Molecular connectivity V: connectivity series concept applied to density. *J Pharm Sci*. 1976 Aug;65(8):1226-30.
12. Kier LB, Hall LH. General definition of valence delta-values for molecular connectivity. *J Pharm Sci*. 1983 Oct;72(10):1170-3.
13. Galvez J, Garcia R, Salabert MT, Soler R. Charge Indexes. New Topological Descriptors. *Journal of Chemical Information and Computer Sciences*.1994;34(3):520-5.
14. Desmol11 software. Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular. Facultad de Farmacia Universitat de Valencia, Spain.
15. Dixon WJ. BMDP Statistical software. Berkeley: University of California; 1990.
16. Galvez J, Garcia-Domenech R, de Gregorio Alapont C, de Julian-Ortiz JV, Popa L. Pharmacological distribution diagrams: a tool for de novo drug design. *Journal of Molecular Graphics*. 1996;14(5):272-6.
17. Leon A, Liu L, Yang Y, Hudock MP, Hall P, Yin F, et al. Isoprenoid biosynthesis as a drug target: bisphosphonate inhibition of Escherichia coli K12 growth and synergistic effects of fosmidomycin. *J Med Chem*. 2006 Dec 14;49(25):7331-41.
18. Kotsikorou E, Song Y, Chan JMW, Faelens S, Tovian Z, Broderick E, et al. Bisphosphonate Inhibition of the Exopolyphosphatase Activity of the Trypanosoma brucei Soluble Vacuolar Pyrophosphatase. *J. Med. Chem*. 2005;48(19):6128-39.
19. Obojska A, Berlicki L, Kafarski P, Lejczak B, Chicca M, Forlani G. Herbicidal pyridyl derivatives of aminomethylene-bisphosphonic acid inhibit plant glutamine synthetase. *Journal of Agricultural and Food Chemistry*. 2004;52(11):3337-44.