

## Chalcones and *N*-acylhydrazones: direct analogues? Exploratory data analysis applied to potential novel antileishmanial agents

Daniela Gonçalves Rando\*, Jeanine Giarolla, Kerly Fernanda Mesquita Pasqualoto,  
Elizabeth Igne Ferreira

Laboratory of design and synthesis of chemotherapeutical agents potentially active against tropical diseases – LAPEN,  
Department of Pharmacy, Faculty of Pharmaceutical Sciences, University of São Paulo

Leishmaniasis is an important health and social problem for which there is limited effective therapy. Chalcones and *N*-acylhydrazones have been studied as promising antileishmanial agents in enzymatic inhibition and *in vitro* assays. Since these chemical classes of compounds also resemble each other structurally, it would be useful to investigate whether they share direct analogy. Exploratory data analysis was applied to a library of chalcones and nitrated *N*-acylhydrazones assayed against *Leishmania donovani* to investigate their similarity. Under the conditions applied in the present study, the two classes did not present functional or structural analogy.

**Uniterms:** Chalcones. *N*-acylhydrazones. Antileishmanial agents. Leishmaniasis. Chemometrics.

As leishmanioses são importantes problemas sociais e de saúde pública para os quais a terapia farmacológica atual é, ainda, limitada. Chalconas e *N*-acilidrazonas têm sido estudadas como promissoras agentes leishmanicidas tanto em ensaios *in vitro* quanto em ensaios de inibição de cisteino-proteases importantes para o parasito. Uma vez que estas classes de compostos apresentam similaridade bidimensional, seria interessante estudar se estes compostos guardariam relação de analogia direta entre si. Análise exploratória de dados foi aplicada, então, à biblioteca de chalconas e *N*-acilidrazonas nitradas ensaiadas contra *Leishmania donovani* para investigar suas relações de similaridade. Os resultados mostraram que, ao menos sob as condições consideradas neste estudo, as duas classes de compostos não apresentam analogia estrutural e funcional simultaneamente, embora elas apresentem alguma similaridade estrutural.

**Unitermos:** Chalconas. *N*-acilidrazonas. Agentes leishmanicidas. Leishmaniose. Quimiometria.

### INTRODUCTION

Leishmaniasis is a neglected disease that affects millions of individuals worldwide causing significant morbidity and mortality, particularly in developing countries. Available drug therapy, however, is limited and frequently inappropriate (Chung *et al.*, 2008). Of the 1393 new drugs marketed between 1975 and 1999, only 13 were targeted for tropical diseases (Trouiller *et al.*, 2002). In fact, the effective drugs are toxic and often require parenteral administration for long treatment courses, increasing the

chances of failure due to emergence of resistance (Croft, 2005). This scenario emphasizes the importance of the search for completely new active compounds.

Screening of natural products offers the promise of discovering new molecules with a unique structure, and both high activity and selectivity (Kayser *et al.*, 2003). Chalcones are natural 1,3-diarylpropanones which exhibit a broad spectrum of potential applications, including antiprotozoal activity (Nielsen *et al.*, 1998; Liu *et al.*, 2001). This activity is likely derived from a cysteine protease inhibition mechanism. Cysteine proteases are key enzymes in many parasitic biochemical pathways and, therefore, constitute potential targets in the search for drugs against several tropical infectious diseases (McKerrow *et al.*, 1999; McKerrow, 1999). For these reasons, sets of chalcone analogues have also been tested as potential cysteine proteases inhibitors.

\*Correspondence: D. G. Rando. Laboratório de planejamento e síntese de quimioterápicos potencialmente ativos contra endemias tropicais - LAPEN, Departamento de Farmácia, Faculdade de Ciências Farmacêuticas - USP. Av. Prof. Lineu Prestes, 580 - Bloco 13 Superior - Cidade Universitária - 05508900 - São Paulo - SP, Brasil. E-mail: [dgrando@usp.br](mailto:dgrando@usp.br)

Studies comparing the inhibitory effect of *N*-acylhydrazones and chalcones on well-known parasitic cysteine proteases, such as cruzain, falcipain, and trypanopain, have been conducted (Li *et al.*, 1995, 1996; Troeberg *et al.*, 2000;). Both classes of compounds present IC<sub>50</sub> values in a micromolar range, suggesting that these compounds could be applied as prototypes in the development of new inhibitor agents. Since these chemical classes of compounds also resemble each other structurally, it would be interesting to verify whether they share direct analogy.

Direct analogues are defined as chemical entities which present, simultaneously, chemical and pharmacological similarities (Wermuth, 2006).

Structural and functional analogy between chalcones and *N*-acylhydrazones would bring some advantages in the search for new antileishmanial agents. 3D and 4D-QSAR studies applied to both chemical classes could be able to provide ligand-receptor interaction information, and a common pharmacophore. Those findings could be useful for designing new and more potent compounds, which accumulate the significant molecular features of both chalcones and *N*-acylhydrazones.

Several chemometric methods have broadened the arsenal of tools that can be applied to QSAR studies. Exploratory data analysis methods of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) (Beebe *et al.*, 1998) number among these methods.

HCA is an important multivariate method whose primary purpose is to emphasize the clusters and patterns of the investigated data, displaying them as a dendrogram, which allows the visualization of the samples or variables in a 2D space. The distances between samples or variables are calculated and transformed into a similarity matrix whose elements correspond to the similarity indexes (Ferreira, 2002).

PCA however, is a data compression multivariate method based on correlation among variables. Its objective is to group correlated original independent variables or descriptors and replace them with a new set called principal components (PCs) onto which the data is projected. The PCs are completely uncorrelated and are built as a simple linear combination of original variables. Furthermore, the PCs contain most of the variability in the data set within a much smaller dimensional space (Ferreira, 2002).

Thus, in this study, an exploratory data analysis using the above-mentioned methods was applied to a set of ninety-four chalcones and nitrated *N*-acylhydrazones to verify if they indeed share **direct analogy**. This is fundamental to subsequent construction of 3D-QSAR models, to explaining their antileishmanial behavior and to finding a common pharmacophore, in order to design new parasitic

agents for this activity.

## METHODOLOGY

A complete series of ninety-four compounds were investigated and are listed in Figure 1. The set encompasses forty-nine 5-nitroheterocyclic benzhydrazides, synthesized by Rando *et al.* (2008), and forty-five chalcone derivatives, which were selected from the work of Nielsen *et al.* (1998).

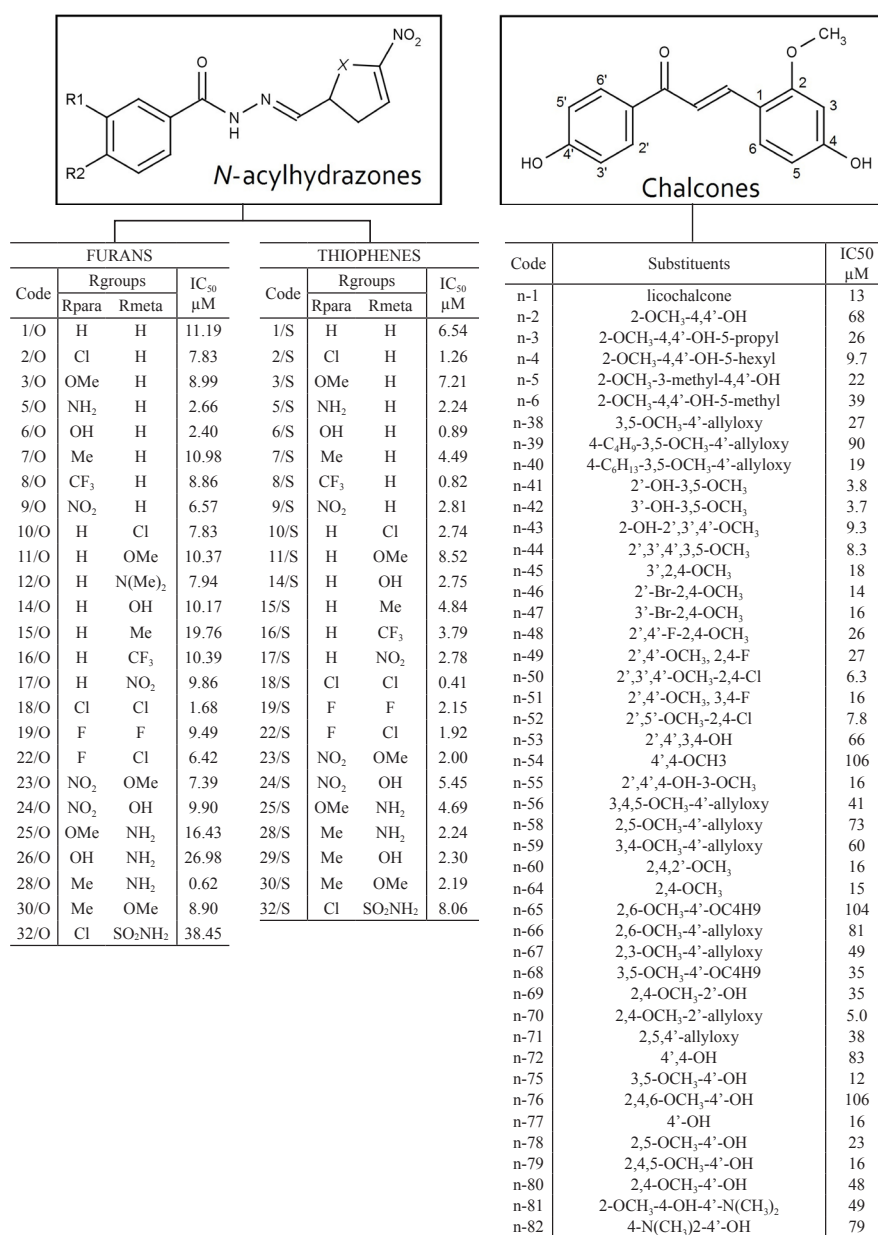
Biological activities were evaluated as the compounds *in vitro* concentration capable of inhibiting 50 percent of *Leishmania donovani* promastigote proliferation (IC<sub>50</sub>, μM) (Figure 1). The IC<sub>50</sub> values were expressed in negative logarithmic units, pIC<sub>50</sub> (-log IC<sub>50</sub>), and comprise the set of dependent variables.

The three-dimensional structures of each of the ninety-four compounds or ligands, in their neutral forms, were constructed employing the HyperChem 7 software (HyperChem, 2002). The crystallized structures of nitrofurazone and 2'-hydroxy-4''-dimethylamino-chalcone were retrieved from the Brookhaven Protein Data Bank (PDB entry code 1yki, resolution at 1.70 Å; Race *et al.*, 2005) and file ob1067.cif (resolution at 1.01 Å; Liu *et al.*, 2002), respectively, and were used as geometry references in the building up of all ligands. Crystallographic information of a nitrothiophene 2-carbaldehyde was retrieved from the file lh6379.cif (resolution at 1.18 Å; McBurney *et al.*, 2005) and also used as a reference particularly in the construction of the nitrothiophene ring moiety.

The energy minimization was carried out employing the HyperChem 7 MM+ force field without any restriction. The MOLSIM 3.2 program (Doherty, 1997) was also used for geometry optimization of each structure investigated. Partial atomic charges were calculated using the PM3 (Stewart, 1989) semi-empirical method, also implemented in the HyperChem program.

The structures modeled as described above were used as the initial structures in each molecular dynamics (MD) simulation (Van Gunsteren, Berendsen, 1990), which are employed to perform the conformational ensemble profile of each ligand (MOLSIM 3.2).

The MD simulation protocol included 100,000 steps, a step size of 0.001 ps (1fs), and a simulation temperature of 300 K. Trajectory files were recorded every 20 simulation steps to generate 5,000 conformations for each molecule. The lowest energy conformation was selected from MD simulation for each ligand and electrostatic partial atomic charges (ChelpG) were computed using the HF/6-31G\* method, implemented in the Gaussian 03W program (Gaussian, 2003).



**FIGURE 1** - Structures and actions of the investigated set of chalcones and nitrated *N*-acylhydrazones

The independent variables or descriptors, related to thermodynamic, steric, electronic, hydrophobic and structural contributions, were calculated employing specific software such as MOLSIM 3.2 (Doherty, 1997), Marvin Beans (MarvinBeans, 2007), Gaussian 03W (Gaussian, 2003), Pallas 3.0 (Pallas, 1995), and HyperChem 7 (HyperChem, 2002). The total data matrix contains 94 rows (number of samples or compounds) and 48 columns, including biological activity. The samples were randomized before performing the exploratory data analysis.

HCA and PCA (Beebe, 1998) were carried out using the Pirouette 3.11 program (Pirouette, 2003). The autoscaling procedure was applied as a preprocessing method,

meaning that each element of data matrix was subtracted by its mean column and divided by the standard deviation of its column before the analysis. After screening using different linkage methods, HCA studies were performed considering the centroid linkage method. The PCA formalism was run up to seven factors.

## RESULTS AND DISCUSSION

Table I shows the complete list of independent variables calculated for the compounds together with their descriptions and other relevant information.

The thermodynamic descriptors found for the lowest

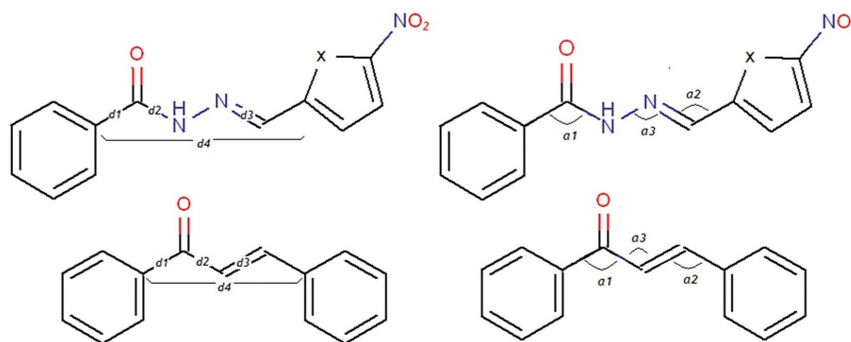
TABLE I - Descriptors calculated for the set studied

Descriptor	Unity	Nature	Description	Software
<b>EL<sub>stretch</sub></b>	kcal/mol	Thermodynamics	Stretching energy contribution	
<b>EL<sub>bend</sub></b>	kcal/mol	Thermodynamics	Bending energy contribution	
<b>EL<sub>tors</sub></b>	kcal/mol	Thermodynamics	Torsional energy contribution	
<b>EL<sub>vdw</sub></b>	kcal/mol	Thermodynamics	Van der Waals energy contribution	
<b>EL<sub>charge</sub></b>	kcal/mol	Thermodynamics	Electrostatic energy contribution	<i>MOLSIM</i>
<b>EL<sub>hb</sub></b>	kcal/mol	Thermodynamics	Hydrogen bound energy contribution	
<b>EL<sub>14</sub></b>	kcal/mol	Thermodynamics	Lennard-Jones energy contribution	
<b>EL<sub>solvate</sub></b>	kcal/mol	Thermodynamics	Solvation energy contribution	
<b>EL<sub>tot</sub></b>	kcal/mol	Thermodynamics	Total ligand energy (sum of all energy contributions)	
<b>E<sub>lumo</sub></b>	kcal/mol	Electronic	LUMO frontier orbital energy	
<b>E<sub>homo</sub></b>	kcal/mol	Electronic	HOMO frontier orbital energy	
<b>GAP</b>	kcal/mol	Electronic	Difference between HOMO and LUMO energies (H-L)	
<b>ChelpG 1 (EPS)</b>	---	Electronic	Electrostatic potential charges under carbonyl carbon calculated by ChelpG	
<b>ChelpG 2 (EPS)</b>	---	Electronic	Electrostatic potential charges under carbonyl oxygen calculated by ChelpG	
<b>ChelpG 3 (EPS)</b>	---	Electronic	Electrostatic potential charges under $\beta$ -carbon (chalcones) or C=N carbon ( <i>N</i> -acylhydrazones) calculated by ChelpG	<i>Gaussian 03W for Windows</i>
<b>ChelpG 4 (EPS)</b>	---	Electronic	Rate between electrostatic potential charges calculated for the carbonyl carbons and C $\beta$ (chalcones) or C=N carbon ( <i>N</i> -acylhydrazones).	
$\mu_x$	debye	Electronic	X component of dipole moment.	
$\mu_y$	debye	Electronic	Y component of dipole moment.	
$\mu_z$	debye	Electronic	Z component of dipole moment.	
$\mu_{tot}$	debye	Electronic	Total dipole moment.	
<b>ClogP<sub>Hyper</sub></b>	---	Hydrophobic	Calculated partition coefficient (water/ <i>n</i> -octanol)	
<b>Volume</b>	Å <sup>3</sup>	Steric	Total volume	
<b>MR</b>	Å <sup>3</sup>	Steric Hydrophobic	Molar refractivity.	
$\alpha$	Å <sup>3</sup>	Steric/Electronic	Polarizability	<i>HyperChem Program Release 7 for Windows</i>
<b>Mass</b>	uma	Steric	Atomic mass	
<b>SAapprox</b>	Å <sup>2</sup>	Steric	Approximated surface area	
<b>SAgrid</b>	Å <sup>2</sup>	Steric	Surface area obtained by GRID method	
<b>ClogP<sub>Pallas</sub></b>	---	Hydrophobic	Calculated partition coefficient (water/ <i>n</i> -octanol)	
<b>LogD 2.0</b>	---	Hydrophobic		
<b>LogD 7.0</b>	---	Hydrophobic	Apparent partition coefficient in different pH systems (2.0, 7.0 and 9.0).	<i>Pallas 3.0</i>
<b>LogD 9.0</b>	---	Hydrophobic		

**TABLE I** - Descriptors calculated for the set studied (cont.)

$I_{\text{balaban}}$	---	Topologic	Index which describes the averaged distance sum connectivity.	
$H_{\text{Donor}}$	---	Topologic	Number of hydrogen bond acceptors and donor sites in structures.	
$H_{\text{Accep}}$	---	Topologic		
<b>Dreiding</b>	---	Steric	Dreiding energy (conformers stability)	
<b>Hind 1</b>	---	Steric	Steric hindrance under carbonyl carbon.	
<b>Hind 2</b>	---	Steric	Steric hindrance under $\beta$ -carbon (chalcones) or under C=N carbon, C=N ( <i>N</i> -acylhydrazones).	
$SA_{\text{VDW}}$	$\text{\AA}^2$	Steric	Van der Waals surface area.	
$SA_{\text{SolvAccess}}$	$\text{\AA}^2$	Steric	Solvent access surface area.	<i>MarvinBeans</i>
<b>PSA</b>	$\text{\AA}^2$	Steric	Polar surface area.	
<b>D1</b>	$\text{\AA}$	Structural	Distance d1	
<b>D2</b>	$\text{\AA}$	Structural	Distance d2	
<b>D3</b>	$\text{\AA}$	Structural	Distance d3	
<b>D4</b>	$\text{\AA}$	Structural	Distance d4	
<b>A1</b>	degrees	Structural	Angle a1	
<b>A2</b>	degrees	Structural	Angle a2	
<b>A3</b>	degrees	Structural	Angle a3	

Referred distances and angles:



energy conformation from MD simulation of each ligand included the following energy contributions: stretching energy ( $EL_{\text{stretch}}$ ), bending energy ( $EL_{\text{bend}}$ ), torsional energy ( $EL_{\text{tors}}$ ), Lennard-Jones or 1,4 interactions energy ( $EL_{1,4}$ ), electrostatic energy ( $EL_{el}$ ), van der Waals energy ( $EL_{\text{vdW}}$ ), hydrogen bonding energy ( $EL_{\text{Hb}}$ ), and solvation energy ( $EL_{\text{solv}}$ ). The total energy of each ligand ( $EL_{\text{tot}}$ ) corresponds to the summation of all these energy contributions.

The correlation between the biological data ( $\text{pIC}_{50}$  values) and the calculated independent variables or descriptors was visualized through the respective scatter plots. This is an important step in exploratory data analysis for acquiring reliable information about the dataset behavior. Correlation coefficients were not employed as a cutoff

factor because other types of correlation, besides the linear type, must also be considered in similarity studies.

The  $D_4$  parameter, for instance, describes the total distance between the two aromatic rings in both subsets of compounds. The  $D_4$  range values found for the nitrated *N*-acylhydrazones and chalcone derivatives were 4.80 to 4.90  $\text{\AA}$  and 4.90 to 5.20  $\text{\AA}$ , respectively. These distance ranges suggest that, although the intermediary chain of *N*-acylhydrazones possesses four atoms, its 3D models are comparable to those of the chalcone derivatives, which have only three atoms in their intermediary chain.

Conversely, the scatter plot  $\text{pIC}_{50}$  versus  $D_4$  indicates that the related structural similarity does not translate to functional similarity (Figure 2).

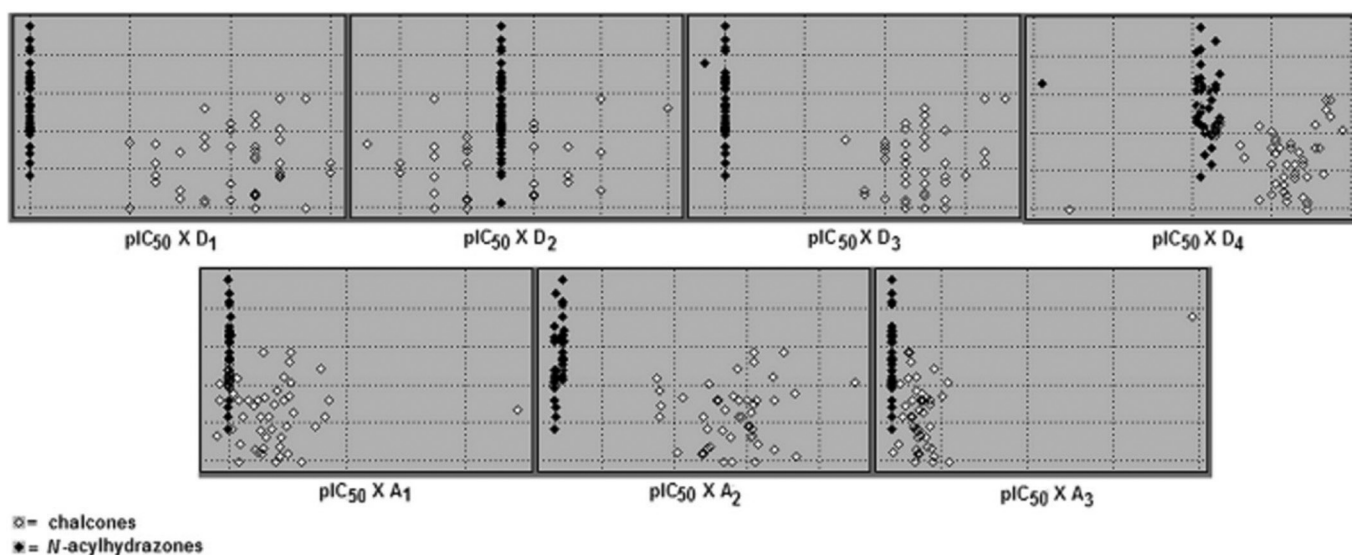


FIGURE 2 - Scatter plots of biological activity ( $pIC_{50}$ ) versus structural descriptors.

Nitrated *N*-acylhydrazones presents a narrow range of  $D_4$  values but a significant  $pIC_{50}$  value range. This indicates that factors other than  $D_4$  are interfering with the biological action of these compounds. Similar behavior was also verified for the other distances and angles measured, particularly the  $D_2$ ,  $A_1$  and  $A_2$  descriptors (Figure 2 and Table II).

The visual inspection of the biological activity versus calculated descriptor scatter plots was also applied as a filter for selecting the independent variables that would be used to perform HCA and PCA. Concerning scatter plot behavior, the independent variables can be classified as shown in Table III.

Among the forty-seven calculated descriptors, thirteen (*Ehb*, *Ehomo*, *ChelpG 4*, *D1*, *D2*, *D3*, *A1*, *A2*, *A3*, *H<sub>Donor</sub>*, *Dreiding*, *Hind 1*, *Hind 2*) were practically constant and did not present significant variability with the

TABLE II - Topological and structural parameters calculated for the investigated compounds

Descriptor	Chalcones	<i>N</i> -acylhydrazones
$D_1$	1.40 – 1.48	1.36
$D_2$	1.33-1.42	1.37
$D_3$	1.32-1.40	1.26
$D_4$	4.90-5.20	4.80-4.90
$A_1$	118.9-144.5	119.9-120.1
$A_2$	123.8-137.4	116.7-117.3
$A_3$	118.7-131.7	118.5

$pIC_{50}$  values. Moreover, six parameters (*ELstrech*, *Elumo*, *GAP*, *CLogP<sub>Hyper</sub>*, *D4*, *PSA*) were grouped in their scatter plots, as opposed to showing a regular dispersion behavior, and were also not considered for further analysis. Thus,

TABLE III - Descriptor classification regarding tendency on scatter plots versus  $pIC_{50}$  values

TENDENCY OF $pIC_{50}$ X DESCRIPTOR				
HIGH		LOW		MODERATE
$EL_{bend}$	$\mu_{tot}$	$EL_{stretch}$	$D_4$	$EL_{solvat}$
$EL_{14}$	Volume	$EL_{hb}$	$A_1$	$ClogP_{Pallas}$
$EL_{tot}$	MR	$E_{lumo}$	$A_2$	$LogD_{2,0}$
$EL_{tors}$	$\alpha$	$E_{homo}$	$A_3$	$LogD_{7,0}$
$EL_{vdw}$	Mass	GAP	$H_{Donor}$	$LogD_{9,0}$
$EL_{charge}$	$SA_{approx}$	ChelpG 4	Dreiding	$\mu_y$
ChelpG 1	$SA_{GRID}$	$ClogP_{Hyper}$	Hind 1	$H_{Accep}$
ChelpG 2	$I_{Balaban}$	$D_1$	Hind 2	
ChelpG 3	$SA_{vdw}$	$D_2$	PSA	
$\mu_x$ $\mu_z$	$SA_{solventaccess}$	$D_3$		

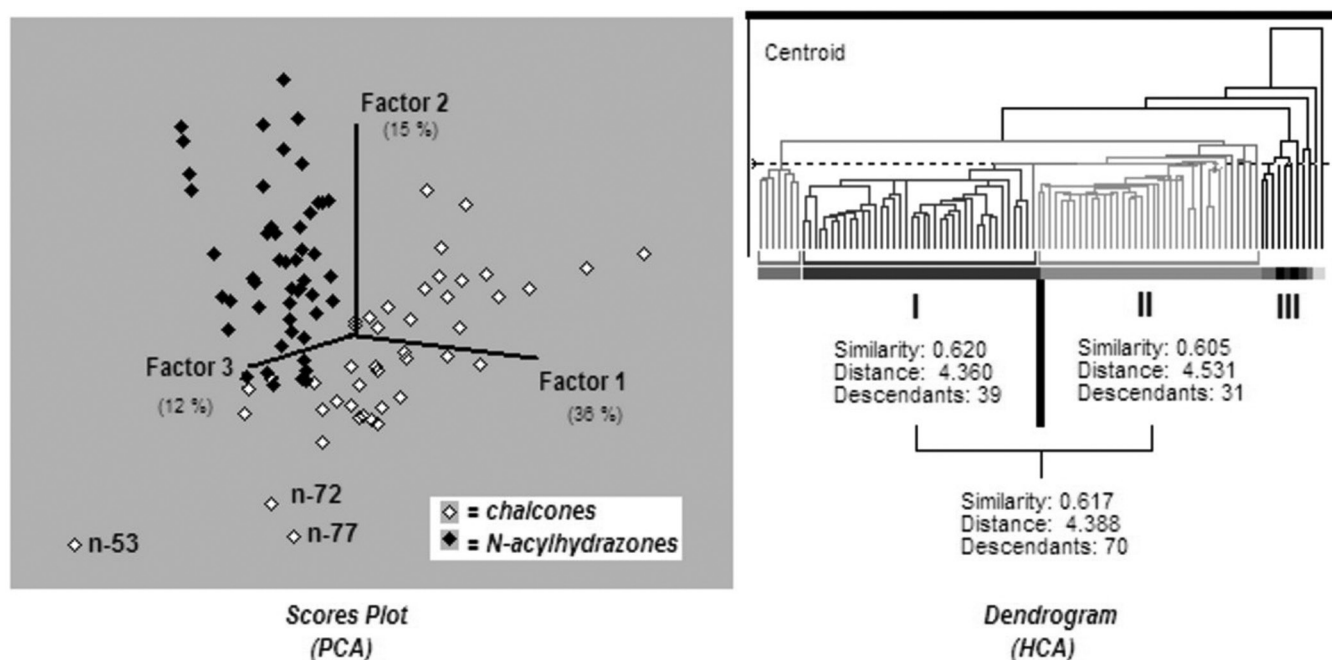


FIGURE 3 - Preliminary PCA and HCA of the investigated set of compounds.

twenty-eight descriptors, which presented good or moderate correlation with biological activity, were employed in the exploratory data analysis (Table III). Volume and atomic mass were excluded from the analysis because they were not considered very expressive.

Preliminary HCA displayed chalcones and *N*-acylhydrazones grouped in two different main sub-clusters: *I*, which comprises exclusively nitrated *N*-acylhydrazones, and *II* comprising only chalcones, as presented in Figure 3. The similarity index between the sub-clusters *I* and *III* is 0.617 (70 selected descendants), suggesting low similarity. Additionally, sub-cluster *III* is composed of chalcone derivatives that are significantly different from the chalcone derivatives grouped in sub-cluster *II*. Moreover, regarding the PCA 3D-diagram, *N*-acylhydrazones are mainly distributed in the PC2 and PC3 spaces whereas chalcones can be found in all factors (PC1, PC2 and PC3 space), indicating distinct behavior among the compounds of this subset.

The PCA scores plot (Figure 3) also reveals three chalcone derivatives (n-53, n-72 and n-77) that possibly differ from the others. These three compounds could be acting as outliers interfering with the overall dataset behavior and leading to miscalculations. A 2D outlier diagnosis plot was therefore constructed to confirm this suspicion (Figure 4).

Considering the outlier diagram (sample residual versus Mahalanobis distance), samples falling outside

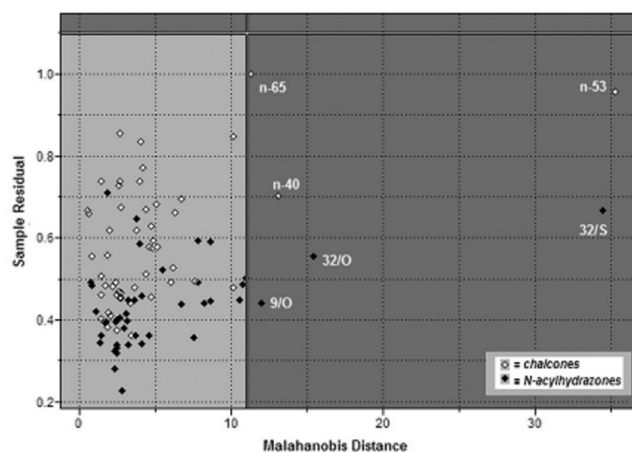


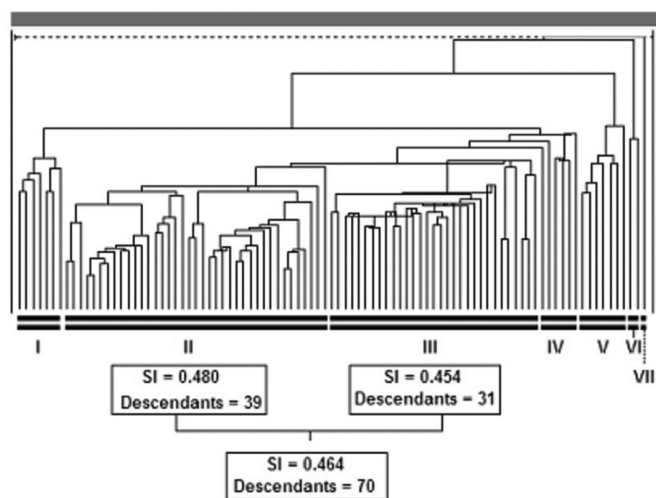
FIGURE 4 - Outlier diagram plot.

one or both of the thresholds are potential outliers. The sample residual threshold is based on a 95% probability limit (Pirouette, 2006), thus 5% of *normal* samples would be expected to fall outside this cutoff. For this reason, samples slightly exceeding only one threshold may be normal. However, samples lying either significantly beyond one threshold, or beyond both, are more likely to be outliers. Accordingly, only compound n-53 can be considered an outlier. In addition, another compound can also be identified as a potential outlier, namely, 32/S.

\* Similarity indexes calculated for the seventy compounds grouped in the sub-clusters *I* (39 descendants) and *II* (31 descendants) (see Figure 3).

After the exclusion of compounds n-53 and 32/S individually and, later, at the same time, HCA procedures were carried out and the similarity indexes between the previous sub-clusters **I** and **II** were computed and evaluated\*. When compound 32/S was individually eliminated, the similarity index obtained (0.618) was identical to that obtained before the exclusions (0.617). Nevertheless, when only compound n-53 was excluded, the similarity index diminished from 0.617 to 0.464, indicating that n-53 interferes significantly in the behavior of the chalcone derivatives. A similar result was found when both compounds n-53 and 32/S were excluded (0.460). This finding suggests that compound n-53 has greater impact on the overall set behavior than compound 32/S, and it should be considered as the only outlier.

Optimization of the HCA and PCA procedures was performed by removing just compound n-53. The HCA results are shown in Figure 5. The new dendrogram reveals two major clusters. The first is composed of the sub-clusters **I**, **II**, **III**, **IV**, **V**, and **VI**, whereas the second was called cluster **VII** and is distinct from the rest of the investigated data. Sub-cluster **II** corresponds to the *N*-acylhydrazones main group while sub-cluster **III** represents the chalcones main group. As previously mentioned, the similarity index between **II** and **III** was 0.464 as shown in Figure 5.



**FIGURE 5** - Optimized dendrogram acquired after n-53 exclusion.

Sub-cluster **I** consists of seven *N*-acylhydrazones, containing more than one nitro group in its molecular structures (9/O, 9/S, 17/S, 23/O, 23/S, 24/O, and 24/S). The sub-clusters **IV** and **V** comprise chalcone derivatives that differ substantially from the chalcones of sub-cluster **III**. Finally, the dendrogram also revealed that compounds 32/O and 32/S did not share any kind of similarity neither

with chalcones nor with the other nitrated *N*-acylhydrazones, being classified as a distinct sub-cluster (**VI**). These compounds possess a sulfonylamide group in their molecular structures. This chemical group can attribute completely different physicochemical features to the structures, such as higher hydrophilic and ionization potential properties, which could be responsible for the distinct behavior of these ligands compared to the other ligands of the investigated set.

It is noteworthy that cluster **VII** is composed of one single ligand, n-65, which does not share any kind of similarity with the investigated set of compounds. Considering the calculated parameters, this derivative differs significantly from the others in terms of its torsional energy contribution ( $EL_{tors}$ ) (see Table A, Supplementary Material).

PCA methodology without the n-53 compound better elucidated the HCA results. The PCA procedure was run up to a maximum of seven PCs because when more than seven PCs were computed the total variance explained remained practically unchanged.

As shown in Table IV, the first three PCs described 62% of the total variance explained. The relative importance of each descriptor for the respective factor or PC (loadings) is also listed in this table.

The 3D scores plot confirmed the classification of the investigated set into two main separate subsets: *N*-acylhydrazones and chalcones. This fact, which can be visualized in Figure 5, is much more evident after the outlier exclusion, corroborating the substantial impact of compound n-53 on the set behavior. Moreover, the partitioning of the investigated set also indicates that chalcones and *N*-acylhydrazones do not share direct analogy or, in other words, do not show both structural and functional similarity, at least regarding the methodology applied in this study.

The clusters found with the HCA technique can also be observed in PCA results. For instance, sub-clusters **I**, **II**, **VI**, and cluster **VII** (compound n-65) are clearly discernible in view 1 of the 3D scores plot.

Parallel analysis considering the loading plot as well as the score (Table B, Supplementary Material) and loadings (Table IV), the tables revealed that the nitrated *N*-acylhydrazones biological behavior depends chiefly on thermodynamic and electronic descriptors, such as total energy ( $EL_{tot}$ ), bending energy ( $EL_{bend}$ ), total dipole moment ( $\mu_{tot}$ ), and the number of acceptor hydrogen bonding sites ( $H_{Accep}$ ).

It is important to consider that nitroderivatives can also exert their action by a nitro group reduction pathway (Orna, Mason, 1989), and this could be the mechanism highlighted in the methodologies applied here.



**TABLE IV** - Principal component analysis runs for 93 samples and 25 independent variables. The loadings of independent variables are listed below

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
<i>Variance %</i>	36.49	15.04	10.56	9.46	5.22	3.82	3.64
<i>Cumulative variance</i>	36.49	51.54	62.10	71.55	76.77	80.60	84.24
<i>EL<sub>bend</sub></i>	-0.07	0.34	0.11	-0.32	-0.07	0.26	0.28
<i>EL<sub>tors</sub></i>	0.04	0.33	-0.07	-0.01	-0.09	-0.26	-0.57
<i>EL<sub>1,4</sub></i>	0.10	-0.18	0.47	-0.13	0.00	-0.09	-0.21
<i>EL<sub>vdW</sub></i>	0.11	-0.09	0.00	0.22	0.25	0.68	-0.13
<i>EL<sub>et</sub></i>	-0.06	0.24	-0.41	0.09	0.12	0.16	0.33
<i>EL<sub>solv</sub></i>	0.26	-0.15	-0.15	-0.22	0.16	-0.08	-0.02
<i>EL<sub>tot</sub></i>	-0.09	0.22	0.41	-0.29	-0.02	0.04	0.02
<i>ChelpG 1</i>	0.05	0.18	-0.26	-0.39	0.17	0.12	-0.10
<i>ChelpG 2</i>	0.10	-0.06	0.14	0.50	0.06	-0.10	0.17
<i>ChelpG 3</i>	-0.01	0.28	-0.20	-0.10	0.24	-0.43	0.14
$\mu_x$	0.06	-0.22	-0.06	-0.11	-0.48	-0.02	0.34
$\mu_y$	0.02	0.29	0.06	0.29	0.22	-0.10	0.00
$\mu_z$	0.01	-0.07	0.25	-0.18	0.57	0.11	-0.02
$\mu_{tot}$	-0.18	0.24	0.20	0.30	0.06	-0.09	0.04
<i>MR</i>	0.29	0.19	0.03	-0.01	-0.10	0.11	0.02
$\alpha$	0.30	0.18	0.01	-0.05	-0.04	0.08	0.04
<i>SAGrid</i>	0.30	0.13	0.04	0.02	-0.15	0.07	-0.03
<i>ClogP<sub>Pallas</sub></i>	0.32	-0.02	0.02	0.01	0.13	-0.09	0.13
<i>LogD 2.0</i>	0.29	-0.01	0.00	0.11	0.16	-0.15	0.19
<i>LogD 7.0</i>	0.32	-0.03	0.03	-0.01	0.11	-0.10	0.14
<i>LogD 9.0</i>	0.29	-0.02	0.14	-0.07	0.07	-0.13	0.18
<i>I<sub>balaban</sub></i>	0.18	0.02	-0.29	0.12	-0.07	0.11	-0.35
<i>H<sub>Accep</sub></i>	-0.22	0.31	0.16	0.10	-0.04	0.09	0.11
<i>SA<sub>VDW</sub></i>	0.29	0.18	0.05	0.03	-0.17	0.08	-0.13
<i>SA<sub>SolvAccess</sub></i>	0.20	0.28	0.19	0.10	-0.24	0.10	0.02

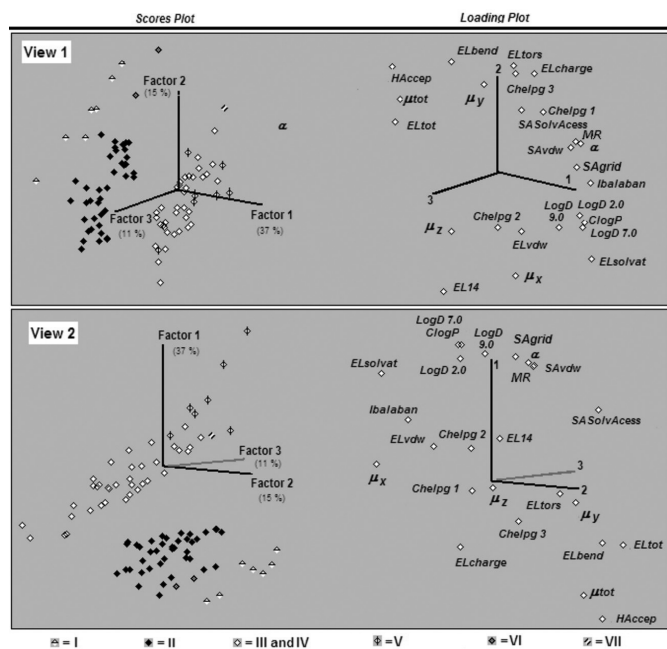
Non-nitrated *N*-acylhydrazones could be more suitable to demonstrate a common mechanism of action between the two classes of compounds investigated. However, as previously mentioned, a set of non-nitrated *N*-acylhydrazones assayed under the same biological conditions as the classes evaluated in this study were not available in the literature.

Nonetheless, if a second mechanism of action takes place and is similar to that of chalcones, descriptors of any other nature should be common to both classes.

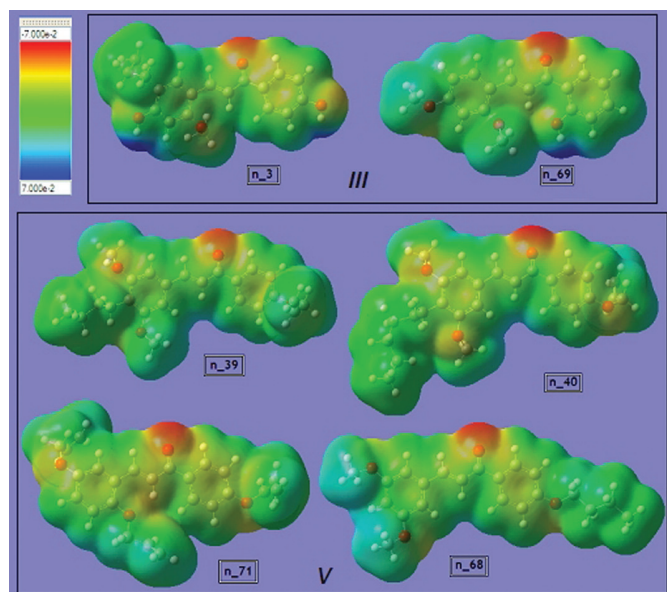
Chalcones, on the other hand, presented a broader distribution in PCs coordinates than did *N*-acylhydrazones (see View 2, Figure 6). The van der Waals energy (*EL<sub>vdW</sub>*) and electronic descriptors, such as dipole moment at the X coordinate ( $\mu_x$ ), and electrostatic potential partial atomic

charges at the carbonyl carbon (*ChelpG 1*) and oxygen (*ChelpG 2*), respectively, are related to chalcone biological behavior. These latter descriptors could indicate the importance of the carbonyl group to chalcone action and also support the cysteine protease inhibitory mechanism of action.

Interestingly, sub-cluster *V* of chalcones differs from the other chalcone derivatives by descriptors related to the molecular volume, such as van der Waals surface areas (*SA<sub>VDW</sub>*), polarizability ( $\alpha$ ), and molar refractivity (*MR*). All compounds in this group presented bigger or more branched side chains (Figure 1). In fact, the Connolly surfaces of these compounds can present volume differences compared to the other unsubstituted chalcone analogues (sub-cluster *III*) (Figure 7).



**FIGURE 6** - Scores and loadings on 3D plots found for the first three factors or PCs.



**FIGURE 7** - Electrostatic potential maps calculated on the Connolly surfaces of the chalcone derivatives from subclusters III and V. Negative electrostatic potential regions are represented in red (high electronic density) while positive electrostatic potential areas are shown in dark blue (low electronic density).

## CONCLUSIONS

The visual examination of the correlation between biological activity and calculated descriptor scatter plots

as well as HCA and PCA formalisms used, indicated that chalcones and nitrated *N*-acylhydrazones do not present functional similarity under the conditions applied in the present study.

Given that the two investigated classes have been reported as potential cysteine protease inhibitors, the hypothesis that these compounds inhibit cysteine proteases enzymes at different sites should be considered. Studies regarding field descriptors are now investigating the same library of compounds to verify if they could share similarities from the energy-interaction point of view, and should be helpful to verify the hypotheses raised by our findings.

## ACKNOWLEDGMENT

The authors would like to thank FAPESP (01/01192-3, 06/00116-5, 07/59416 and 07/5461-0) for the financial support, and the Chem21 Group, Inc., for the MOLSIM 3.2 academic license.

## REFERENCES

- BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. *Chemometrics: A practical guide*. New York: Wiley, 1998. 360 p.
- CHUNG, M. C.; FERREIRA, E. I.; SANTOS, J. L.; GIAROLLA, J.; RANDO, D. G.; ALMEIDA, A. E.; BOSQUESI, P. L.; MENEGON, R. N.; BLAU, L. Prodrugs for the treatment of neglected diseases. *Molecules*, v.13, p.616-677, 2008.
- CROFT, S. L. Public-private partnership: From there to here. *Trans. R. Soc. Trop. Med. Hyg.*, v.99, p.9-14, 2005.
- DOHERTY, D. *MOLSIM: Molecular Mechanics and Dynamics Simulation Software - User's Guide*, version 3.2. Lake Forest: The Chem21 Group Inc., 1997. CD-Rom
- FERREIRA, M. M. C. Multivariate QSAR. *J. Braz. Chem. Soc.*, v.13, p.742-753, 2002.
- GAUSSIAN 03W for Windows, version 6; Pittsburgh: Gaussian Inc., 1995-2003. CD-Rom
- HYPERCHEM Program. Versão 7.0 for Windows, Gainesville: Hypercube, Inc., 2002. CD-Rom
- KAYSER, O.; KIDERLEN, A. F.; CROFT, S. L. Natural products as antiparasitic drugs. *Parasitol. Res.*, v.90, p.S55-S62, 2003.

- LI, R.; CHEN, X.; GONG, B.; SEIZER, P. M.; LI, Z.; DAVIDSON, E.; KURZBAN, G.; MILLER, R. E.; NUZUM, E. O.; MCKERROW, J. H.; FLETTERICK, R. J.; GILLMOR, S. A.; CRAIK, C. S.; KUNTZ, I. D.; COHEN, F. E.; KENYON, G. L. Structure-based design of parasitic protease inhibitors. *Bioorg. Med. Chem.*, v.4, p.1421-1427, 1996.
- LI, R.; KENYON, G. L.; COHEN, F. E.; CHEN, X.; GONG, B.; DOMINGUEZ, J. N.; DAVIDSON, E.; KURZBAN, G.; MILLER, R. E.; NUZUM, E. O.; ROSENTHAL, P. J.; MCKERROW, J. H. *In vitro* antimalarial activity of Chalcones and their derivatives. *J. Med. Chem.*, v.38, p.5031-5037, 1995.
- LIU, M.; WILAIRAT, P.; GO, M. L. Antimalarial alkoxyated and hydroxylated chalcones: structure-activity relationship analysis. *J. Med. Chem.*, v.44, p.4443-4452, 2001.
- LIU, Z.; FANG, Q.; YU, W.; XUE, G.; CAO, D.; JIANG, M. 2'-Hydroxy-4''-dimethylamino-chalcone. *Acta Crystallogr. C*, v.C58, p.o445-o446, 2002.
- MARVINBEANS, Version 4.1.8; Budapest: ChemAxon Ltd, 1998-2007. CD-Rom
- MCBURNEY, B.; CRUNDWELL, G.; UPDEGRAFF III, J. B.; ZELLERB, M.; HUNTERB, A. D. 5-Nitrothiophene-2-carbaldehyde. *Acta Crystallogr. E*, v.E61, p.o1221-o1222, 2005.
- MCKERROW J. H.; ENGEL, J. C.; CAFFREY, C. R. Cysteine protease inhibitors as chemotherapy for parasitic infections. *Bioorg. Med. Chem.*, v.7, p.639-644, 1999.
- MCKERROW, J. H. Development of cysteine protease inhibitors as chemotherapy for parasitic diseases: insights on safety, target validation, and mechanism of action. *Int. J. Parasitol.*, v.29, p.833-837, 1999.
- NIELSEN, S. F.; CHRISTENSEN, S. B.; CRUCIANI, G.; KHARAZMI, A.; LILJEFORS, T. Antileishmanial Chalcones: Statistical design, synthesis, and three-dimensional quantitative structure-activity relationship analysis. *J. Med. Chem.*, v.41, p.4819-4832, 1998.
- ORNA, M. V.; MASON, R. P. Correlation of kinetic parameters of nitroreductase enzymes with redox properties of nitroaromatic compounds. *J. Biol. Chem.*, v.264, p.12379-12384, 1989.
- PALLAS, version 3.0, Sedona: CompuDrug Chemistry Ltd, 1994-1995. CD-Rom
- PIROUETTE, version 3.11, Woodinville: Infometrix Inc., 1990-2003. CD-Rom
- RACE, P. R.; LOVERING, A. L.; GREEN, R. M.; OSSOR, A.; WHITE, S. A.; SEARLE, P. F.; WRIGHTON, C. J.; HYDE, E. I. Structural and mechanistic studies of *Escherichia coli* nitroreductase with the antibiotic nitrofurazone: reversed binding orientations in different redox states of the enzyme. *J. Biol. Chem.*, v.280, p.13256-13264, 2005.
- RANDO, D. G.; AVERY, M. A.; TEKWANI, B. L.; KHAN, S. I.; FERREIRA, E. I. Antileishmanial activity screening of 5-nitro-2-heterocyclic benzylidene hydrazides. *Bioorg. Med. Chem.*, v.16, p.6724-6731, 2008.
- STEWART, J. J. P. Optimization of parameters for semi-empirical methods I-method. *J. Comput. Chem.*, v.10, p.209-220, 1989.
- TROEBERG, L.; CHEN, X.; FLAHERTY, T. M.; MORTY, R. E.; CHENG, M.; HUA, H.; SPRINGER, C.; MCKERROW, J. H.; KENYON, G. L.; LONSDALE-ECCLES, J. D.; COETZER, T. H. T.; COHEN, F. E. Chalcone, acyl hydrazide, and related amides kill cultured *Trypanosoma brucei brucei*. *Mol. Med.*, v.6, p.660-669, 2000.
- TROUILLER, P.; OLLIARO, P.; TORREELE, E.; ORBINSKI, J.; LAING, R.; FORD, N. Drug development for neglected diseases: a deficient market and a public-health policy failure. *Lancet*, v.359, p.2188-2194, 2002.
- VAN GUNSTEREN, W. F.; BERENDSEN, H. J. C. Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed.*, v.29, p.992-1023, 1990.
- WERMUTH, C. G. Similarity in drugs: reflections on analogue design. *Drug Discov. Today*, v.11, p.348-354, 2006.

Received for publication on 23<sup>rd</sup> July 2008  
Accepted for publication on 16<sup>th</sup> October 2009