5-Nitro-2-furfuriliden derivatives as potential anti-Trypanosoma cruzi agents: Design, synthesis, bioactivity evaluation, cytotoxicity and exploratory data analysis

ARTICLE *in* BIOORGANIC & MEDICINAL CHEMISTRY · JUNE 2013

Impact Factor: 2.79 · DOI: 10.1016/j.bmc.2013.06.017 · Source: PubMed

9 AUTHORS, INCLUDING:

Fanny Palace-Berl
University of São Paulo
9 PUBLICATIONS **50** CITATIONS
SEE PROFILE

Salomão Dória Jorge
University of São Paulo
12 PUBLICATIONS **81** CITATIONS
SEE PROFILE

Jose Ângelo L Lindoso
Instituto de Infectologia Emílio Ribas
26 PUBLICATIONS **430** CITATIONS
SEE PROFILE

Leoberto Costa Tavares
University of São Paulo
51 PUBLICATIONS **634** CITATIONS
SEE PROFILE

Bioorganic & Medicinal Chemistry 21 (2013) 5395–5406

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

5-Nitro-2-furfuriliden derivatives as potential anti-*Trypanosoma cruzi* agents: Design, synthesis, bioactivity evaluation, cytotoxicity and exploratory data analysis

Fanny Palace-Berl [a,*], Salomão Dória Jorge [a], Kerly Fernanda Mesquita Pasqualoto [b], Adilson Kleber Ferreira [b,c], Durvanei Augusto Maria [c], Rodrigo Rocha Zorzi [a], Leandro de Sá Bortolozzo [a], José Ângelo Lauletta Lindoso [d], Leoberto Costa Tavares [a]

[a] *Department of Biochemical and Pharmaceutical Technology, Faculty of Pharmacy, University of São Paulo, Av. Prof. Lineu Prestes, 580, São Paulo, SP 05508-000, Brazil*
[b] *Biochemistry and Biophysics Laboratory, Butantan Institute, SP, Brazil*
[c] *Experimental Physiopathology, Faculty of Medicine, University of Sao Paulo, Sao Paulo, Brazil*
[d] *Laboratory of Seroepidemiology, LIM 38-HC FMUSP, Institute of Tropical Medicine and Institute of Infectology Emílio Ribas, SP, Brazil*

ARTICLE INFO

Article history:
Received 25 March 2013
Revised 29 May 2013
Accepted 6 June 2013
Available online 15 June 2013

Keywords:
Trypanosoma cruzi
5-Nitro-2-furfuriliden
Nifuroxazide
Azomethine derivatives
Oxadiazoline derivatives
Molecular modeling

ABSTRACT

The anti-*Trypanosoma cruzi* activity of 5-nitro-2-furfuriliden derivatives as well as the cytotoxicity of these compounds on J774 macrophages cell line and FN1 human fibroblast cells were investigated in this study. The most active compounds of series **I** and **II** were 4-butyl-[*N*′-(5-nitrofuran-2-yl) methylene] benzidrazide (**3g**; IC_{50} = 1.05 μM ± 0.07) and 3-acetyl-5-(4-butylphenyl)-2-(5-nitrofuran-2-yl)-2,3-dihydro,1,3,4-oxadiazole (**4g**; IC_{50} = 8.27 μM ± 0.42), respectively. Also, compound **3g** was more active than the standard drugs, benznidazole (IC_{50} = 22.69 μM ± 1.96) and nifurtimox (IC_{50} = 3.78 μM ± 0.10). Regarding the cytotoxicity assay, the **3g** compound presented IC_{50} value of 28.05 μM (SI = 26.71) against J774 cells. For the FN1 fibroblast assay, **3g** showed IC_{50} value of 98 μM (SI = 93.33). On the other hand, compound **4g** presented a cytotoxicity value on J774 cells higher than 400 μM (SI >48), and for the FN1 cells its IC_{50} value was 186 μM (SI = 22.49). Moreover, an exploratory data analysis, which comprises hierarchical cluster (HCA) and principal component analysis (PCA), was carried out and the findings were complementary. The molecular properties that most influenced the compounds' grouping were *C*log*P* and total dipole moment, pointing out the need of a lipophilic/hydrophilic balance in the designing of novel potential anti-*T. cruzi* molecules.

1. Introduction

Chagas disease (CD) is a chronic and debilitating parasitic infection caused by the intracellular protozoan *Trypanosoma cruzi*. CD, also called American Trypanosomiasis, affects around 8–10 million people worldwide, mainly in Latin America.[1] However, as a result of the influx of immigrants, for instance, CD is increasingly being reported in several non-endemic areas of Europe and North America as well, becoming a serious global problem.[1,2]

Currently there are only two drugs available for the treatment of CD, benznidazole (BZ) and nifurtimox (NFX) (Fig. 1A), which are effective in the acute phase and initial chronic phase of the infection, but not in its late stage.[2–4] Chemotherapy can shorten the acute phase of disease and may achieve the parasitological cure in 50–80% cases. However, for the initial chronic phase the expectative of cure decreases to 20–60%. Moreover, due to the side effects associated with BZ and NFX, patients under treatment require a careful monitoring.[2]

Considering the global incidence of CD and the lack of novel therapeutic alternatives, it is still necessary the search for new anti-*T. cruzi* agents to improve the cure rate and reduce the adverse effects on chagasic patients. Many classes of compounds are being investigated as anti-trypanocidal, and among them are the nitroheterocyclic compounds.[4,5] In this regard, our research group focused the investigation on nifuroxazide (NF), 4-hydroxy-*N*′-(5-nitro-2-furfuryliden) benzhydrazide (Fig. 1B), as a lead compound, since it has shown quite promising biological responses by considering molecular modifications on its chemical structure.[6,7] The NF mechanism of action is not fully elucidated, but previous studies have suggested that its antimicrobial activity would be related to the reduction of the nitro group and formation of free radical toxic species.[8,9] Then, in this study, two sets (series I and II, Fig. 1C) of NF analogues were designed envisaging the identification of the most active compound against *T. cruzi*.

* Corresponding author. Tel.: +55 11 30913693x38156386, mobile: +55 11 998673742.
E-mail address: palaceberlf@usp.br (F. Palace-Berl).

0968-0896/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2013.06.017

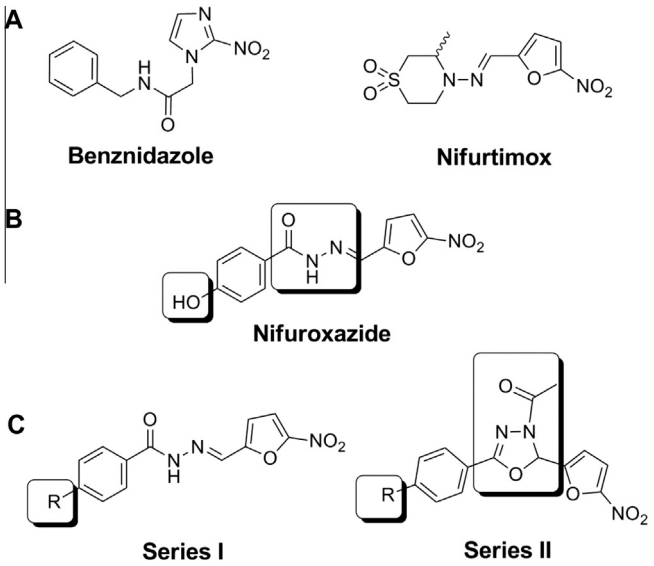


Figure 1. (A) Chemical structures of the drugs used in the CD treatment (benznidazole and nifurtimox). (B) Chemical structure of lead compound (nifuroxazide). (C) Chemical structures of the azomethine (series I) and 3-acetyl-2,5-disubstituted-2,3-dihydro-1,3,4-oxadiazole derivatives (series II), pointing out the regions where chemical modifications were carried out in the present study.

The structure of NF was maintained to design the series **I**, and molecular modifications were performed at position 4 of the benzene ring, and consequently the phenolic hydroxyl group was replaced by different substituents. The same kind of substitution was also performed to the compounds of series **II**. The substituents' selection was based upon the Craig's diagram, a rational method where the physicochemical properties (hydrophobicity and electronic) are interrelated.[10]

The series **II** was planned primarily considering the cyclization of the central portion of the compounds from series **I** (azomethynic group) to form the 3-acetyl-2,5-disubstituted-2,3-dihydro-1,3,4-oxadiazoline ring. Besides the improvement of pharmacokinetics efficacy due to the ability to establish hydrogen bonding interactions, the oxadiazoline ring can provide important structural features to the molecules, as conformational changes, for instance. Recently, several 3-acetyl-2,5-disubstituted-2,3-dihydro-1,3,4-oxadiazoline ring systems have been associated with diverse biological activities, that is, antibacterial, antifungal,[6,11] anti-*T. cruzi*,[6] anticancer.[12] They have also been reported as important ring systems for Parkinson disease regarding the inhibitory activity on MAO-B.[13]

Furthermore, cytotoxicity assays on mammalian cells, using J774 murine macrophages,[14,15] and on FN1 human fibroblast cells[16] were performed to evaluate the investigated compounds.

Molecular modeling methods were also applied, herein, to select an energetically favourable conformation of each compound and, then, calculate the molecular properties of different nature (electronic, steric, hydrophobic, topological, geometric), which were used to carry out the exploratory data analysis (hierarchical cluster analysis, HCA, and principal components analysis, PCA).[17,18] Exploratory analysis allows the investigation of intersamples relationships through either similarity indices, or linear combinations from the original data. The findings can be used to establish some qualitative structure–activity relationships, since the calculated molecular properties are directly related to the compounds' chemical structures.

2. Results and discussion

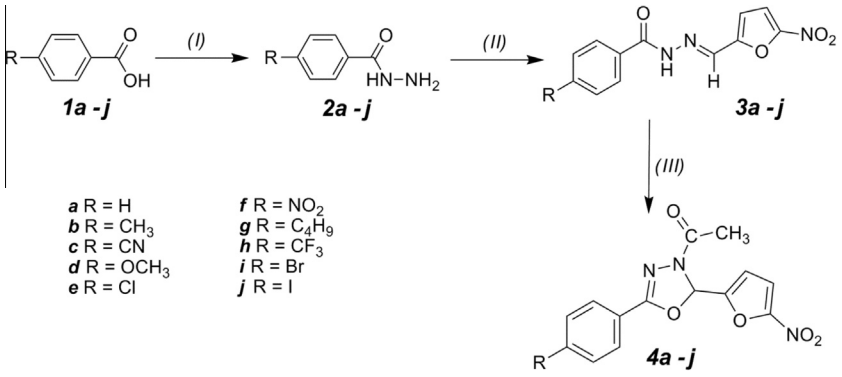
2.1. Chemistry

The two designed sets of 5-nitro-2-furfuriliden derivatives were obtained as shown in Scheme 1. The substituent groups were selected according to the influence of their physicochemical properties (electronic (σ) and hydrophobic (π) character) employing the Craig diagram.[10]

The benzhydrazides (**2a**–**j**) were obtained from benzoic acids through consecutive reaction of esterification and hydrazinolysis, without isolation of the corresponding methyl esters.[19] Furthermore, the azomethine derivatives of series **I** (**3a**–**j**) were obtained by nucleophilic addition of the derivatives **2a**–**j** with 5-nitro-2-furaldehyde diacetate.[20] The 3-acetyl-2,5-disubstituted-2,3-dihydro-1,3,4-oxadiazole derivatives of series **II** (**4a**–**j**) were obtained by oxidative cyclization of azomethine derivatives (**3a**–**j**) with acetic anhydride.[6,11,21]

The hydrazinolysis reactions provided satisfactory yields for all benzhydrazides (values between 81% and 97%). The synthetic procedure was modified to obtain 4-nitrobenzhydrazide, **2f**. The reactions were performed in a controlled temperature (−3 to 2 °C) to avoid the reduction of the compounds by the hydrazine hydrate, which acts as a reducing agent.[19,22]

The synthesis of the azomethynic series (**3a**–**j**) presented yield values between 89% and 99%, as already reported in previous studies.[20] The identification of compounds **3a**–**j** was carried out through the comparison of the experimental melting point values to those previously reported in the literature (see Supplementary data section) as well as through the ^{1}H and ^{13}C NMR spectroscopy and elemental analysis (CNH). The 3-acetyl-2,5-disubstituted-1,3,4-oxadiazoline derivatives (**4a**–**j**) obtained yields between



Scheme 1. Synthetic route for series I (**3a**–**j**) and II (**4a**–**j**). Reaction conditions: (I): CH_3OH, H_2SO_4/reflux/4 h; N_2H_4 80% aq/reflux/30 min; (II) 5-nitro-2-furaldehyde diacetate, H_2O, H_2SO_4, CH_3COOH, C_2H_5OH/reflux/1 h; (III) $(CH_3CO)_2O$/110–120 °C/12 h.

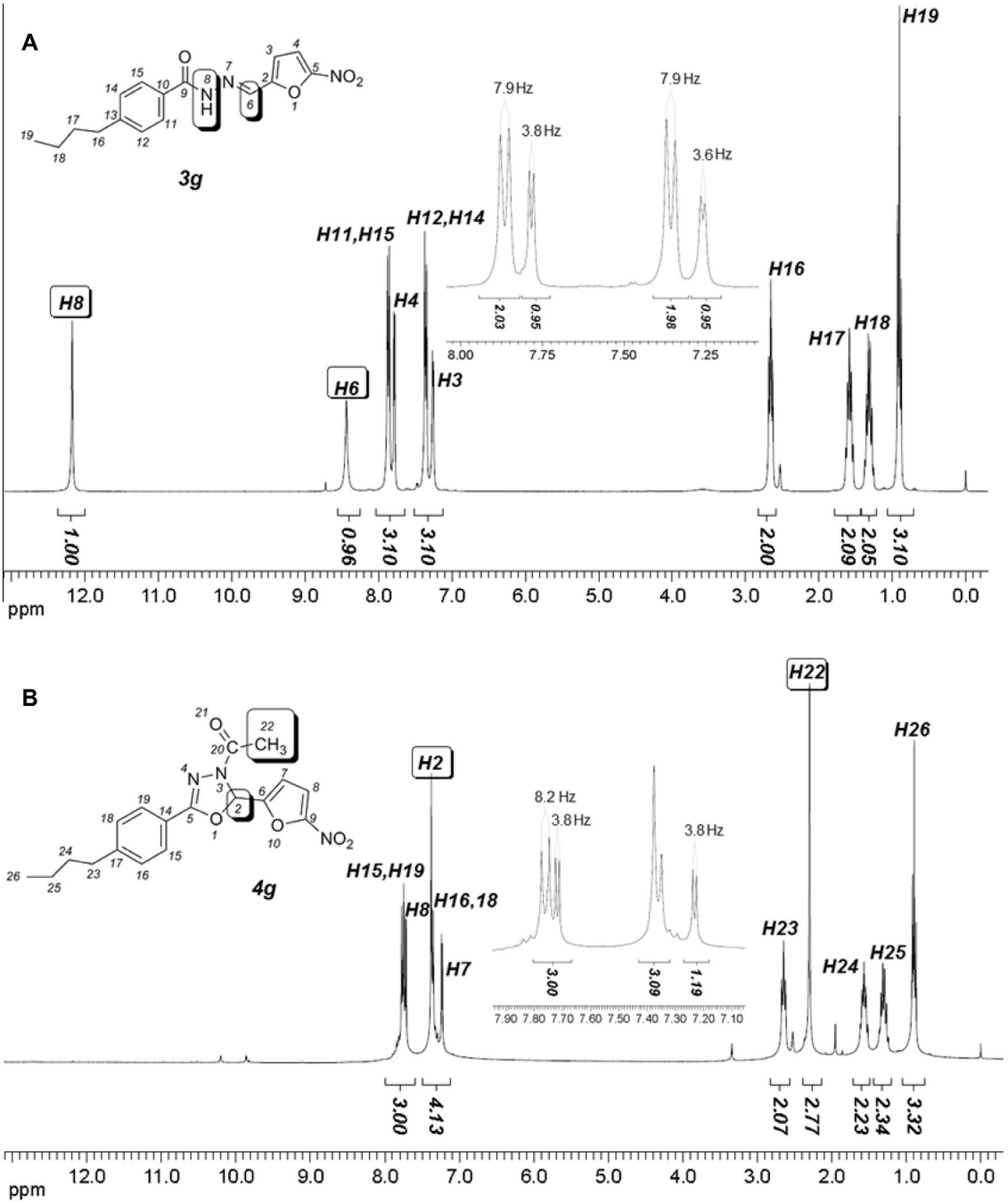


Figure 2. ^{1}H NMR spectra (DMSO-d_6, 300 MHz) of compounds **3g** and **4g**. Chemical shifts (δ) highlighting the differences between the azomethynic group and oxadiazolinic ring: (A) compound **3g**—H8 (12.17 ppm) and H6 (8.44 ppm); (B) compound **4g**—H2 (7.37 ppm) and H22 (2.30 ppm). NMR attributions of compounds were done according to the numbered chemical structure.

52% and 86%. During the synthesis procedure, thin layer chromatography (TLC) analysis was performed to verify the formation of 3-acetyl-2,5-disubstituted-1,3,4-oxadiazolines.[6]

The structural elucidation of these derivatives was confirmed through ^{1}H and ^{13}C NMR spectra analysis, considering chemical shifts (δ) related to either the residual solvent peak, or the internal standard. Compounds of the series **I** showed characteristic NMR spectra, as illustrated in Figure 2A (^{1}H NMR for the **3g** compound). The presence of two singlet signals were observed around δ 12 ppm and δ 8 ppm, respectively, indicating the protons of the azomethine group (see Fig. 2A; H8 and H6). For the compounds of series **II**, two characteristic peaks were found, as displayed in Figure 2B. The ^{1}H NMR spectrum of compound **4g** presented one signal around δ 7.4 ppm, evidencing the formation of oxadiazoline ring, and another singlet signal near δ 2.3 ppm, which confirms the three protons of the acetyl group (see Fig. 2B; H2 and H22, respectively). Analyzing the ^{13}C NMR spectra of compounds from series **I**, the signals were near 162 and 135 ppm, respectively, indicating the presence of the azomethine group. The compounds of series **II** displayed a characteristic peak around 21 ppm related to the presence of the acetyl group in the oxadiazoline ring. Moreover, elemental analysis (CNH) was carried out, and taken together with ^{1}H and ^{13}C NMR, the data were in agreement with the proposed structures (see Supplementary data, p. S2 and S4).

Table 1
Anti-*T. cruzi* activity of 5-nitro-2-furfuriliden derivatives, **3a–j** and **4a–j**, benznidazole, nifurtimox and nifuroxazide

Compounds (R)		IC_{50}[a] (μM)	Compounds (R)		IC_{50}[a] (μM)
3a	H	4.44 ± 0.48	**4a**	H	10.22 ± 0.41
3b	CH_3	3.71 ± 0.14	**4b**	CH_3	9.84 ± 0.59
3c	CN	12.51 ± 0.94	**4c**	CN	13.44 ± 0.80
3d	OCH_3	3.94 ± 0.36	**4d**	OCH_3	9.26 ± 1.02
3e	Cl	4.03 ± 0.23	**4e**	Cl	11.46 ± 0.81
3f	NO_2	9.45 ± 0.54	**4f**	NO_2	13.91 ± 0.95
3g	C_4H_9	1.05 ± 0.07	**4g**	C_4H_9	8.27 ± 0.42
3h	CF_3	4.42 ± 0.14	**4h**	CF_3	11.14 ± 1.07
3i	Br	3.74 ± 0.15	**4i**	Br	9.59 ± 0.31
3j	I	3.21 ± 0.19	**4j**	I	11.52 ± 0.99
Benznidazole		22.69 ± 1.96			
Nifurtimox		3.59 ± 0.46			
Nifuroxazide		120.46 ± 4.06			

IC_{50}: half maximal inhibitory concentration.
[a] Values corresponding to the average of triplicates.

It is noteworthy that, among the twenty compounds synthesized and identified, herein, there are five new compounds, and all from series **II** (**4c**, **4g–4j**) (see ^1H and ^{13}C NMR spectra in Supplementary data, p. S6).

2.2. Anti-*T. cruzi* activity assays

The biological activity of the designed compounds was evaluated through the percentage of growth inhibition of the epimastigote forms of *T. cruzi* (Y strains), according to the experimental protocols. Results were expressed as IC_{50} values, which represent the concentration of compound necessary to inhibit fifty percent of the parasite growth. Parasites in the stationary phase were incubated, during 5 days, increasing the concentration of the investigated compounds from series **I** and **II**, reference drugs (NFX and BZN) and lead compound (NF). All assays were performed using DMSO as control, at the same concentration employed to solubilise the compounds and reference drugs, and which did not exceeded 1.0%.[14,23] No effect on the epimastigotes' growth was observed in the culture media by the presence of up to 1.0% DMSO. The findings are presented in Table 1 (see also Supplementary data, p. S16).

Accordingly, the most active compounds of each series were the 4-C_4H_9 (**3g** and **4g**) substituted derivatives while the lowest activity was observed for the 4-CN (**3c** and **4c**) and 4-NO_2 (**4f** and **5f**) substituted derivatives. Comparing the two series, compounds of series **I** were more active than those from series **II**. Even so, all compounds showed higher activity in comparison to the reference drug BZ. The compound **3g** presented threefold higher activity than NFX (**3g**, IC_{50} = 1.05 μM ± 0.07; NFX, IC_{50} = 3.78 μM ± 0.10). The lead compound, NF, did not show significant activity against *T. cruzi* (NF, IC_{50} = 120.46 μM ± 4.06), evidencing that the molecular modifications performed in the phenyl and azomethynic group generated satisfactory results.

2.3. Cytotoxicity assays

Compounds presenting high, moderate and low biological activity against epimastigote forms were elected for the cytotoxicity tests using mice macrophages and human fibroblast cells, and the findings are presented in Table 2.

Regarding series **I**, compounds **3g** and **4g** displayed higher selectivity indices (SI). Also, the results pointed out the compounds from series **I** as more cytotoxic than those from series **II**. The J774 mice macrophages were chosen as a model to obtain an unspecific mammalian cytotoxicity, since the macrophages are host cells for the parasite. Moreover, it was performed cytotoxicity assays using FN1 human fibroblast cells to determine the in vitro toxicity of the synthesized compounds against human normal cells (not infected). Compounds **3e**, **3g** and **4f** showed lower cytotoxicity for human fibroblasts than for mice macrophages cells. Compound **3g** also presented an interesting SI value regarding both cell lines (SI for J774 cells = 27.58; SI for FN1 cells = 93.33). The SI value for compounds **3e**, **4e** and **4f** was not exactly determined because the maximum concentration considered in the assay for those compounds was not toxic.

2.4. Molecular modeling approach, calculation and selection of descriptors

The three-dimensional (3D) molecular models of compounds were built up and energy-minimized as reported in Section 4. The lowest-energy conformer of each reference template, NF (series **I**) and compound **4a** (series **II**), was selected from the equilibrium region of the conformational ensemble profile (CEP) from molecular dynamics (MD) simulations (see Section 4.4 and Supplementary data p. S18, Figs. 1S and 2S). The root mean square deviation (RMSD) values between the atomic positions of the selected conformers from CEP and of the energy-minimized structures were computed as a criterion to verify whether the structural integrity was maintained after simulations. Higher RMSD values indicate loss of structural integrity during MD simulations if there is considered a crystallographic structure as template or reference.[24]

The RMSD value found between the lowest-energy conformer and the initial energy-minimized structure was 0.07 and 0.14 Å for NF and compound **4a**, respectively. The higher RMSD value for compound **4a** was already expected since the molecular model of this compound was built up using different crystallographic fragments not an entire experimental structure (X-ray diffraction or RMN). However, regarding the main portion of molecular structure (5-nitrofuran moiety), for instance, the RMSD value was

Table 2
Cytotoxicity of 5-nitro-2-furfuriliden derivatives (**3e**, **3f**, **3g**, **4e**, **4f** and **4g**), benznidazole and nifurtimox against J774 mice macrophages, FN1 human fibroblast cells and their respective selectivity index

Compounds (R)		*T. cruzi* IC_{50}[a] (μM)	J774 cells IC_{50}[a] (μM)	SI[b]	FN1 cells IC_{50}[a] (μM)	SI[b]
3e	Cl	4.03 ± 0.23	57.11 ± 1.29	14.17	>200	>49.63
3f	NO_2	9.45 ± 0.54	79.09 ± 7.89	8.37	17,800 ± 11.93	18.83
3g	C_4H_9	1.05 ± 0.07	28.96 ± 1.05	27.58	9800 ± 4.72	93.33
4e	Cl	11.46 ± 0.81	>400	>34	>200	>17
4f	NO_2	13.91 ± 0.95	114.27 ± 2.05	8.21	>200	>14
4g	C_4H_9	8.27 ± 0.43	>400	>48	18,600 ± 9.50	22.49
Benznidazole		22.69 ± 1.96	>400	>18	>200	>9
Nifurtimox		3.78 ± 0.10	280.48 ± 2.44	74.2	>200	>53

IC_{50}: half maximal inhibitory concentration.
[a] Values corresponding to the average of triplicates.
[b] SI—selectivity index (SI = IC_{50} of cells line/IC_{50} of *T. cruzi*).

0.07 Å, reinforcing the structural integrity at the pharmacophore group region. Additionally, the selected conformer for NF and for compound **4a** presented total potential energy value of 19.90 and 30.86 kcal/mol, respectively, considering the intramolecular energy contribution of solvation (E_{solv}) and of hydrogen bonding (E_{Hb}) (see Supplementary data, p. S18, Table 1S).

Then, forty descriptors or molecular properties of different nature were calculated (thermodynamic, hydrophobic, electronic, topological, geometric, and steric) for each compound. A matrix composed of 21 rows (samples or compounds plus NF) and 41 columns (independent variables or descriptors or molecular properties plus biological activity) was generated and used as input for the exploratory data analysis (HCA and PCA). All calculated descriptors and their methods are shown in Supplementary data (p. S19, Table 2S; p. S21, Table 3S).

The biological data (anti-*T. cruzi* activity) was transformed into potency ($\log 1/IC_{50}$ or $-\log IC_{50}$ or pIC_{50}) and corresponds to the dependent variable. Compounds of the two series were distributed according to a gradient of biological activity as follows: less, moderately, and highly active compounds. It was considered each series separately to emphasize the substituent groups' contribution during the analysis. For series **II**, **4c** and **4f** were considered as less active compounds (pIC_{50} values of 4.86 and 4.87); **4a**, **4b**, **4d**, **4e**, **4g–j** as moderately active compounds (pIC_{50} values from 4.87 to 5.03), and **4g** was the most active compound of series **II** (pIC_{50} value of 5.08). Similar distribution of biological activity was observed for series **I**, where **3c** and **3f** were considered less active compounds (pIC_{50} values of 4.90 and 5.02); **3a**, **3b**,**3d**, **3e**, **3g–j** as moderately active compounds (pIC_{50} values from 5.35 to 5.49); and **3g** as the most active compound (pIC_{50} value of 5.98).

Before proceeding the exploratory data analysis, a preliminary variable selection was carried out using as criteria (i) the Pearson's linear correlation coefficient between descriptors and inhibitory activities (cut off value $R = 0.2$), and (ii) a visual inspection of scatter plots for the biological activity (pIC_{50} values) versus each calculated molecular property or descriptor.[17] These findings are presented in Table 4S (Supplementary data, p. S23). Ten calculated descriptors showed uniform distribution/dispersion with linear tendency, and they were: the intramolecular energy of solvation (E_{solv}); the electrostatic potential atomic charge (ESP or CHELPG) of nitrogen in the N=C bond for series **I** (N_7) and nitrogen in oxadiazolinic ring for series **II** (N_4) (CHELPG_$N_{7,4}$), using the HF/3-21G* method;[25,26] total dipole moment (μ); the solvent accessible surface area of all hydrophobic atoms (ASA_h);[27] the STERIMOL parameter L defined as the length of a substituent along the axis of the bond between the first atom of the substituent and the parent molecule;[28] the calculated *n*-octanol/water partition coefficient applying the Ghose et al., 1998-SLN method[29] ($C\log P_{SG\text{-}SLN}$), the weighted method using Viswanadhan et al.,1989,[30] Klopman et al., 1994[31] and PHYSPROP© databases[32] ($C\log P_{WM}$), and the Viswanadhan et al., 1989[30] method ($C\log P_{SV}$); molar refractivity (MR);[33] and the hydrophobic Hansch constant (π).[34] The final matrix was used as input for the exploratory data analysis. The data matrix was autoscaled because the calculated independent variables presented different magnitude orders.

2.5. Exploratory data analysis

As already mentioned, exploratory data analysis is a multivariate method of classification, which comprises hierarchical cluster analysis (HCA) and principal components analysis (PCA). HCA is an unsupervised pattern recognition analysis, where the distance between similar samples is smaller than for unrelated samples, and classifies the results according to similarity indices that range from zero to one. The similarity matrix generates a dendrogram, which is a tree-diagram constructed from the Euclidean distances data (see equation in 4.5.1 Hierarchical cluster analysis section).[17,18,35] Herein, HCA analysis was performed for (i) a matrix considering the biological activity as dependent variable allowing the grouping of compounds or samples (see Fig. 3), and (ii) another matrix considering the biological activity also as independent variable to verify intervariables relationships (see Fig. 4).

Regarding the dendrogram of samples (Fig. 3), the compounds were split into three main clusters: A (45.2% similarity), B (52.0% similarity), and C (53.9% similarity), using a similarity cursor of 0.4 (dashed line) as reference. Clusters A and B share 35% of similarity whereas cluster C has no similarity with those two.

The compounds were grouped, in general, accordingly to their substituent groups and, then, subdivided in each series (see cluster A and B). Considering the gradient of biological activity [less (1), moderately (2) and highly (3) active compounds of each series], cluster A is contained of moderately active compounds from series **II** and **I**. Also, it presents more lipophilic compounds and has two subclusters, regarding each series: A' (**4i -**Br, **4j -**I, **4e -**Cl, **4b –** CH_3, and **4h –**CF_3) and A'' (**3j -**I, **3i -**Br, **3e -**Cl, **3h –**CF_3, **3b –**CH_3, and **3d** –OCH_3), which share 71.7% and 59.1% of similarity, respectively. The values of Hansch constant (π) for the substituent groups in subclusters A' and A'' range from 0.56 to 1.12 and from -0.02 to 1.12, respectively, evidencing a more hydrophobic contribution to those compounds. Cluster B have grouped moderately and less active compounds and it was subdivided into subclusters B' (**4f** –NO_2, **4d** –OCH_3, **4c –**CN, and **4a -**H) and B'' (**3f** –NO_2, **3c –**CN, **3a –**H, and NF), which displayed similarity indices of 0.58 and 0.64, respectively. The NO_2 and CN substituent groups present more negative values for the Hansch constant, π equal to -0.28 and -0.57, meaning a more hydrophilic contribution. The lead compound, NF, was put together into cluster B'' along with the non-substituted moderately active compound, **3a**, probably due to the structural resemble with compounds from series **I** as well as the character more hydrophilic of its substituent group (-OH, $\pi = -0.67$). The most active compounds [**3g** and **4g**, *n*-butyl as substituent group, $\pi = 2.13$], regarding the both series, were grouped into cluster C and share a similarity index of 0.54.

The dendrogram of variables or descriptors (see Fig. 4) showed that the electronic molecular properties, μ and CHELPG_$N_{7,4}$, are more close to the biological activity and grouped in the same cluster. The cluster shares 37.0% of similarity, and CHELPG_$N_{7,4}$ and μ display a similarity index of 0.74. Dipole moment, μ, is a global molecular property, which is related to the molecule's polarity. Based upon this property, the compounds of series **I** are more polar (hydrophilic) than those from series **II**. The average value of μ found for series **I** (including NF) was 12.68 whereas for series **II** was 6.75. ESP or CHELPG of nitrogen in the N=C bond for series **I** (N_7) ranged from -0.3297 to -0.3932 whereas for the corresponding nitrogen in series **II** (N_4 in the oxadiazolinic ring) varied from -0.5099 to -0.5971.

The unsupervised principal components analysis (PCA) aimed to correlate the descriptors to find other uncorrelated variables, the principal components (PCs), which can describe the information from the original data set.[17,18]

The PCA findings are showed in Fig. 5. According to the factors selection, the two first factors or PCs discriminated more than seventy percent (76.68%) of total variance from the original data. Also, regarding the scores' plot (Fig. 5), complementary findings were obtained for the both methods, PCA and HCA, since the compounds were split in the same previous groups (A', blue rectangle; A'', cyan rectangle; B', light pink square; B'', dark pink square; and C, yellow oval circle). Interestingly, PC2 or factor 2 (26.37%) seems to be responsible for split the compounds into the two series, **I** and **II** (horizontal gray line in the scores' plot). Additionally, the most active compounds of each series, **3g** and **4g**, were a little bit more distant from the rest of data set and from one another, evidencing the

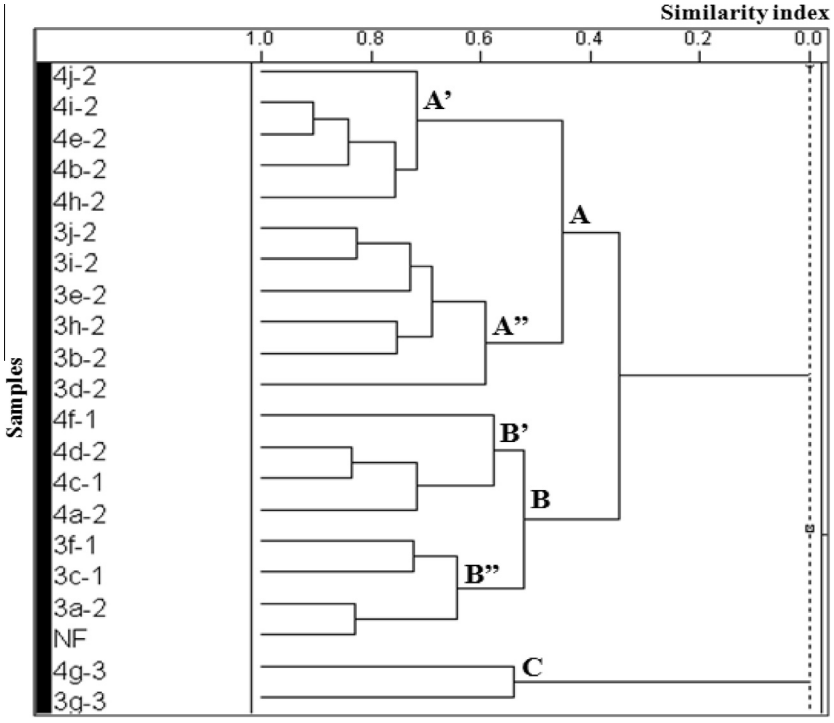


Figure 3. HCA of samples: considering a similarity cursor of 0.4 (dashed line), the compounds were grouped into three main clusters: A (moderately active compounds; 45.2% similarity), B (less and moderately active compounds; 52.0% similarity), and C (highly active compounds (53.9% similarity); *n*-butyl substituent group). Numbers 1, 2, or 3 were added after the name of each compound to define its gradient of biological activity as less, moderately, or highly active, respectively.

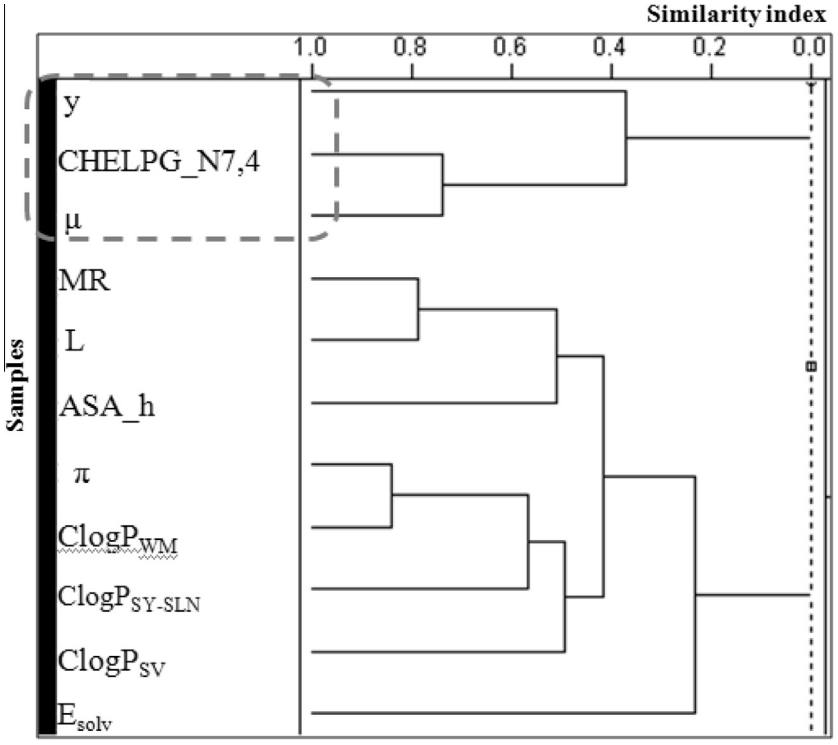


Figure 4. HCA of variables: the electronic descriptors grouped in the same cluster of biological activity were CHELPG_$N_{7,4}$ and μ. They are highlighted in a light gray dashed rectangle.

two series classification. Regarding the HCA of samples, these two compounds share 54% of similarity.

The loadings' table, also presented in Fig. 5, provides information about which descriptors or molecular properties most influenced the separation procedure of samples/compounds. In PC1, or factor 1 (50.31%), hydrophobic ($C\log P_{SG\text{-}SLN}$, $C\log P_{SV}$, $C\log P_{WM}$, π), steric/hydrophobic (MR), steric/geometric (ASA_h) and steric (L) molecular properties displayed higher loading values (absolute

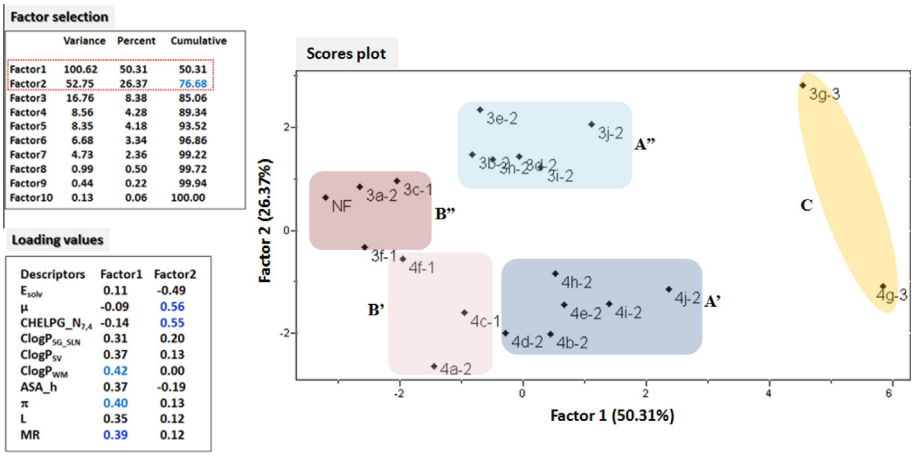


Figure 5. PCA findings: factor selection, scores' plot and loadings' table. In the scores' plot, the clusters observed in HCA are highlighted as blue rectangle (A′), cyan rectangle (A″), light pink square (B′), dark pink square (B″), and yellow oval circle (C).

Table 3
Calculated values found for the two molecular properties that displayed the highest loading value in each PC or factor

Compounds	R	ClogP_{WM} (PC1)	μ (Debye) (PC2)
NF	OH	1.75	13.19
3a-2	H	2.05	14.52
3b-2	CH_3	2.57	15.06
3c-1	CN	1.91	10.99
3d-2	OCH_3	1.89	13.77
3e-2	Cl	2.66	12.76
3f-1	NO_2	1.99	9.80
3g-3	**C_4H_9**	**3.90**	**15.15**
3h-2	CF_3	2.93	11.07
3i-2	Br	2.82	11.46
3j-2	I	2.98	11.74
4a-2	H	2.36	5.57
4b-2	CH_3	2.87	6.72
4c-1	CN	2.21	7.22
4d-2	OCH_3	2.20	6.20
4e-2	Cl	2.96	6.40
4f-1	NO_2	2.30	9.34
4g-3	**C_4H_9**	**4.21**	**6.75**
4h-2	CF_3	3.24	6.61
4i-2	Br	3.13	6.30
4j-2	I	3.29	6.41

The calculated partition coefficient (*n*-octanol/water) ClogP_{WM} was obtained applying the weighted method using Viswanadhan et al., 1989, Klopman et al., 1994 and PHYSPROP© databases.[30-32]
The values of total dipole moment (μ) were calculated employing the HF/3.21G* method.[25,26]
Numbers 1, 2, or 3 were added after the name of each compound to define its gradient of biological activity as less (1), moderately (2), or highly active (3), respectively. Clog *P* and dipole moment of the most active 5-nitro-2-furfuriliden derivatives are in bold.

values). ClogP_{WM} and π, for instance, presented the highest loading values, 0.42 and 0.40, respectively, reinforcing the discrimination pattern obtained, which was closely the same observed in HCA. In PC2, μ and CHEPLG_$N_{7,4}$, which are electronic molecular properties, showed the highest loading values, 0.56 and 0.55, respectively. Regarding the investigated compounds, the calculated values for the two molecular properties presenting the highest loading values in each PC, or factor, are shown in Table 3.

It is noteworthy that the Clog*P* and μ values found for the most active compound of each set, **3g** and **4g**, could be considered as the optimum values to the anti-*T. cruzi* activity for these investigated sets, for instance. The μ measures the delocalized charge when atoms with different electronegativity are interconnected. Considering that, substituent groups with different electronegativity can change the molecular properties of a compound, and consequently, can influence its activity.[36] This kind of influence was observed in studies based on the physicochemical properties of NFX analogues aiming to explore substituent groups in the opposite side of the nitro-furan group. The μ property was directly related to the anti-*T. cruzi* activity, meaning more reactive molecules had larger values of total dipole moment.[37] As we know, hydrophobic properties, such as Clog*P*, are quite important to the diffusion of compounds through the biological system. Additionally, the environment of a binding site in enzymes or receptors is generally hydrophobic and the partition complementarity with the respective ligand is crucial for the biological response. In this regard, the compounds of series **I** and **II** with higher Clog*P* values presented also higher activity values. Therefore, the hydrophobic/hydrophilic balance and electronic density distribution were pointed out as important molecular properties for the biological activity in some studies with similar compounds, which had the mechanism of action by reduction.[37–39]

The outliers' diagnosis, implemented in Pirouette 3.11 software,[40] was also performed through the Mahalanobis distance,[41] as can be seen in Fig. 6. The sample residual threshold was based upon a ninety-five percent of confidence level interval set internally in Pirouette 3.11.[40] The compounds did not exceed the considered threshold, meaning the calculated properties were sufficient to describe the structural features of the both data sets (see Fig. 6).

The lipophilic and electrostatic potential maps (MLP and MEP) were calculated onto the molecular surface of the most active compound of each set to visualize the molecular properties responsible for the compounds' discrimination, and are shown in Fig. 7.

MLP (Fig. 7A) and MEP (Fig. 7C) can be interpreted through a color range scheme, which varies depending on the software used for calculation. It is well-known that molecular surfaces can translate the shape of any molecular system, and the most active compound from series **II**, **4g**, is shorter and more bent than the most active compound from series **I**, **3g**. Furthermore, the total dipole moment vector was also displayed to the most active compounds (blue arrows in Fig. 7B).

The MLP property gives information of the molecular hydrophobic/hydrophilic balance, where the blue color corresponds to hydrophilic regions and the brown color to hydrophobic regions.[42] The lipophilicity can be numerically expressed by the calculated *n*-octanol/water partition coefficient (Clog*P*), applying Ghose et al.,—SLN, Sybyl Line Notation.[29,42] The brown color is mainly concentrated onto the substituent group region (**–**C_4H_9) for the both

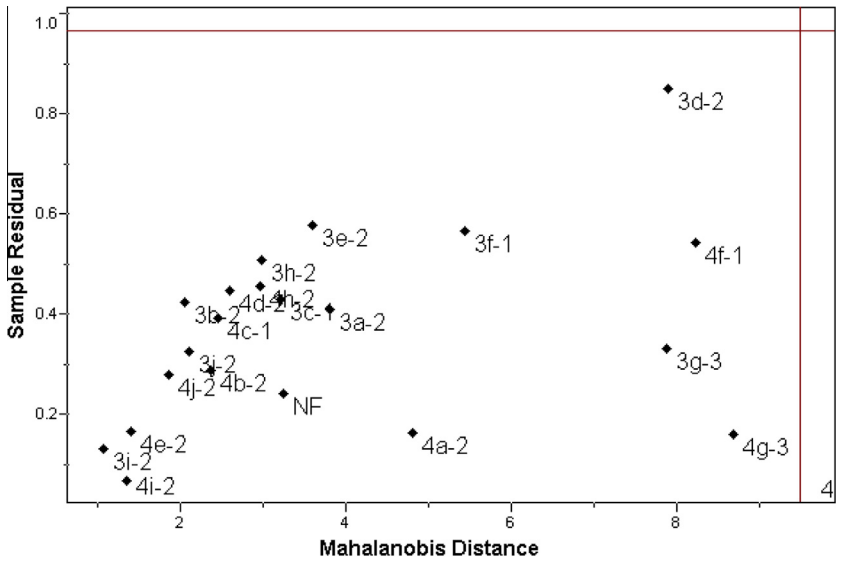


Figure 6. Plot of sample residual versus Mahalanobis distance. The sample residual threshold (red line) is based upon a 95% probability limit (set internally in Pirouette program).

compounds, **3g** and **4g** (see Fig. 7A). Regarding the MEP property, regions with higher electronic density distribution are displayed in the red color while regions with lower electronic density distribution are in blue.[43,44] The nearest neighbor moiety to the $-C_4H_9$ substituent group is less negatively charged (green/light blue color) in compound **4g** than in compound **3g** (yellow/green color). Thus, the findings indicate that the hydrophobic/hydrophilic balance and electronic density distribution seem to play an important role for the anti-*T. cruzi* activity.

3. Conclusion

The findings indicated that the anti-*T. cruzi* activity of the investigated compounds regarding the two series is strongly influenced by the physicochemical properties of the substituent group attached to the benzene ring.

The cytotoxicity assay shown that some compounds have presented interesting selectivity indices, highlighting especially compound **3g**, which had SI value between parasites and cell lines of 27.58 for J774 cells, and of 93.33 for FN1 cells.

The complementary findings generated in the HCA and PCA approach suggested primarily the hydrophobic and electronic properties as the responsible for the compounds grouping and, consequently, for the anti-*T. cruzi* activity of the 5-nitro-2-furfuriliden derivatives.

4. Experimental sections

4.1. Chemistry

NMR spectra were recorded on a Bruker ADPX Advanced (300 MHz) spectrometer employing DMSO-d_6 solutions with tetramethylsilane as internal standard. Melting points were determined using Büchi M-560 apparatus and elemental analysis was performed on a Perkin–Elmer 24013 CHN Elemental Analyzer.

4.1.1. General procedure for the synthesis of benzhydrazides (2a–j)

Each substituted benzoic acid (1) (0.02 mol) was refluxed for 4 h in 20.0 mL (0.49 mol) of anhydrous methanol and 0.5 mL (0.01 mol) of sulphuric acid. The reaction mixture was cooled down to room temperature and the hydrazine hydrate 80% (v/v) (10.0 mL, 0.13 mol) was added. The system was maintained by vigorously stirring for more 30 min in reflux. In the case of compound with 4-nitro substituent group attached to the benzene moiety, after the addition of hydrazine hydrate 80% (v/v) at room temperature, the reaction mixture was cooled down in ice bath and maintained into stirring during 1 h. After this period, the mixture was maintained at low temperature to give **2**.[19,22]

4.1.2. General procedure for the synthesis of azomethine derivatives (3a–j)

Compounds of series **I** were synthesized by refluxing 5-nitro-2-furaldehyde diacetate 98% (5 mmol) and benzhydrazides (3) (5 mmol) in water, sulphuric acid, acetic acid, and methanol (8:7:8:20 v/v) for 1 h. After cooling, the mixture was poured into cold water to precipitate the azomethine derivatives[20] (see structural elucidation of the compounds of series **I** in Supplementary data, p. S2).

4.1.3. General procedure for the synthesis of 1,3,4-oxadiazoline derivatives (4a–j)

Compounds of series **II** (1 mmol) were dissolved in acetic anhydride (10 mL, 0.10 mol), and the mixture was heated between 110 and 120 °C for 12 h under stirring and nitrogen gas atmosphere. After this period, 10 mL of distilled water was added to the system under constant stirring until its precipitate[6,21,43] (see structural elucidation of the compounds of series **II** in Supplementary data, p. S4). Identification of five new chemical entities, **4c** and **4g–4j**, which belong to series **II**, are described as follows (see numbered structure in Fig. 2 and ^{1}H and ^{13}C NMR spectra in Supplementary data, p. S6).

4.1.3.1. 3-Acetyl-5-(4-ciano-phenyl)-2-(5-nitrofuran-2-yl)-2,3-dihydro-1,3,4-oxadiazole (4c). Light yellow solid (55%); mp 177.0–178.0 °C. ^{1}H NMR (DMSO-d_6, 300 MHz): δ (ppm): 7.99 (d, 2H, H15, H19, J = 8.7 Hz), 7.96 (d, 2H, H16, H18, J = 8.7 Hz), 7.65 (d, 1H, H8, J = 3.8 Hz), 7.41 (s, 1H, H2), 7.20 (d, 1H, H7, J = 3.8 Hz), 2.32 (s,3H, H22); ^{13}C NMR {H} (DMSO-d_6, 75 MHz): δ (ppm): 167.4 (C20), 153.3 (C6), 151.8 (C5), 150.3 (C9), 132.9 (C16, C18),

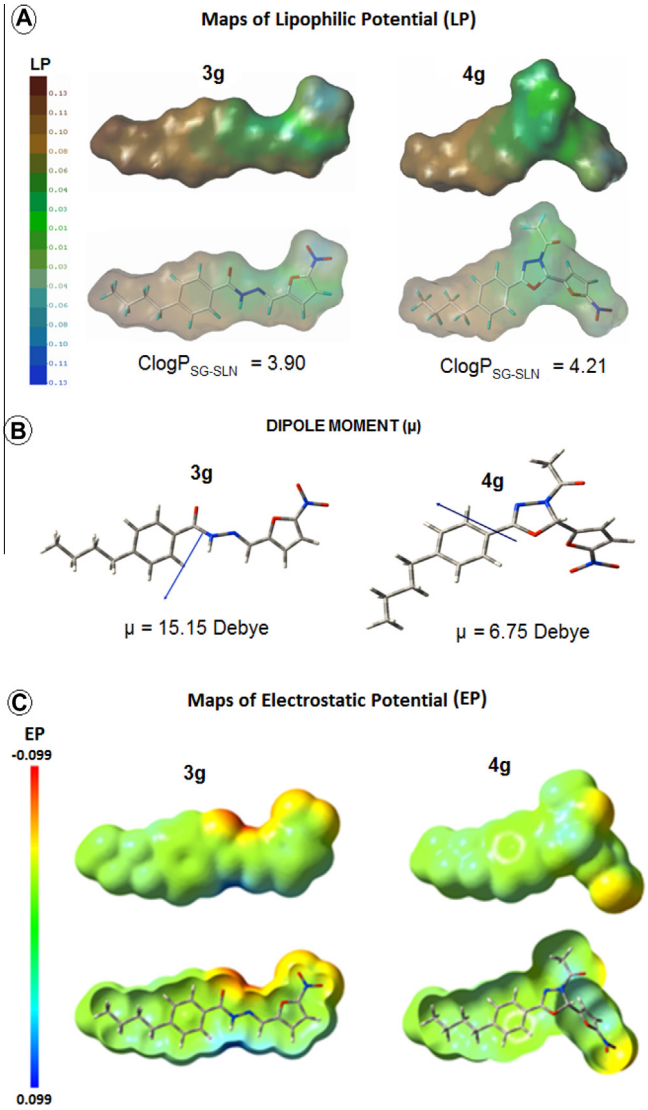


Figure 7. Visualization of lipophilic and electronic molecular properties for the most active 5-nitro-2-furfuriliden derivatives, **3g** (**–**C_4H_9) and **4g** (**–**C_4H_9): (A) the MLP property and Clog*P* values were calculated using the SYBYL 8.0 package,[42] and the color range scheme shows hydrophobic regions in brown (0.13) and hydrophilic regions in blue color (−0.13). The compounds are displayed in stick capped model (carbon atoms are in light gray color, oxygen in red, nitrogen in blue, and hydrogen in white). (B) Total dipole moment vector (μ) (blue arrows) was calculated using the HF/3-21G* method.[25,26] (C) The MEP property was calculated using GAUSSIAN 03W[43] and the visualization was generated employing GaussView 3.0.[44] The color range scheme indicates positive regions (0.099) in blue color and negative regions (−0.099) in red. The molecules are displayed in stick capped model (carbon atoms are in light gray color, oxygen in red, nitrogen in blue, and hydrogen in white).

127.7 (C14), 127.2 (C15, C19), 119.0 (CN), 117.8 (C17), 114.3 (C7), 112.9 (C8), 85.1 (C2), 20.9 (C22); Anal. Calcd for ($C_{15}H_{10}N_4O_5$): C, 55.22; H, 3.09; N, 17.17. Found: C, 54.88; H, 2.92; N, 17.18.

4.1.3.2. 3-Acetyl-5-(4-buthyl-phenyl)-2-(5-nitrofuran-2-yl)-2,3-dihydro-1,3,4-oxadiazole (4g). Light yellow solid (52%); mp 89.0–90.0 °C. ^{1}H NMR (DMSO-d_6, 300 MHz): δ (ppm): 7.76 (d, 2H, H15, H19, J = 8.2 Hz), 7.72 (d, 1H, H8, J = 3.8 Hz), 7.37 (s, 1H, H2), 7.36 (d, 2H, H16, H18, J = 8.0 Hz), 7.23 (d, 1H, H7, J = 3.8 Hz), 2.65 (t, 2H, CH_2), 2.30 (s, 3H, H22), 1.56 (t, 2H, CH_2), 1.31 (m, 2H, CH_2), 0.89 (s, 3H, CH_3). ^{13}C NMR {H} (DMSO-d_6, 75 MHz): δ (ppm): 167.2 (C20), 154.6 (C6), 151.7 (C5), 150.6 (C9), 147.0 (C17), 129.3 (C14), 128.9 (C15), 128,3 (C19), 126,6 (C16, C18), 114.6 (C7), 113.1 (C8), 84.2 (C2), 34,7 (CH_2), 32,6 (CH_2), 21,6 (CH_2), 21,0 (C22), 13,6 (CH_3); Anal. Calcd for ($C_{18}H_{19}N_3O_5$): C, 60.50; H, 5.36; N, 11.76. Found: C, 60.11; H, 5.47; N, 11.58.

4.1.3.3. 3-Acetyl-5-(4-trifluoromethyl-phenyl)-2-(5-nitrofuran-2-yl)-2,3-dihydro-1,3,4-oxadiazole (4h). Light yellow solid (86%); mp 166.0–167.0 °C. ^{1}H NMR (DMSO-d_6, 300 MHz): δ (ppm): 8.06 (d, 2H, H15, H19, J = 8.1 Hz), 7.92 (d, 2H, H16, H18, J = 8.4 Hz), 7.73 (d, 1H, H8, J = 3.8 Hz), 7.45 (s, 1H, H2), 7.28 (d, 1H, H7, J = 3.8 Hz), 2.33 (s,3H, H22); ^{13}C NMR {H} (DMSO-d_6, 75 MHz): δ (ppm): 172.6 (C20), 158.6 (C6), 157.0 (C5), 155.5 (C9), 137.1 (C14), 133.7 (C17), 132.6 (C15, C19), 131,3 (C16), 131,4 (C18), 127.0 (CF_3), 119.9 (C7), 118.3 (C8), 90.2 (C2), 26.3 (C22); Anal. Calcd for ($C_{15}H_{10}F_3N_3O_5$): C, 48.79; H, 2.73; N, 11.38. Found: C, 48.80; H, 2.66; N, 11.25.

4.1.3.4. 3-Acetyl-5-(4-bromo-phenyl)-2-(5-nitrofuran-2-yl)-2,3-dihydro-1,3,4-oxadiazole (4i). Light yellow solid (67%); mp 160.0–161.0 °C. ^{1}H NMR (DMSO-d_6, 300 MHz): δ (ppm): 7.76 (d, 2H, H15, H19), 7.76 (d, 2H, H16, H18), 7.70 (d, 1H, H8, J = 3.7 Hz), 7.38 (s, 1H, H2), 7.23 (d, 1H, H7, J = 3.7 Hz), 2.28 (s,3H, H22); ^{13}C NMR {H} (DMSO-d_6, 75 MHz): δ (ppm): 167.2 (C20), 153.8 (C6), 151.8 (C5), 150.3 (C9), 132.3 (C16, C18), 128.5 (C15, C19), 125,8 (C14), 122.6 (C17), 114.2 (C7), 112.9 (C8), 84.7 (C2), 21.1 (C22); Anal. Calcd for ($C_{14}H_{10}BrN_3O_5$): C, 44.23; H, 2.65; N, 11.05. Found: C, 44.13; H, 2.55; N, 10.73.

4.1.3.5. 3-Acetyl-5-(4-iodo-phenyl)-2-(5-nitrofuran-2-yl)-2,3-dihydro-1,3,4-oxadiazole (4j). Light yellow solid (78%); mp 175.0–176.0 °C. ^{1}H NMR (DMSO-d_6, 300 MHz): δ (ppm): 7.91 (d, 2H, H16, H18, J = 8.2 Hz), 7.70 (d, 1H, H8, J = 3.8 Hz), 7.58 (d, 2H, H15, H19, J = 8.2 Hz), 7.38 (s, 1H, H2), 7.23 (d, 1H, H7, J = 3.8 Hz), 2.29 (s,3H, H22); ^{13}C NMR {H} (DMSO-d_6, 75 MHz): δ (ppm): 167.2 (C20), 154.1 (C6), 151.8 (C5), 150.4 (C9), 138.0 (C16, C18), 128.2 (C15, C19), 122.8 (C14), 114.5 (C7), 113.1 (C8), 99.7 (C17), 84.6 (C2), 21.1 (C22); Anal. Calcd for ($C_{14}H_{10}IN_3O_5$): C, 39.37; H, 2.36; N, 9.84. Found: C, 39.47; H, 2.29; N, 9.68.

4.2. Anti-*T. cruzi* activity assay

In vitro anti-*T. cruzi* activity assays were performed against the epimastigote forms of Y strain of *T. cruzi* isolated by Silva and Nussenzweig.[45] The epimastigote form, at the stationary phase, and from a culture incubated for 10 days at 28 °C, was diluted into liver infusion tryptose medium (LIT) with 10% of fetal calf serum to obtain a final concentration of 2.0×10^6 parasites per mL. The nitro derivatives, the standard drugs, BZ and NFX, and the lead compound, NF, were separately dissolved in DMSO and diluted into LIT culture medium to obtain a concentration from 0.25 to 180.0 μM, depending on the biological activity of each compound. An aliquot of the inocula containing 2.0×10^6 viable parasites per mL was added to the compounds being evaluated and incubated at 28 °C for 5 days. After this period, the amount of viable parasites was determined using a spectrophotometer (T80 plus, PG Instruments, United Kingdom) at 580 nm (10 mm light path length). The assays were performed in triplicate and the anti-*T. cruzi* activity was determined by the relation: $PGI = \{1-[(A_p - A_{0p})/(A_c - A_{0c})]\} \times 100$, where: PGI = percentage of growth inhibition; $A_p = A_{580}$ of the culture containing the drug at day 5; $A_{0p} = A_{580}$ of the culture containing the compound after the addition of the inoculums at day 0; $A_c = A_{580}$ of the culture in the absence of any compound at day 5; $A_{0c} = A_{580}$ of the culture in the absence of any compound at day 0. The inhibitory concentration 50%, IC_{50}, was calculated from the percentage of inhibition growth using nonlinear model growth/sigmoid dose–response for each case, employing at least, five concentrations.[6,23]

4.3. Cytotoxicity assays

4.3.1. Cell culture

Cell culture of murine macrophage J774 cell line and FN1 human fibroblast cells were maintained in RPMI 1640 containing 10% FBS, 100 UI penicillin and 100 μg/ml gentamycin at 37 °C in 95% air and 5% CO_2.

4.3.2. J-774 murine macrophage-like cells

It were seeded (1.0×10^4 cells/well) in 96 well microplates with 200 mL of RPMI 1640 medium supplemented with 5% heat-inactivated fetal calf serum. Cells were allowed to attach for 24 h in a humidified 5% CO_2/95% air atmosphere at 37 °C. Then, cells were exposed to compounds **3c**, **3f**, **3g**, **4c**, **4f**, **4g**, the standard drugs, BZ and NFX, and the lead compound, NF, in concentrations between 0.8 and 400.0 μM for 48 h. Afterwards, cell viability was assessed by measuring the mitochondrial-dependent reduction of MTT (Sigma) to formazan. For that purpose, MTT was added to cells to a final concentration 0.4 mg/mL and cells were incubated at 37 °C for 3 h. After removing the media, formazan crystals were dissolved in DMSO (180 mL), and the absorbance at 595 nm was read using a microplate reader (Titertek Uniscience, Multiskan MCC/340, USA). Results are expressed as IC_{50} (compound concentration that reduce 50% control absorbance at 595 nm). Every IC_{50} is the average of six different experiments and was calculated from the percentage of inhibition growth using nonlinear model growth/sigmoid dose–response for each case using, at least, five concentration values. Cytotoxicity percentages (%*C*) were determined as following: $\%C = 100-[(OD_d - OD_{dm})/(OD_c - OD_{cm}) \times 100]$, where OD_d is the mean of OD_{595} of wells with macrophages and different concentrations of the compounds; OD_{dm} is the mean of OD_{595} of wells with different compounds concentration in the medium; OD_c is the growth control and OD_{cm} is the mean of OD_{595} of wells with medium only.[14,46,47]

4.3.3. FN1 human fibroblast cells

It were plated in triplicate in 96-well plates at a density of 1×10^4 cells/well into flat microtiter plates and incubated for overnight at 37 °C in a humidified incubator containing 5% CO_2. Then, the cells were treated with the compounds from 12.5 to 200 μM in a final volume of 100 μL. After 24 h of treatment, cells were exposed to 5 mg/mL of MTT for 2 h, and the precipitated formazan was dissolved in 0.1 N HCl in isopropanol and measured at 540 nm with a microplate reader Thermo Plate (Thermoplate TP Reader, Tokay Hit, Japan) and the IC_{50} values determined.[27]

4.4. Molecular modeling approach, calculation and selection of descriptors

The 3D molecular models of the compounds from series **I** were built up using the Cartesian coordinates of crystallized structure of the lead compound NF (LEQTAC code, $R_{factor} = 0.11$)[48] retrieved from the Cambridge Structural Database (CSD) as reference geometry.[49] For compounds from series **II**, the 3D molecular models were constructed from the crystallized structure of NF and a similar fragment, 2-[4-acetyl-5-(biphenyl-4-yl)-4,5-dihydro-1 3,4-oxadiazol-2-yl]-phenyl acetate, containing the oxadiazolinic ring ($R_{factor} = 0.035$),[50] giving the 3D molecular model of compound **4a**.[49] For the reference drugs, the Cartesian coordinates of crystallized structure of BZ ($R_{factor} = 0.025$)[51] and NFX analogue ($R_{factor} = 0.042$)[52] were used as template.[49]

Each molecular model was energy-minimized in the MM+ force field,[53] without any constrains. The partial atomic charges were calculated employing the AM1 (Austin Model 1) semiempirical method.[54]

The energy minimization of the 3D molecular models (**4a** and NF) was also performed employing the MOLSIM 3.2 program[55] by applying the steepest descent and, subsequently, the conjugated gradient method, using a convergence criterion of 0.01 kcal/mol. The energy-minimized structures were the input to the MD simulations of 1 ns (1,000,000 steps; step size of 1 fs) at 301 K (28 °C), the same temperature of the biological assay. Dielectric constant of 3.5 was used to simulate the environment of the biological membranes and enzyme or receptor binding sites. It was assigned a fictitious atomic mass of 5000 u.m.a. to some atoms position in order to maintain the structural integrity of the molecular models during simulation. Trajectory files were recorded every 20 simulation steps to generate 50,000 conformations. The lowest-energy conformation was selected from the equilibrium region of conformational ensemble profile (CEP) and compared to the initial energy-minimized molecular model to verify whether the structural integrity was maintained after simulation. The RMSD value was used as criterion to verify the maintenance of structural integrity employing the Hyperchem 8.0 program.[49]

The selected lowest-energy conformation for NF and compound **4a** were used as template for building up the molecular models of compounds from series **I** and **II**, respectively.[49] The geometry of each model was optimized employing the MM+ force field,[53] and the partial atomic charges were calculated using the AM1 semiempirical method.[54] The MOLSIM 3.2 program[55] was also employed to energy-minimization procedure, applying the steepest descent and conjugated gradient methods. The convergence criterion was 0.01 kcal/mol. At this time, some thermodynamic properties were obtained as follows: the total potential energy from minimization step (E_{total}); the intramolecular energy contribution of solvation (E_{solv}), applying the hydration shell model proposed by Forsythe and Hopfinger;[56] and, the intramolecular energy contribution of hydrogen bonding (E_{Hb}). The resulting 3D molecular model for each compound was the input structure to the calculation of physicochemical and structural properties, which were explored in this study.

The electronic properties as Charges from Electrostatic Potentials using a Grid based method (CHELPG),[26] maps of electrostatic potential (MEP), dipole moment (total and x, y, z), and frontier molecular orbital energy (E_{HOMO} and E_{LUMO}) were calculated using the HF/3-21G* method,[25,26] implemented in the Gaussian 03W program.[43]

The maps of lipophilic potential (MPL) and the respective values of the calculated *n*-octanol/water partition coefficient ($C\log P$) were obtained using the Sybyl 8.0 package.[42] The maps were calculated onto Connolly molecular surfaces, using a probe with 1.4 Å radii (water molecule). Two methods available in the Sybyl 8.0 program were employed to compute $C\log P$: Ghose et al., 1998-SLN ($C\log P_{SG\text{-}SLN}$)[42] and Viswanadhan et al. ($C\log P_{SV}$).[30]

The Marvin 5.2.1_1 program[57] was used to calculate steric, topological/geometric, and lipophilic properties. $C\log P_{WM}$ was calculated by the weighted method, assigning equal weight for each method.[30–32]

Properties related to the substituent groups were retrieved from the literature.[33] Further detailed information regarding all descriptors, methods, and respective software used to perform these calculations are listed in Supplementary data, p. S21, Table 3S.

The independent variables or descriptors or molecular properties of different nature were calculated and generated a matrix containing 41 columns (independent variables or descriptors plus biological activity (log1/C)) and 21 rows corresponding to the number of compounds (11 from series **I**, including NF, and 10 from series **II**). The compounds were distributed in gradients regarding the biological activity (less, moderately and highly active compounds).

Procedure to select the most representative molecular properties: two filters were used as criteria to previously select the independent variables: (i) the Pearson correlation coefficient and (ii) a visual inspection of scatter plots for biological activity versus each variable or descriptor or molecular property.[17] These data are presented as Supplementary data (p. S23, Table 4S). The Pearson correlation coefficient value of 0.22 was formerly used as cut off for selecting the calculated independent variables. Regarding the visual inspection, only variables presenting uniform distribution and linear tendency with the biological data were selected to compose the final matrix.

The final matrix was used as input for the exploratory data analysis. Due to the distinct magnitude orders among the calculated variables, the autoscaling procedure was applied as a preprocessing method.[35]

4.5. Exploratory data analysis[17,18]

4.5.1. Hierarchical cluster analysis, HCA

HCA was performed using the Pirouette 3.11 program,[40] employing the complete linkage method and Euclidean distance. In HCA, distances between pairs of samples (or variables) are calculated and compared. When distances between samples are relatively small, this implies that the samples are similar. The calculated distances between samples were set on a similarity matrix whose elements are called similarity indices, ranging between 0 and 1, where 1 is equivalent to a maximum similarity. The results are usually presented as a two-dimensional chart, called dendrogram.[40]

The multivariate distance d_{ab} between two samples vectors, a and b, was determined by computing differences at each of the m variables:

$$d_{ab} = \left[\sum_{i}^{m}(X_{aj} - X_{bj})^{M}\right]^{1/M}$$

M is the order of the distance, and here corresponds to the Euclidean distance ($M = 2$).[40] Because of inter-sample distances can vary with the type and number of measurements, it is customary to transform them onto a somewhat more standard scale of similarity, where d_{max} is the largest distance in the data set:

$$\text{similarity}_{ab} = 1 - \frac{d_{ab}}{d_{max}}$$

HCA was carried out for samples and variables. In the first case, biological activity was considered as dependent variable. In the second option, the biological activity is considered as independent variable.[17,18,40]

4.5.2. Principal components analysis, PCA

The exploratory analysis of PCA was also carried out employing the Pirouette 3.11 software.[40] For PCA, the biological activity was considered as the dependent variable. The data were decomposed into two matrices, one of scores related to the samples, and another of loadings, related to the variables.[17,18]

The new set of axes generates the PCs, or factors, into which are the information related to the original descriptors. Thus, the number of PCs that explain most of the variability in the data set can be determined, considering that these PCs are uncorrelated and mutually orthogonal variables built up as simple linear combinations from the original data.[17,18]

In this exploratory data analysis, PCA was run up to ten factors or PCs. The outliers' diagnosis, implemented in Pirouette 3.11 software,[40] was also performed through the Mahalanobis distance.[41]

Acknowledgements

The authors thank Professor Maria Júlia Manso Alves of the Laboratory of Biochemical Parasitology, Department of Biochemistry, Institute of Chemistry, University of São Paulo (USP) for providing the epimastigote form of *T. cruzi* strain. Prof. Elizabeth Igne Ferreira of the LAPEN laboratory, Faculty of Pharmacy, University of São Paulo (USP), which had kindly provided access to MOLSIM 3.2 program. The authors thank Brazilian scientific committees, FAPESP, CAPES and CNPq for the financial support.

Supplementary data

Supplementary data (structure elucidation of 5-nitro-2-furfuriliden derivatives (series **I**); structure elucidation of 3-acetyl-2,5-disubstituted-2,3-dihydro-1,3,4-oxadiazole derivatives (series **II**); ^{1}H and ^{13}C NMR spectra of new compounds synthesized (**4c** and **4g**–**j**); dose–response curve of anti-*T. cruzi* activity assay; conformational ensemble profile of nifuroxazide **NF**; conformational ensemble profile of compound **4a** (series **II**); values of the thermodynamic parameters obtained for the lowest-energy conformer of nifuroxazide (NF) and of **4a** compound, selected in CEP from MD simulations; Independent variables obtained from literature and from calculation for nifuroxazide (**NF**), azomethine (**3a**–**g**) and oxadiazoline derivatives (**4a**–**g**); descriptors, methods and software used to perform calculation for the two series of 5-nitro-2-furfuriliden derivatives; Pearson's correlation coefficient and scatter plots of the descriptors versus biological activity) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2013.06.017.

References and notes

1. Schmunis, G. *Lancet Infect. Dis.* **13, 2013, 283**.
2. Rassi, A.; Marin-Neto, J. A. *Lancet* **2010**, *375*, 1388.
3. Astelbauer, F.; Walochnik, J. *Int. J. Antimicrob. Agents* **2011**, *38*, 118.
4. Coura, J. R. *Mem. Inst. Oswaldo Cruz* **2009**, *104*, 549.
5. Cerecetto, H.; González, M. *Pharmaceuticals* **2010**, *3*, 810.
6. Ishii, M.; Jorge, S. D.; Oliveira, A. A.; Palace-Berl, F.; Sonehara, I. Y.; Pasqualoto, K. F. M.; Tavares, L. C. *Bioorg. Med. Chem.* **2011**, *19*, 6292.
7. Jorge, S. D.; Ishii, M.; Palace-Berl, F.; Ferreira, A. K.; De-Sá-Júnior, P. L.; Oliveira, A. A.; Sonehara, I.; Pasqualoto, K.; Tavares, L. *Med. Chem. Commun.* **2012**, *3*, 824.
8. Paula, F. R.; Serrano, S. H. P.; Tavares, L. C. *Quím. Nova* **2009**, *32*, 1013.
9. Viodé, C.; Bettache, N.; Cenas, N.; Krauth-Siegel, R. L.; Chauvière, G.; Bakalara, N.; Périé, J. *Biochem. Pharmacol.* **1999**, *57*, 549.
10. Craig, P. N. *J. Med. Chem.* **1971**, *14*, 680.
11. Rollas, S.; Gulerman, N.; Erdeniz, H. *IL Farmaco* **2002**, *57*, 171.
12. Lee, L.; Robb, L. M.; Lee, M.; Davis, R.; Mackay, H.; Chavda, S.; Babu, B.; O'Brien, E. L.; Risinger, A. L.; Mooberry, S. L. *J. Med. Chem.* **2009**, *53*, 325.
13. Maccioni, E.; Alcaro, S.; Cirilli, R.; Vigo, S.; Cardia, M. C.; Sanna, M. L.; Meleddu, R.; Yanez, M.; Costa, G.; Casu, L.; Matyus, P.; Distinto, S. *J. Med. Chem.* **2011**, *54*, 6394.
14. Romeiro, N. C.; Aguirre, G.; Hernández, P.; González, M.; Cerecetto, H.; Aldana, I.; Pérez-Silanes, S.; Monge, A.; Barreiro, E. J.; Lima, L. M. *Bioorg. Med. Chem.* **2009**, *17*, 641.
15. Ralph, P.; Nakoinz, I. *Nature* **1975**, *257*, 393.
16. Ferreira, A. K.; Meneguelo, R.; Pereira, A.; Mendonça, F. O. R.; Chierice, G. O.; Maria, D. A. *Anticancer Res.* **2012**, *32*, 95.
17. Ferreira, M. *J. Braz. Chem. Soc.* **2002**, *13*, 742.
18. Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. *Chemometrics: A Practical Guide*; Wiley-Interscience: New York, 1998.
19. Jorge, S. D.; Palace-Berl, F.; Masunari, A.; Cechinel, C. A.; Ishii, M.; Pasqualoto, K. F. M.; Tavares, L. C. *Bioorg. Med. Chem.* **2011**, *19*, 5031.
20. Tavares, L. C.; Penna, T. C.; Amaral, A. T. *Boll. Chim. Farm.* **1997**, *136*, 244.
21. Somogyi, L. *J. Heterocycl. Chem.* **2007**, *44*, 1235.
22. Li, C.-K.; Ma, Y.-J.; Cao, L.-H. *J. Chin. Chem. Soc.* **2009**, *56*, 182.
23. Cerecetto, H.; Di Maio, R.; González, M.; Risso, M.; Saenz, P.; Seoane, G.; Denicola, A.; Peluffo, G.; Quijano, C.; Olea-Azar, C. *J. Med. Chem.* **1999**, *42*, 1941.
24. Baber, J. C.; Thompson, D. C.; Cross, J. B.; Humblet, C. *J. Chem. Inf. Model.* **2009**, *49*, 1889.
25. Fock, V. *Z. Phys. A: Hadron Nucl.* **1932**, *75*, 622.
26. Brenneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **2004**, *11*, 361.
27. Ferrara, P.; Apostolakis, J.; Caflisch, A. *Proteins* **2002**, *46*, 24.
28. Verloop, A. *The STERIMOL Approach to Drug Design*; Marcel Dekker: New York, 1987.

29. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Phys. Chem.* **1998**, *102*, 3762.
30. Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163.
31. Klopman, G.; Li, J.-Y.; Wang, S.; Dimayuga, M. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 752.
32. *PHYSPROP database*. http://www.syrres.com/what-we-do/databaseforms.aspx?id=386.
33. Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*; American Chemical Society: Washington, 1995.
34. Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* **1964**, *86*, 1616.
35. Ferreira, M.; Antunes, A. M.; Melgo, M. S.; Volpe, P. L. O. *Quim. Nova* **1999**, *22*, 724.
36. Arroio, A.; Honório, K. M.; da Silva, A. B. F. *Quim. Nova* **2010**, *33*, 694.
37. Soriano-Correa, C.; Raya, A.; Esquivel, R. O. *Int. J. Quantum Chem.* **2008**, *108*, 1369.
38. Jorge, S. D.; Palace-Berl, F.; Pasqualoto, K. F. M.; Ishii, M.; Ferreira, A. K.; Berra, C. M.; Bosch, R. V.; Maria, D. A.; Tavares, L. C. *Eur. J. Med. Chem.* **2013**, *64*, 200.
39. Boiani, M.; Cerecetto, H.; Gonzalez, M.; Gasteiger, J. *J. Chem. Inf. Model.* **2008**, *48*, 213.
40. *PIROUETTE 3.11*, Infometrix Inc.: Woodinville, WA, 1990–2003.
41. Mahalanobis, P. C. *J. Assist. Soc. Bengal* **1930**, *26*, 541.
42. *SYBYL 8.0 for Linux*; Tripos, Inc.: 2007. 1699 South Hanley Rd., St. Louis, MO 63144–2917, USA.
43. GAUSSIAN 03W-revision B.02 for Windows, version 6; Gaussian Inc.: Pittsburgh, PA, 1995–2003.
44. *GAUSSVIEW 3.0*; Gaussian Inc.: Pittsburgh, PA, 2000–2003.
45. Silva, L. H. P.; Nussenzweig, V. *Folia Clín. Biol.* **1953**, *20*, 191.
46. Mosmann, T. *J. Immunol. Methods* **1983**, *65*, 55.
47. Denizot, F.; Lang, R. *J. Immunol. Methods* **1986**, *89*, 271.
48. Pniewska, B.; Januchowski, M. *Pol. J. Chem.* **1998**, *72*, 2629.
49. *HYPERCHEM Program for Windows*; Hypercube, Inc.: Gainesville, FL, 2008.
50. Yehye, W. A.; Ariffin, A.; Rahman, N. A.; Ng, S. W. *Acta Crystallogr. Sect. E: Struct. Rep. Online* **2010**, *66*, o878.
51. Soares-Sobrinho, J. L.; Cunha-Filho, M. S.; Rolim Neto, P.; Torres-Labandeira, J. J.; Dacunha-Marinho, B. *Acta Crystallogr. Sect. E: Struct. Rep. Online* **2008**, *64 (Pt. 3)*, o634.
52. Caracelli, I.; Stamato, F.; Mester, B.; Paulino, M.; Cerecetto, H. *Acta Crystallogr. C* **1996**, *52*, 1281.
53. Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.
54. Dewar, M. J.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
55. Doherty, D. *MOLSIM: Molecular Mechanics and Dynamics Simulation Software-User's Guide, version 3.2*; The Chem21 Group Inc.: Lake Forest, 1997.
56. Forsythe, K.; Hopfinger, A. *Macromolecules* **1973**, *6*, 423.
57. *Marvin Beans, version 5.2.1_1; ChemAxon Ldt.*: 1998–2010, Budapest, Hungary.