New Antitrichomonal Drug-like Chemicals Selected by Bond (Edge)-Based TOMOCOMD-CARDD Descriptors

ALFREDO MENESES-MARCEL,[1,2] OSCAR M. RIVERA-BORROTO,[1,3] YOVANI MARRERO-PONCE,[1,4] ALINA MONTERO,[1] YANETSY MACHADO TUGORES,[1] JOSÉ ANTONIO ESCARIO,[2] ALICIA GÓMEZ BARRIO,[2] DAVID MONTERO PEREIRA,[2] JUAN JOSÉ NOGAL,[2] VLADIMIR V. KOUZNETSOV,[5] CRISTIAN OCHOA PUENTES,[5] ARNOLD R. BOHÓRQUEZ,[5] RICARDO GRAU,[3] FRANCISCO TORRENS,[4] FROYLÁN IBARRA-VELARDE,[6] and VICENTE J. ARÁN[7]

Bond-based quadratic indices, new TOMOCOMD-CARDD molecular descriptors, and linear discriminant analysis (LDA) were used to discover novel lead trichomonacidals. The obtained LDA-based quantitative structure-activity relationships (QSAR) models, using nonstochastic and stochastic indices, were able to classify correctly 87.91% (87.50%) and 89.01% (84.38%) of the chemicals in training (test) sets, respectively. They showed large Matthews correlation coefficients of 0.75 (0.71) and 0.78 (0.65) for the training (test) sets, correspondingly. Later, both models were applied to the virtual screening of 21 chemicals to find new lead antitrichomonal agents. Predictions agreed with experimental results to a great extent because a correct classification for both models of 95.24% (20 of 21) of the chemicals was obtained. Of the 21 compounds that were screened and synthesized, 2 molecules (chemicals G-1, UC-245) showed high to moderate cytocidal activity at the concentration of 10 μg/ml, another 2 compounds (G-0 and CRIS-148) showed high cytocidal activity only at the concentration of 100 μg/ml, and the remaining chemicals (from CRIS-105 to CRIS-153, except CRIS-148) were inactive at these assayed concentrations. Finally, the best candidate, G-1 (cytocidal activity of 100% at 10 μg/ml) was in vivo assayed in ovariectomized Wistar rats achieving promising results as a trichomonacidal drug-like compound. (*Journal of Biomolecular Screening* 2008:785-794).

Key words: bond-based TOMOCOMD-CARDD quadratic indices, LDA-based QSAR model, virtual screening, lead generation, trichomonacidal, cytostatic and cytocidal activity

INTRODUCTION

Trichomonas vaginalis (Tv) is the causative agent of the most common, nonviral, sexually transmitted disease.[1] In 1995, the World Health Organization estimated the number of adults with trichomoniasis at 170 million worldwide, more than the numbers for gonorrhea, syphilis, and chlamydia combined.[2]

In 1959, a nitroimidazole derivative of a *Streptomyces* antibiotic, azomycin, was found to be highly effective in the systemic treatment of trichomoniasis.[3] This derivative was a,b-hydroxyethyl-2-methyl-5-nitroimidazole, commonly referred to as metronidazole (MTZ) and marketed under the trade name Flagyl.

The recommended MTZ regimen results in cure rates of approximately 95%.[4] In addition, it is remarkably safe compared with the most toxic antiprotozoal products.[5] However, resistance to MTZ has been proven to be geographically widely distributed, and no clustering or temporal trends in patients have been observed.[6] Also, in patients who do not respond to high-dose MTZ therapy, a variety of regimens have

[1]Unit of Computer-Aided Molecular "Biosilico" Discovery and Bioinformatic Research (CAMD-BIR Unit), Faculty of Chemistry-Pharmacy, Central University of Las Villas, Villa Clara, Cuba.
[2]Departamento de Parasitología, Facultad de Farmacia, Universidad Complutense de Madrid, Madrid, Spain.
[3]Bioinformatics Group, Informatics Research Center (CEI), Faculty of Mathematics, Physics and Computer Science, Central University of Las Villas, Villa Clara, Cuba.
[4]Institut Universitari de Ciència Molecular, Universitat de València, Edifici d'Instituts de Paterna, València, Spain.
[5]Laboratorio de Química Orgánica y Biomolecular, Escuela de Química, Universidad Industrial de Santander, Bucaramanga, Colombia.
[6]Department of Parasitology, Faculty of Veterinarian Medicinal and Zootecnic, Universidad Nacional Autónoma de México, Mexico City, Mexico.
[7]Instituto de Química Médica, CSIC, Madrid, Spain.

Received Jun 25, 2007, and in revised form Apr 10, 2008. Accepted for publication Apr 29, 2008.
Journal of Biomolecular Screening 13(8); 2008
DOI: 10.1177/1087057108323122

been evaluated for possible effectiveness, with rare or only occasional success. These include zinc sulfate, povidone-iodine douche, arsenicals, nonoxynol-9 cream, mebendazole, albendazole, furazolidone, and rifabutin.[7–12] These agents, although they demonstrate considerable in vitro activity, have been clinically disappointing. Paromomycin was previously reported to be useful in the management of resistant trichomoniasis. It was used fairly effectively (cure rate, 58%) in 12 patients and remains an important option; however, local side effects were considerable and can be quite severe.[13, 14]

Currently, it is clear that new trichomonacidals are needed to combat resistant Tv organisms; therefore, the main objective of this work was to use TOMOCOMD-CARDD nonstochastic and stochastic bond-based quadratic indices[15] to generate predictive linear discriminant analysis (LDA)–based quantitative structure-activity relationships (QSAR) models enabling the selection of new hits and lead drug-like compounds with antitrichomonal activity. The in vitro and in vivo evaluation of a new lead series of heterocyclic compounds with antitrichomonal activity is also presented.

MATERIALS AND METHODS

TOMOCOMD-CARDD approach

TOMOCOMD (TOpological MOlecular COMputer Design) is an interactive program for molecular design and bioinformatic research.[16] It is composed of 4 subprograms; each 1 of them allows drawing the structures (drawing mode) and calculating molecular 2D/3D (calculation mode) descriptors. In the present report, we outline salient features concerned with only 1 of its subprograms, CARDD, and with the calculation of nonstochastic and stochastic 2D bond-based quadratic indices.

Computational strategies

The main steps for the application of present method in QSAR and drug design can be briefly summarized in the following set of steps: (1) Draw the molecular pseudographs for each molecule of the data set, using the software drawing mode. (2) Use appropriated atomic properties to weight and differentiate the molecular bonds. In this study, the properties are some of those previously proposed for the calculation of the DRAGON descriptors, that is, atomic mass (M), atomic polarizability (P), atomic Mullinken electronegativity (K), van der Waals atomic volume (V).[17] The values of these atomic labels are shown in **Supplemental Table 1** (**Table SM1**). **(See Supplemental Table 1 online at http://jbx.sagepub.com/supplemental/.)** (3) Compute the total and local (bond and bond-type) nonstochastic and stochastic quadratic indices. (4) Find a QSAR equation by using LDA with quadratic indices as independent variables. In this sense, we can find a quantitative relation between an activity **A** and the linear indices having, for instance, the following appearance, $\mathbf{A} = a_0\boldsymbol{q}_0(\overline{w}) + a_1\boldsymbol{q}_1(\overline{w}) + a_2\boldsymbol{q}_2(\overline{w}) + \ldots + a_k\boldsymbol{q}_k(\overline{w}) +$ c, where **A** is the measured activity, $\boldsymbol{q}_k(\overline{w})$ are the kth total bond-based quadratic indices, the $\boldsymbol{a}_k s$ are the coefficients obtained by the linear descriminant analysis, and $\overline{w}$ represent the molecular vector. (5) Test the robustness and predictive power of the QSAR equation by using internal (cross-validation) and external (using a test set and an external predicting set) validation techniques. And (6) apply the obtained LDA-based QSAR models as cheminformatic tool for identifying leads through ligand-based virtual screening–drug discovery process.

The bond-based TOMOCOMD-CARDD descriptors computed in this study were the following:

(1) $k^{\text{th}}(k = \overline{0, 5})$ total nonstochastic bond-based quadratic indices not considering and considering H-atoms in the molecular graph (G) [$\boldsymbol{q}_k(\overline{w})$ and $\boldsymbol{q}_k^H(\overline{w})$, respectively].

(2) $k^{\text{th}}(k = \overline{0, 5})$ total stochastic bond-based quadratic indices not considering and considering H-atoms in the molecular graph (G) [${}^s\boldsymbol{q}_k(\overline{w})$ and ${}^s\boldsymbol{q}_k^H(\overline{w})$, respectively].

(3) $k^{\text{th}}(k = \overline{0, 5})$ bond-type local (group = heteroatoms: S, N, O) nonstochastic quadratic indices not considering and considering H-atoms in the molecular graph (G) [$\boldsymbol{q}_{kL}(\overline{w}_E)$ and $\boldsymbol{q}_{kL}^H(\overline{w}_E)$, correspondingly]. These local descriptors are putative atomic charges, dipole moment, and H-bonding acceptors.

(4) $k^{\text{th}}(k = \overline{0, 5})$ bond-type local (group = heteroatoms: S, N, O) stochastic quadratic indices not considering and considering H-atoms in the molecular graph (G) [${}^s\boldsymbol{q}_{kL}(\overline{w}_E)$, and ${}^s\boldsymbol{q}_{kL}^H(\overline{w}_E)$, correspondingly]. These local descriptors are putative atomic charges, dipole moment, and H-bonding acceptors.

Database selection

Taking into account that the most critical aspect in the construction of a training data set is the molecular diversity of the included compounds, we selected a group of 123 organic chemicals having as much structural variability as possible. The 50 antitrichomonals considered in this study are representative of families with diverse structural patterns and action modes. **Figure 1** shows a representative sample of such active compounds. On the other hand, 73 compounds having different clinical uses were selected for the set of inactive compounds, through a random selection, also guaranteeing a great structural variability. All these chemicals were taken from the Negwer handbook[18] and *Merck Index*,[19] where their names, synonyms, and structural formulas can be found. From these 123 chemicals, 91 were chosen at random

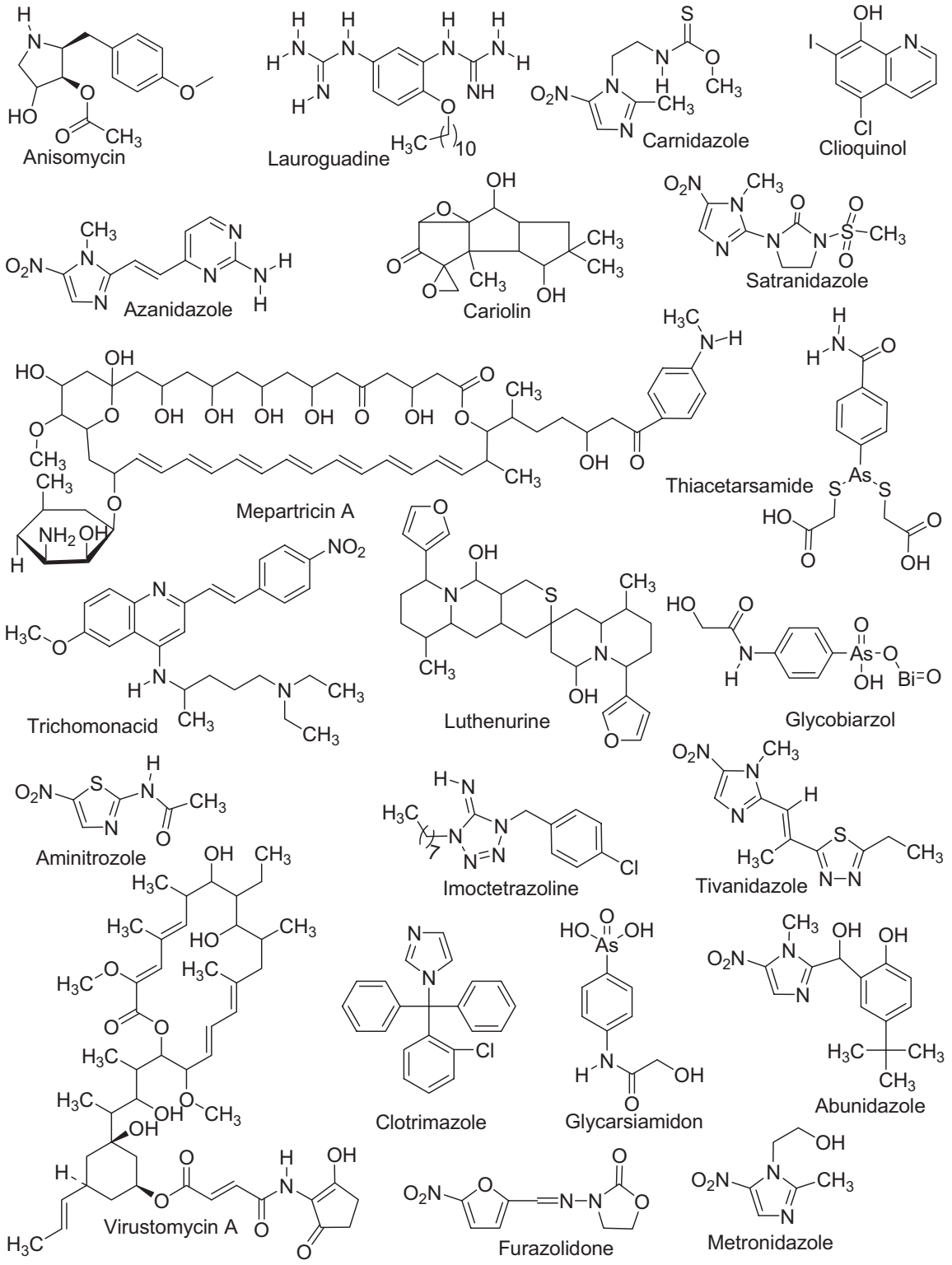


FIG. 1. Random sample of the molecular families of trichomonacidal agents studied here.

to form the training set, 40 of them active and 51 inactive. The remaining subseries consisting of 10 trichomonacidals and 22 nontrichomonacidals were prepared as test sets for the external validation of the models (32 chemicals). These compounds were never used in the development of the classification models.

Data analysis and processing: linear discriminant analysis

The discriminant functions were obtained by using the LDA[20–22] as implemented in STATISTICA (version 6.0, StatSoft Inc., Tulsa, OK). The default parameters of this program were used in the development of the model. Forward stepwise was fixed as the strategy for variable selection. The quality of the models were determined by examining Wilks' λ parameter (*U*-statistic), squared Mahalanobis distance (D^2), Fisher ratio (F), and the corresponding *p*-level (p(F)) as well as the percentage of good classification in the training and test sets. Models with a proportion between the number of cases and variables in the equation lower than 5 were rejected.[20]

By using the models, 1 compound can then be classified as either active, if $\Delta P\% > 0$, being $\Delta P\% = [P(\text{Active}) - P(\text{Inactive})] \times 100$ or inactive otherwise. P(Active) and P(Inactive) are the probabilities with which the equations classify a compound as active and inactive, respectively.

Also a leave-group-out (LGO) cross-validation strategy was carried out. In this case, 10% of the data set was used as group size, that is, groups including 10% of the training data set were left out and predicted by the model based on the remaining 90%. This process was carried out 10 times on 10 unique independent subsets.[23]

Finally, the calculation of percentages of global good classification (accuracy), sensibility, specificity (also known as "hit rate"), false positive rate (also known as "false alarm rate"), and Matthews correlation coefficient (*C*) in the training and test (predicting) sets permitted the assessment of the model.[24]

Determination of in vitro trichomonacidal activity

The biological activity was assayed on TvJH31A #4 Ref. No. 30326 (ATCC, Bethesda, MD) in modified Diamond medium supplemented with equine serum and grown at 37°C (5% CO_2). The compounds were added to the cultures at several concentrations (100, 10, and 1 μg/ml) after 6 h of the seeding (0 h). Viable protozoa were assessed at 24 and 48 h after incubation at 37°C by using the Neubauer chamber. MTZ (Sigma-Aldrich SA, Madrid, Spain) was used as reference drug at concentrations of 2, 1, and 0.5 μg/ml. Cytocidal and cytostatic activities were determined by calculation of percentages of cytocidal (%C) and cytostatic activities (%CA), in relation to controls as previously reported.[25, 26]

Table 1. Infection with *Trichomonas vaginalis* (Tv), Treatment with G-1, and Diagnosis of the Rats

	Day								
	0	*2–3*	*6*		*7*		*8*		*9; 11; 13; 15*
Group	*Estrogen.*	*Infection*	*Diag*	*T 1*	*Diag*	*T 2*	*Diag*	*T 3*	*Diag*
I	Estradiol 10 mg/kg	Tv	1	G-1	2	G-1	3	G-1	4–7
II				M		M		M	
III				—		—		—	
IV				Mtz		Mtz		Mtz	

T, treatment number; Estrogen., estrogenization; Diag, diagnostic number; Tv, 5×10^6 Tv/ml/intravaginal inoculation 200 μl; M, Migliol; G-1, 2-bromo-5-(2-bromo-2-nitrovinil)furano/0.125%; Mtz, metronidazole.

Determination of in vivo trichomonacidal activity

In this study, we examined the protective efficacy of G-1 in ovariectomized rats and dealt with estradiol.[27] The excipient used to dissolve the active principle G-1 was Migliol® 810 N (Degussa-Huls AG, Sarnia, Ontario, Canada). The taking of samples for the diagnosis of the infection was carried out previous to the intravaginal application of the 1st dose of G-1, which allowed us to know the number of rats infected before the 1st treatment. On days 7 and 8 a similar procedure was carried out, which allowed us to evaluate the effect of the 1st and 2nd doses (see **Table 1**).

RESULTS AND DISCUSSION

Development and validation of the discriminant functions

By making use of the LDA technique implemented in the STATISTICA software the following linear models were obtained, in which total as well as local nonstochastic and stochastic bond-based quadratic indices were used as independent variables:

$$\begin{aligned} Class = {} & -4.92 - 1.40 \times 10^{-3M} \boldsymbol{q}_1(\overline{w}) + 0{,}22^{P} \boldsymbol{q}_0^H(\overline{w}) \\ & + 2.01 \times 10^{-2P} \boldsymbol{q}_1(\overline{w}) + -1.44 \times 10^{3V} \boldsymbol{q}_{0L}(\overline{w}_E) \\ & + 0.13^{E} \boldsymbol{q}_{0L}^H(\overline{w}_E) + 5.82 \times 10^{-2E} \boldsymbol{q}_{1L}^H(\overline{w}_E) \qquad (1) \\ N = 91; \quad \lambda = 0.44; \quad & D^2 = 5.02; \quad F(6.84) = 17.71; \quad p < 0.05 \end{aligned}$$

$$\begin{aligned} Class = {} & -5.50 + 2.78 \times 10^{-3Ms} \boldsymbol{q}_{0L}(\overline{w}_E) \\ & + 7.87 \times 10^{-3Ms} \boldsymbol{q}_{1L}(\overline{w}_E) - 1.49 \times 10^{-2Ms} \boldsymbol{q}_{3L}(\overline{w}_E) + \\ & - 0.25^{Ps} \boldsymbol{q}_{1L}^H(\overline{w}_E) + 0.49^{Es} \boldsymbol{q}_{3L}^H(\overline{w}_E) \qquad (2) \\ N = 91; \quad \lambda = 0.36; \quad & D^2 = 6.86; \quad F(5.85) = 29.39; \quad p < 0.05 \end{aligned}$$

where N is the number of compounds, λ is Wilks' statistics, D^2 is the square of the Mahalanobis' distance, F is the Fisher's ratio, and *p* is the significance level.

Table 2. Prediction Performances for 2 LDA-Based QSAR Models (Using Nonstochastic and Stochastic Atom-Type Linear Indices) in the Training and Test Sets

	Matthews Correlation Coefficient (C)	*Accuracy "QTotal" (%)*	*Sensitivity "Hit Rate" (%)*	*Specificity (%)*	*False Positive Rate "False Alarm Rate" (%)*
Nonstochastic bond-based quadratic indices (Eq. 1)					
Learning set	0.75	87.91	85.00	87.00	9.80
Predicting set	0.71	87.50	80.00	80.00	9.00
Stochastic bond-based quadratic indices (Eq. 2)					
Learning set	0.78	89.01	88.00	88.00	10.00
Predicting set	0.65	84.38	80.00	73.00	14.00

LDA, linear discriminant analysis; QSAR, quantitative structure-activity relationships.

Model 1 classifies correctly 85% of active and 90.20% of inactive compounds in the training set for a global good classification (accuracy) of 87.91%. Model 2 classifies correctly 89.01% of the compounds in the training set. Specifically, the model correctly classifies 35 of 40 (87.50%) trichomonacidal compounds and 46 of 51 (90.20%) inactive chemicals in the training series. On the other hand, equations 1 and 2 show an 87.50% (30 of 32) and 84.38% (27 of 32) of global predictability in the prediction series, respectively. These results validate the models for use in the ligand-based virtual screening, taking into consideration that 85.0% is considered as an acceptable threshold limit for this kind of analysis.[28]

In **Tables SM2** and **SM3** (see supplemental material) we give the names of all compounds in the training and test active and inactive sets together with their posterior probabilities calculated from the Mahalanobis distance using both equations. The same information of all compounds in the training and test inactive set appears in **Table 2**, which summarizes the results of the classifications for both models in the training and test groups.

A more serious analysis was carried out by calculating most of the parameters commonly used in medical statistics (accuracy, sensitivity, specificity, and false positive rate) and the Matthews correlation coefficient (*C*). **Table 2** also lists these parameters for both obtained models.[24,29] Whereas the sensitivity is the probability of correctly predicting a positive example, the specificity is the probability that a positive prediction is correct. On the other hand, *C* quantifies the strength of the linear relation between the molecular descriptors and the classifications, and it may often provide a much more balanced evaluation of the prediction than, for instance, the percentages.[24,29] The obtained models, equations 12 and 13, showed a high *C* of 0.75 (0.71) and 0.78 (0.65) in training (test) sets, which correspondingly prove the existence of a strong linear relationship because a value of + 1 implies a total linear agreement between the variables under consideration.

Although the most important criterion for the quality of the discriminant model is based on the statistics for the external prediction set, for a more exhaustive testing of the predictive power of the models, we carried out a leave-10-*fold* full-out (LGO) cross-validation procedure. The statistical results are depicted in **Table SM4** (as supplemental material). The overall mean of the correct classification in training (test) set for this process for equations 1 and 2 was 87.69% (85.64%) and 89.03% (87.86%), correspondingly. The result of predictions on the 10% *full* cross-validation test evidenced the quality (robustness, stability, and predictive power) of the obtained models.

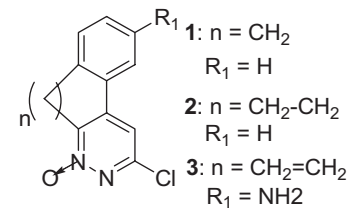
"Virtual screening" as promissory alternative for drug discovery

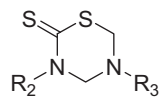
One of the main features that any theoretical approach to drug discovery needs is the identification of active compounds from never-used databases of chemicals. This search can be understood as an alternative to screening approaches to drug discovery. By means of the 1st mentioned procedures, instead of essaying a large number of chemicals in a series of biological tests, one "virtually essays" these compounds by evaluating their activities with the models developed to this effect; this process is known today as computational (*virtual* or *in silico*) screening.[30–36]

To prove the possibilities of the present approach for the ligand-based virtual screening of antitrichomonal compounds, we have selected a series of 12 compounds, as a 2nd external test set, whose activities against Tv have already been proved by several researchers.[37–39] They all were evaluated with models 1 and 2 as active/inactive ones. Its structures as well as the results of the classification are shown in **Table 3**.

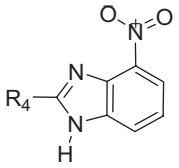
As can be seen, both models classify correctly most of the 12 selected compounds. The 1st model (Eq. 1) classifies only 2 chemicals incorrectly (1 of them as false positive and the other as false negative) for yielding 83.33% of correct classification, whereas the 2nd model (Eq. 2) classifies 3 chemicals incorrectly (2 of them as false positive and the other as false negative) for yielding 75.00% of correct classification. This result is a more important validation criterion for the models developed here because they were able to detect a series of

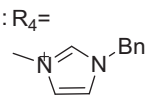
Table 3. Lead Identification among Chemicals Extracted from Literature as Active or Inactive toward the Antitrichomonal Activity by Using Linear Discriminant Analysis–based QSAR Models to Simulate Virtual Screening

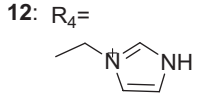




	R_2	R_3
4	furfuril	$(CH_2)_5$-COOH
5	furfuril	CH_2-COOH
6	furfuril	$CH(CH_2Ph)$-COOH
7	furfuril	CH_2-CONH-CH_2-COOH
8	furfuril	$CH(CH_2CONH_2)$-COOH
9	furfuril	$CH[CH_2$-$CH(CH_3)_2]$-COOH
10	$(CH_2)_5$-COOH	$(CH_2)_5$-COOH







Comp.[a]	*Ref.*[b]	$\Delta P\%$[c]	$\Delta P\%$[d]	*Antitrichomonal Activity*
1		−14.70	−36.09	Inactive
2	Gavini et al., 2000	−18.99	−42.54	Inactive
3		−22.78	−24.75	Inactive
4		66.49	−28.97	100 μg/ml = 100[e] 10 μg/ml = (100)[f] 1 μg/ml = (100)[f]
5		76.66	42.66	100 μg/ml = 100[e] 10 μg/ml = (100)[f] 1 μg/ml = (97)[f]
6		86.39	79.25	100 μg/ml = 100[e] 10 μg/ml = (18)[f] 1 μg/ml = (12)[f]
7	Ochoa et al., 1999	91.03	60.63	100 μg/ml = 100[e] 10 μg/ml = (100)[f] 1 μg/ml = (73)[f]
8		87.67	83.14	100 μg/ml = 100[e] 10 μg/ml = (100)[f] 1 μg/ml = (93)[f]
9		68.30	72.92	100 μg/ml = 100[e] 10 μg/ml = (33)[f] 1 μg/ml = (94)[f]
10		−5.25	−97.46	100 μg/ml = 100[e] 10 μg/ml = (25)[f] 1 μg/ml = (65)[f]
11	Alcalde et al., 1995	90.34	83.38	Inactive
12		83.68	73.30	MLC[g] = 50 μg/ml LD_{50}[h] = 50 μg/ml

QSAR, quantitative structure-activity relationships.
a. The molecular structures of the compounds represented with numbers are shown at the top of this table.
b. Bibliographical references from where molecules together with in vitro activities were taken.
c, d. Antitrichomonal activity predicted by Eq. 1 and Eq. 2; $\Delta P\% = [P(\text{Active}) - P(\text{Inactive})] \times 100$.
e. Percentage of reduction of *T. vaginalis* or cytocidal activity at the indicated doses at 24 h.
f. Specific activity against *T. vaginalis* (in parentheses) expressed as percentages of growth inhibition or cytostatic activity at 24 h.
g. MLC, minimum lethal concentration that killed all the parasites by 24 h.
h. LD_{50}, minimum concentration that reduced the number of parasites at least 50%.

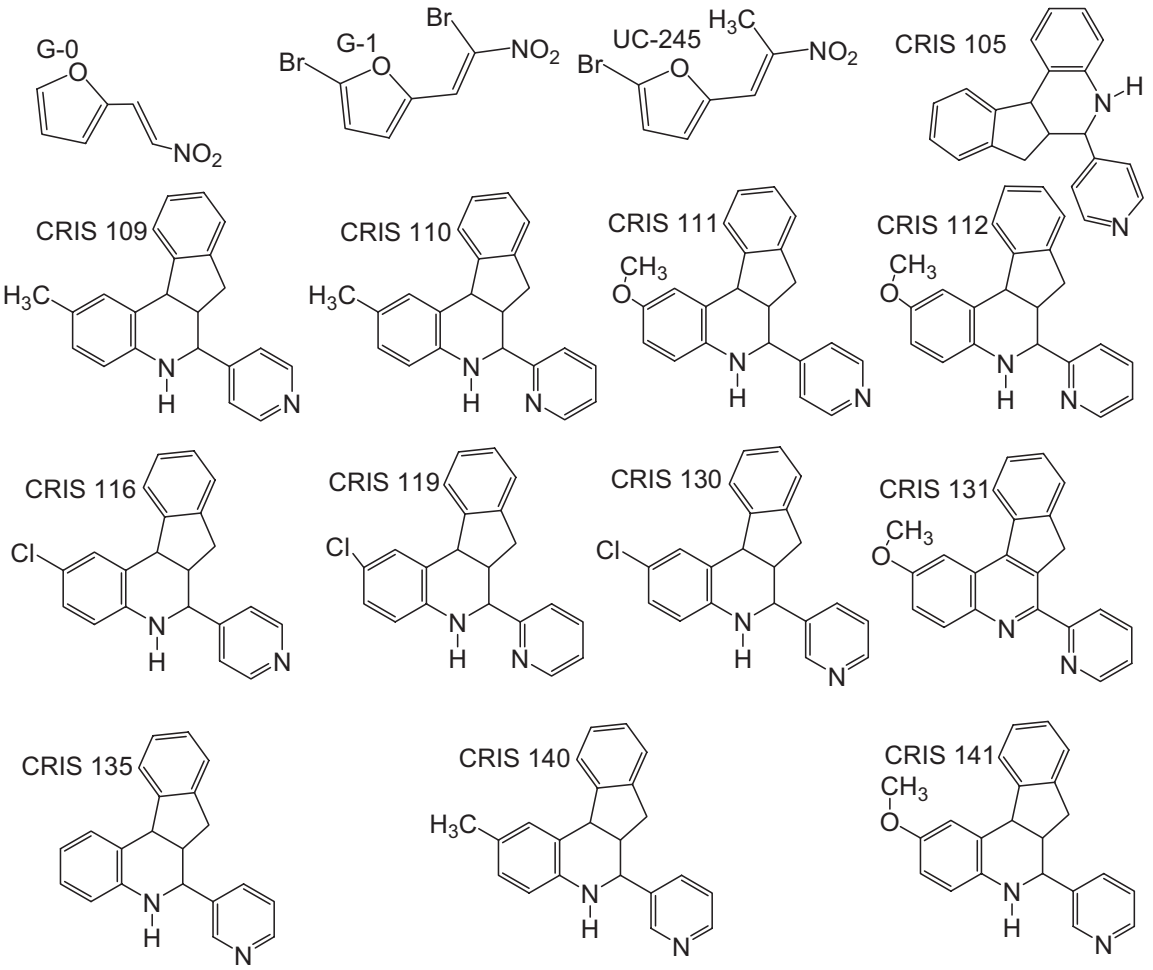


FIG. 2. Structures of nitrovinyl-furans and pyridinyl substituted 7*H*-indeno[2,1-*c*]quinoline derivatives for lead discovery by ligand-based in silico screening.

compounds as active from a database composed of compounds selected from literature, and these chemicals have shown the predicted activity.

Lead discovery by ligand-based in silico screening: from dry selection to wet evaluation

To test the potential of the TOMOCOMD-CARDD method and LDA for detecting novel antiprotozoan leads, we predicted the biological activity of all the chemicals contained in our "in-house" collection of nitrovinyl-furans and pyridinyl substituted 7*H*-indeno[2.1-*c*]quinoline derivatives, which were provided by 2 of our synthesis research teams.[38,40–43] The structures of these compounds are presented in **Figure 2**.

All these compounds were evaluated initially with the QSAR models 1 and 2 and then they were evaluated in vitro, to corroborate the predictions against Tv. The results for the classification and the ΔP% values of the compounds in these series are summarized in **Table 4**. At the same time, this table also depicts the in vitro antitrichomonal activity of these 21 compounds on Tv.

In general, a good correspondence was observed between the theoretical predictions and the observed activity for both active and inactive compounds. Our trained LDA-based QSAR models (Eq. 1 and Eq. 2) successfully classified 20 of 21 compounds yielding (both) an accuracy of the 95.24%.

In these experiments, compounds G-0, G-1, UC-245, and CRIS-148 exhibited pronounced cytocidal activities at the concentrations of 100 μg/ml at 24 h, and 48 h, almost all of them showed cytocidal activity of 100%. Compounds G-1 and UC-245 maintained a good trichomonacidal (cytocidal) activity at 10 μg/ml, although only G-1 maintained a high level of percentage of reduction of Tv at concentrations of 10 mg/ml at 24 h and 48 h—100% in both periods of time. On the contrary, chemicals from CRIS-105 to CRIS-153 but

Table 4. Results of the Computational Evaluation Using LDA-Based QSAR Models and Percentages of Cytostatic and/or Cytocidal Activity [Brackets] for the 3 Concentrations Assayed in Vitro against *Trichomonas vaginalis*

						In Vitro Activity (μg/ml)[f]					
	Theoretical Results					*%CA$_{24h}$ [%C$_{24h}$]*			*%CA$_{48h}$ [%C$_{48h}$]*		
Compound[*]	*Class*[a]	*ΔP%*[b]	*Class*[c]	*ΔP%*[d]	*Class*[e]	*100*	*10*	*1*	*100*	*10*	*1*
G-0	+	61.11	+	46.27	+	**[100]**	50.8	24.5	**[100]**	12.08	0.8
G-1	+	77.34	+	99.72	+	**[100]**	**[100]**	22.42	**[100]**	**[100]**	0
UC-245	+	62.08	+	97.80	+	**[100]**	33.33	0	**[100]**	**[94.18]**	6.84
CRIS 105	—	−93.16	−	−95.33	—	41.18	12.53	1.79	14.75	4.26	0.00
CRIS 109	—	−94.23	—	−94.96	—	90.28	0.00	0.00	93.77	0.00	0.00
CRIS 110	—	−93.19	—	−96.38	—	42.28	4.86	0.00	23.61	0.00	0.00
CRIS 111	—	−87.83	—	−53.01	—	90.28	13.55	0.00	84.90	0.00	0.00
CRIS 112	—	−85.74	—	−64.09	—	84.65	28.39	0.26	85.90	0.00	0.00
CRIS 116	—	−70.66	—	−77.55	—	20.23	2.64	0.00	3.82	0.00	0.00
CRIS 119	—	−66.14	—	−83.47	—	40.18	0.00	0.00	25.19	0.00	0.00
CRIS 130	—	−70.58	—	−79.34	—	74.19	40.18	18.48	95.11	5.34	1.53
CRIS 131	—	−65.10	—	−48.22	—	26.69	0.00	0.00	0.00	0.00	0.00
CRIS 135	—	−93.14	—	−95.73	—	88.86	7.92	0.00	83.59	1.53	0.00
CRIS 140	—	−94.21	—	−95.40	—	17.12	3.30	0.00	13.15	0.00	0.00
CRIS 141	—	−87.79	—	−56.27	—	68.17	0.00	0.00	86.81	0.00	0.00
CRIS 142	—	−79.28	—	−94.70	—	49.43	0.00	17.11	0.00	0.00	0.00
CRIS 143	—	−82.23	—	−94.30	—	59.09	3.42	11.43	43.27	0.00	0.00
CRIS 147	—	−82.28	—	−94.21	—	48.67	0.00	0.00	13.45	0.00	0.00
CRIS 148	—	−82.32	—	−94.40	+	**[96.58]**	2.66	5.70	**[97.93]**	0.00	0.00
CRIS 149	—	−84.88	−	−93.98	—	8.58	0.00	0.00	0.00	0.00	0.00
CRIS 153	—	−28.13	—	−74.22	—	0.00	0.00	0.00	0.00	0.00	0.00
MTZ	+	69.08	+	91.09	+	99.63	99.18	98.19	100	99.72	98.79

LDA, linear discriminant analysis; QSAR, quantitative structure-activity relationships.

[*]The molecular structures of the compounds represented with codes are shown in **Figure 2**.

a, c. In silico classification obtained from Eq. 1 and Eq. 2 using nonstochastic and stochastic bond-type quadratic, respectively.

b, d. Results for the classification of compounds obtained from models Eq. 1 and Eq. 2, correspondingly: $\Delta P\% = [P(\text{Active}) - P(\text{Inactive})] \times 100$.

e. Observed (experimental activity) classification against *Trichomonas vaginalis*.

f. Pharmacological activity of each tested compound, which was added to the cultures at doses of 100, 10, and 1 μg/ml: $\%CA_{\#} = \text{cytostatic activity}_{(24 \text{ or } 48 \text{ hours})}$ and $[\%C_{\#}] = \text{cytocidal activity}_{(24 \text{ or } 48 \text{ hours})}$. MTZ, metronidazole (concentrations for MTZ were 2, 1, and 0.5 mg/ml, respectively).

CRIS-148 were inactive at all assayed concentrations; coinciding with model predictions. It is remarkable that these compounds did not show toxic activity in macrophages cultivations at these concentrations (see **Table 4**).

These last results can be considered as a promising starting point for the future design and refinement of novel compounds with higher antitrichomonal activity with low toxicity. Although compounds G-1, UC-245, G-0, and CRIS-148 were active at higher doses than MTZ (reference drug), this result leaves a door open to a virtual variational study of the structure of these compounds to improve their activity.

Biological "in vivo" assays of G-1

Wistar ovariectomized rats were used in the in vivo experiment of G-1. The results are shown in **Table 5**. A total of 95% of the rats were infected at the beginning of the experiment, higher than in previous reports.[44] The product showed 100% effectiveness from the concentration of 0.125%. With 2 treatments a reduction of 50% in the infected animals was observed, and with the application of the 3rd treatment infected animals were not observed. MTZ was also used as a control in this experiment. With the application of the 2nd treatment of MTZ-infected rats were not observed. The infection controls and excipient remained infected until the end of the experiment.

CONCLUSION

The obtained *biosilico* models permit us to classify new "physical" or "virtual" chemicals as active or inactive ones in the chemotherapy of the trichomoniasis, and they will contribute to a more rational discovery of new lead compounds with antitrichomonal activity. In fact, this report showed that by following the current procedure 4 new chemicals with potentialities (at least in in vitro assays) in antitrichomonal therapeutics were found. One of them (G-1) was even a promissory drug-like compound in the treatment of this disease in more complex living systems (rats). The identification of this new family, making use of the TOMOCOMD-CARDD approach, constitutes an

Table 5. Results of the Activity of G-1 in Wistar Rats Inoculated with *Trichomonas vaginalis*

		Diagnostic/% of Infection						
Group	*Treatment*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
I	G-1	100	50	10	0	0	0	0
II	Migliol	80	80	80	80	80	80	80
III	Control without treatment	100	100	100	100	100	100	100
IV	Metronidazol	100	20	0	0	0	0	0
	% of total infection	95						

example of how this rational computer-aided method can help to reduce cost and to increase the rate in which novel chemical entities progress through the pipeline.

ACKNOWLEDGMENTS

The authors express their gratitude to Prof. Dr. Jorge Gálvez for his attention to this work and valuable suggestions. The authors also offer their sincere thanks to the 2 anonymous referees for their critical opinions about the manuscript, which have significantly contributed to improving its presentation and quality. Y.M.-P. thanks the Valencia University (Spain) for partial financial support as well as the program "Estades Temporals per an Investigadors Convidats" for a fellowship to work at Galvez's group, Pharmacy Faculty, Valencia University (2008). Support for some authors was received from Spanish MEC (Project Reference: SAF2006-04698). F.T. is grateful for support from Spanish MEC DGI (Project No. CTQ2004-07768-C02-01/BQU) and Generalitat Valenciana (DGEUI INF01-051 and INFRA03-047, and OCYT GRUPOS03-173). Finally, but not least, the CAMD-BIR Unit is grateful for the partial support of this research from a project funded by the Flemish Interuniversity Council (VLIR) of Belgium: "*Strengthening postgraduate education and research in Pharmaceutical Sciences.*"

REFERENCES

1. Heine P, McGregor JA: Trichomonas vaginalis: a reemerging pathogen. *Clin Obstet Gynecol* 1993;36:137-144.
2. World-Health-Organization: an overview of selected curable sexually transmitted diseases. *Global program on AIDS*. Geneva: World Health Organization, 1995:2-27.
3. Cosar C, Julou L: Activity of 1-(29-hydroxyethyl)-2-methyl-5-nitroimidazole (8823 rp) against experimental trichomonas vaginalis infection. *Ann Inst Pasteur* 1959;96:238-241.
4. Centers for disease control and prevention: sexually transmitted diseases treatment guidelines. *Morb Mortal Wkly Rep* 1993;42(RR-14):70.
5. Knight R: The chemotherapy of amoebiasis. *J Antimicrob Chemother* 1980;6:577-593.
6. Gillette H, Schmid GP, Moswe D, et al: Metronidazole-resistant *Trichomonas* vaginalis, a case series, 1985–1998 [abstract 067]. Presented at the XIIIth Meeting of the International Society of Sexually Transmitted Disease Research, Denver, July 11-14, 1999.
7. Narcisi EM, Secor WE: In vitro effect of tinidazole and furazolidone on metronidazole-resistant trichomonas vaginalis. *Antimicrob Agents Chemother* 1996;40:1121-1125.
8. Houang ET, Ahmet Z, Lawrence AG: Successful treatment of four patients with recalcitrant vaginal trichomoniasis with a combination of zinc sulfate douche and metronidazole therapy. *Sex Transm Dis* 1997;24:116-119.
9. Pattman RS, Sprott MS, Kerns AM, Earnshaw M: Failure of mebendazole to cure trichomonal vaginitis resistant to metronidazole: case reports. *Genitourin Med* 1989;65:274-275.
10. Wong CA, Wilson PD, Chew TA: Providone-iodine in treatment of metronidazole-resistant trichomonas vaginalis. *Aust N Z J Obstet Gynaecol* 1990;30:169-171.
11. Livengood CHI, Lossick JG: Resolution of resistant vaginal trichomoniasis associated with the use of intravaginal nonoxynol-9. *Obstet Gynecol* 1991;78:954-956.
12. Watson PG, Pattman RS: Arsenical pessaries in the successful elimination of metronidazole-resistant trichomonas vaginalis. *Int J STD AIDS* 1996;7:296-297.
13. Nyirjesy P, Sobel JD, Weitz MV: Difficult to treat trichomoniasis: results with paromomycin cream. *Clin Infect Dis* 1998;26:986-988.
14. Nyirjesy P, Weitz MV, Gelone SP, Fekete T: Paromomycin for nitroimidazole-resistant trichomonosis. *Lancet* 1995;346:1110.
15. Marrero-Ponce Y, Torrens F, Alvarado YJ, Rotondo R: Bond-based global and local (bond, group and bond-type) quadratic indices and their applications to computer-aided molecular design. 1. QSPR studies of diverse sets of organic chemicals. *J Comp Aided Mol Design* 2006; 20:685.
16. Marrero-Ponce Y, Romero V: *TOMOCOMD* [Software]. Villa Clara, Cuba: Central University of Las Villas, 2002.
17. Kier LB, Hall LH: *Molecular Connectivity in Structure–Activity Analysis*. Letchworth, UK: Research Studies Press, 1986.
18. Negwer M: *Organic-Chemical Drugs and their Synonyms*. Berlin: Akademie-Verlag, 1987.
19. Budavari S, O'Neil MJ, Smith A, Heckelman PE, Kinneary JF (eds): *The Merck Index* [CD-ROM]. Whitehouse Station, NJ: Chapman & Hall/CRC, 1999.
20. van de Waterbeemd H: In van Waterbeemd H (ed): *Chemometric Methods in Molecular Design*. Weinheim: VCH Publishers, 1995:265-288.
21. Gálvez J, Garcia-Domenech R, de Julián-Ortiz JV, Soler R: Topological approach to drug design. *J Chem Inf Comput Sci* 1995;35:272-284.
22. Cercos-del-Pozo RA, Pérez-Giménez F, Salabert-Salvador MT, Garcia-March FJ: Discrimination and molecular design of new theoretical hypolipaemic agents using the molecular connectivity functions. *J Chem Inf Comput Sci* 2000;40:178-184.
23. Wold S, Erikson L: Statistical validation of QSAR results. Validation tools. In van de Waterbeemd H (ed): *Chemometric Methods in Molecular Design*. New York: VCH Publishers, 1995:309-318.
24. Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H: Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics* 2000;16:412.
25. Kouznetsov VV, Rivero CJ, Ochoa PC, Stashenko E, Martínez JR, Montero PD, et al: Synthesis and antiparasitic properties of new 4-N-benzylamino-4-hetarylbut-1-enes. *Arch Pharm* 2005;338:32-37.
26. Kouznetsov VV, Vargas MLY, Tibaduiza B, Ochoa C, Montero PD, Nogal RJJ, et al: 4-Aryl(benzyl)amino-4-heteroarylbut-1-enes as building blocks in heterocyclic synthesis. 4. Synthesis of 4,6-dimethyl-5-nitro(amino)-2-pyridylquinolines and their antiparasitic activities. *J Arch Pharm* 2004;337:127-132.

27. Meinsgassner JG, Georgopolus A, Patoschka M: Intravaginale infektionen der ratte mit trichomonas vaginalis und candida albicans. Ein modell zur experimentellen chemotherapie. *Trop Parasitol* 1975;26:395-398.
28. Gálvez J, García R, Salabert MT, Soler R: Charge indexes. New topological descriptors. *J Chem Inf Comput Sci* 1994;34:520-525.
29. Johnson RA, Wichern DW: *Applied Multivariate Statistical Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1988.
30. Xu J, Hagler A: Chemoinformatics and drug discovery. *Molecules* 2002;7:566-700.
31. Seifert MHJ, Wolf K, Vitt D: Virtual highthroughput in silico screening. *Biosilico* 2003;1:143.
32. Watson C: Predictive in silico models in drug discovery. *Biosilico* 2003; 1:83.
33. Scott RK: Informatics integration: the bedrock of nce selection. *Biosilico* 2003;1:14.
34. Estrada E, Uriarte E, Montero A, Teijeira M, Santana L, De Clercq E: A novel approach for the virtual screening and rational design of anticancer compounds. *J Med Chem* 2000;43:1975-1985.
35. McFarland JW, Gans DJ: Cluster significance analysis. In Waterbeemd H (ed): *Chemometric Methods in Molecular Design*. New York: VCH Publishers, 1995:295–307.
36. Estrada E, Uriarte E: Recent advances on the role of topological indices in drug discovery research. *Curr Med Chem* 2001;8:1573-1588.
37. Alcalde E, Pérez L, Dinarés I, Frigola J: Heterocyclic betaines. XXII. Azinium(azolium) 4-nitrobenzimidazolate inner salts and their derivatives with several interannular spacers. Synthesis, characterization and antitrichomonal activity. *Chem Pharm Bull (Tokyo)* 1995;43:493-498.
38. Ochoa A, Pérez E, Pérez R, Suárez M, Ochoa E, Rodríguez H, et al: Synthesis and antiprotozoan properties of new 3,5 disubstitued tetrahydro-2h-1,3,5-thiadiazine-2-thione derivatives. *Arzneim Forsch* 1999;49: 764-769.
39. Gavini E, Juliano C, Mulé A, Pirisino G, Murineddu G, Pinna A: Synthesis and in vitro antimicrobial properties of n-oxide derivates based on tricyclic indeno[2,1-c] pyridazine and benzo [f] cinnoline systems. *Arch Pharm (Weinheim)* 2000;333:341-346.
40. González JB, Creus A, Marcos R: Genetoxic evaluation of the furylethylene derivative 2-furyl-1-nitroethene in cultured human lymphocytes. *Mutat Res* 2001;497:177-184.
41. Castañedo NR, Goizueta RD, Gónzales O, Peréz JA, Gónzales J, Silveira EA: Procedure for the obtención of 1-(5-bromofur-2-il)-2-bromo-2-nitroeteno and its action as micricide. *European solicitud 95 500056 7 / 1270* 1995.
42. Pérez GM, González JB, Casteñedo NC, Creus A, Marcos R: In vitro genotoxicyty testing of the furylethylene derivate uc-245 in human cells. *Mutagenesis* 2004;19:75-80.
43. Kouznetsov VV, Ochoa C, Romero-Bohórquez AR, Zacchino SA, Sortino M, Gupta M, et al: A straightforward synthetic approach to antitumoral pyridinyl substituted 7h-indeno[2.1-c]quinoline derivatives via three-component imino diels-alder reaction. *Lett Org Chem* 2006;3: 300-304.
44. Meinsgassner JG, Georgopolus A, Patoschka M: Intravaginale infektionen der ratte mit trichomonas vaginalis und candida albicans. Ein modell zur experimentellen chemotherapie. *TropParasitol* 1975;26: 395-398.

Address correspondence to:
Yovani Marrero-Ponce, Ph.D.
Unit of Computer-Aided Molecular "Biosilico"
Discovery and Bioinformatic Research (CAMD-BIR Unit)
Faculty of Chemistry-Pharmacy
Central University of Las Villas
Santa Clara, 54830 Villa Clara, Cuba

E-mail: ymarrero77@yahoo.es