Accepted Manuscript

Model for High-Throughput Screening of drug immunotoxicity - study of the anti-microbial G1 over peritoneal macrophages using flow cytometry

Esvieta Tenorio-Borroto, Claudia G. Peñuelas-Rivas, Juan C. Vásquez-Chagoyán, Nilo Castañedo, Francisco J. Prado-Prado, Xerardo García-Mera, Humberto González-Díaz

PII: S0223-5234(13)00552-7

DOI: 10.1016/j.ejmech.2013.08.035

Reference: EJMECH 6382

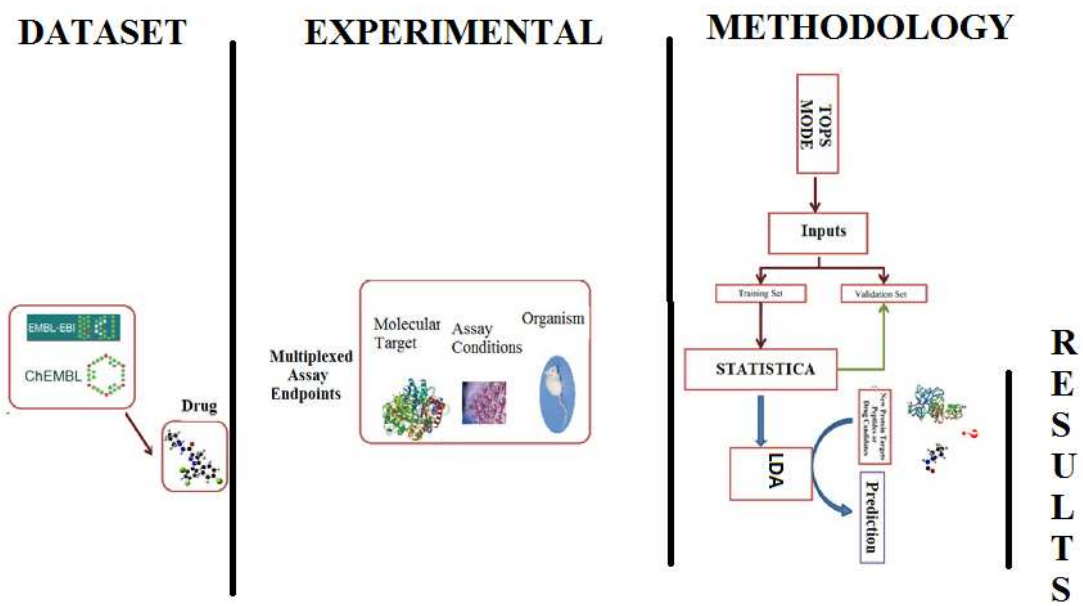
To appear in: *European Journal of Medicinal Chemistry*

Received Date: 30 August 2012

Revised Date: 29 August 2013

Accepted Date: 31 August 2013

Please cite this article as: E. Tenorio-Borroto, C.G. Peñuelas-Rivas, J.C. Vásquez-Chagoyán , N. Castañedo, F.J. Prado-Prado, X. García-Mera, H. González-Díaz, Model for High-Throughput Screening of drug immunotoxicity - study of the anti-microbial G1 over peritoneal macrophages using flow cytometry, *European Journal of Medicinal Chemistry* (2013), doi: 10.1016/j.ejmech.2013.08.035.



Highlights:

1. We train and validate by the first time one High-throughput mt-QSAR model using TOPS-MODE.
2. The model correctly classifies 8,258 out of 9,000 multiplexing assay endpoints of 7903 drugs..
3. We determined the values of EC_{50} and Cytotoxicity for the anti-microbial / anti-parasite drug G1.

Model for High-Throughput Screening of drug immunotoxicity - study of the anti-microbial G1 over peritoneal macrophages using flow cytometry

Esvieta Tenorio-Borroto [a,b], Claudia G. Peñuelas-Rivas [b], Juan C. Vásquez-Chagoyán [b], Nilo Castañedo [c], Francisco J. Prado-Prado [d], Xerardo García-Mera [a], Humberto González-Díaz [e,f,*]

[a] *Department of Organic Chemistry, Faculty of Pharmacy, USC, 15782, Santiago de Compostela, Spain.*
[b] *Centro de Investigación y Estudios Avanzados en Salud Animal, UAEM, México, 50200.*
[c] *Chemical Bioactive Center (CBQ), Central University of Las Villas (UCLV), 50200, Santa Clara, Cuba.*
[d] *Biomedical Sciences Department, Health Science Division, University of Quintana Roo, 77039, Chetumal, Mexico.*
[e] *Department of Organic Chemistry II, University of the Basque Country UPV/EHU, 48940, Leioa, Spain*
[f] *IKERBASQUE, Basque Foundation for Science, 48011, Bilbao, Spain*

ARTICLE INFO

Article history:

Keywords:
High-throughput model
Drug immunotoxicity
Multiplex assay endpoints
Flow cytometry
Macrophage
QSAR model
ChEMBL

ABSTRACT

Quantitative Structure-Activity (mt-QSAR) techniques may become an important tool for prediction of cytotoxicity and High-throughput Screening (HTS) of drugs to rationalize drug discovery process. In this work, we train and validate by the first time mt-QSAR model using TOPS-MODE approach to calculate drug molecular descriptors and Linear Discriminant Analysis (LDA) function. This model correctly classifies 8,258 out of 9,000 (Accuracy = 91.76%) multiplexing assay endpoints of 7903 drugs (including both train and validation series). Each endpoint correspond to one out of 1418 assays, 36 molecular and cellular targets, 46 standard type measures, in two possible organisms (human and mouse). After that, we determined experimentally, by the first time, the values of EC_{50} = 21.58 μg/mL and Cytotoxicity = 23.6 % for the anti-microbial / anti-parasite drug G1 over Balb/C mouse peritoneal macrophages using flow cytometry. In addition, the model predicts for G1 only 7 positive endpoints out 1,251 cytotoxicity assays (0.56% of probability of cytotoxicity in multiple assays). The results obtained complement the toxicological studies of this important drug. This work adds a new tool to the existing pool of few methods useful for multi-target HTS of ChEMBL and other libraries of compounds towards drug discovery.

1. Introduction

The drug cytotoxicity tests are screening methods that typically uses permanent cell lines for ranking acute toxicities of parent compounds based on the basal cytotoxicity theory chemicals exert their acute toxic effects by interfering with basic cellular functions that are common to all mammalian cells [1]. In vitro drug cytotoxicity may be variable among different cell lines; thus one parameter for cell death is the integrity of the cell membrane, which can be measured by the cytoplasmic enzyme activity released by damaged cells

Abbreviations: QSAR/QSTR, Quantitative-Structure/Toxicity Relationship; LDA, Linear Discriminant Analysis; ChEMBL.

* Corresponding author at: *Department of Organic Chemistry II, University of the Basque Country UPV/EHU, 48940, Leioa, Spain.*
E-mail address: humberto.gonzalez@usc.es (H. González-Díaz).

[2]. Specifically, macrophages are phagocytic cells that recognize and kill microbial and tumor targets by cell-to-cell contact or through secretion of a wide array of products including reactive oxygen species, reactive nitrogen intermediates, cytokines, chemokines, *etc*. [3]. The macrophages execute numerous functions such as antigen presentation, cytokine production, phagocytosis, migration, and the production of ROS [4]. Important drug cytotoxicity assays measure the increase in the number of macrophages due to injection of one sterile irritant agent, such as thioglycollate, several days prior to harvesting the cells. The resulting peritoneal cells are referred to as elicited macrophages. A commonly used source of mouse and rat macrophages is the peritoneal cavity. Two types of macrophages from the peritoneal cavity are used, resident and elicited [5]. The experimental results obtained by many groups worldwide that carry out assays of drug cytotoxicity / biological effects over different cell lines including macrophage cells are available for public research in the database ChEMBL at: https://www.ebi.ac.uk/ChEMBLdb [6]. ChEMBL contains >10,000 outcomes for assays of drugs related somehow to macrophage with different degrees of curation (outputs obtained after using macrophage as keyword in a simple search). In this context, the search of computational models to predict the possible results for new drugs in all these assays have become a goal of the major importance to reduce experimentation costs. Besides, many drugs have been assayed only for some selected tests out of the many possible assays described so data mining of ChEMBL is a very interesting source of new knowledge [7].

In special, Quantitative Structure-Activity Relationships (QSAR) and specifically Quantitative Structure-Toxicity Relationships (QSTR) have been widely used to predict toxicity from chemical structure and corresponding physicochemical properties [8]. Unfortunately, almost current QSAR/QSTR models are able to predict new outcomes only for one specific assay. In our opinion, we can circumvent this problem using multi-target QSAR (mt-QSAR) techniques to model complex datasets determined in multiplexing assay conditions (m_j) as is the case of ChEMBL [9, 10]. There are many types of molecular descriptors that can be used, in principle, to seek mt-QSAR or mt-QSTR models able to predict drug activity and cytotoxicity. For instance, entropy measures are very flexible parameters that can be used in many situations in QSAR/QSTR modeling. In a very recent work [11], we used entropy measures to train and validate for the first time a QSTR model that correctly classifies 8,806 out of 9,001 (Accuracy = 91.1%) multiplexing assay endpoints of 7903 drugs (including both training and validation series). We have fit the classifier using LDA. The best equation found was:

$$S_i(m_j) = -2.6103^{*}\theta_5^i - 0.9343\Delta\theta_5^i(o) - 0.2561\cdot\Delta\theta_5^i(t) + 0.0051\cdot\Delta\theta_5^i(s) - 2.6103 \quad (1)$$

$$N = 6611 \qquad R_c = 0.74 \qquad \chi^2 = 5438.21 \qquad p < 0.05$$

This model is expected to give different classification probabilities of the compound for different conditions m_j: organisms (o_t), biological assays (a_u), molecular or cellular targets (t_e), or standard type of activity measure (s_x). Consequently, $S(m_j) = S(d_i, a_u, c_l, o_t, t_e, s_x)$ is a real-valued variable used here to score the propensity of the drug to be active in multiplex pharmacological assays in selected m_j. The different parameters in the equation were introduced to codify specific information that is known to be determinant in the final value of biological activity. The statistical parameters for the above equation are: Number of cases (N), Canonical Regression coefficient (R_c), Chi-square statistic (χ^2), and error level (p-level) < 0.05 [12]. Each endpoint corresponds to one out of 1443 assays, 32 molecular and cellular targets, 46 standard type measures, in two possible organisms (human and mouse). In **Table 1**, we resumed all the statistical parameters of this model, other models published before, and a new model we are going to present in this work.

Table 1 comes about here

We have also determined experimentally, for the first time, the values of EC_{50} = 8.21 μg/mL and Cytotoxicity = 26.25 % for the antimicrobial / antiparasitic drug G1 on Balb/C mouse thymic macrophages using flow cytometry. In addition, we have used the new model to predict G1 endpoints in 1,283 assays finding a low average probability of p(1) = 0.50% in 152 cytotoxicity assays. Last, we have used the model to predict average probability of the interaction of G1 with different proteins in macrophages. Interestingly, the Macrophage colony-stimulating factor receptor, the Macrophage colony-stimulating factor 1 receptor, the Macrophage migration inhibitory factor, Macrophage scavenger receptor types I and II, and the Macrophage-stimulating protein receptor, have also very low average predicted probabilities of p(1) = 0.092, 0.038, 0.077, 0.026, 0.2, 0.106, respectively. Both experimental and theoretical results show a moderate cytotoxicity of G1 over thymic macrophages.

In other work [13], we have developed one mt-QSTR model based on the spectral moments of on stochastic matrix. The model with only four variables is able to assign each drug to 1 out of 2 possible activity classes: active (C = 1) or non-active compounds (C = 0), given the molecular structure and several multiplex assay conditions m_j. A general data set consisting of >10,000 multiplexing assay endpoints has been downloaded from the public ChEMBL database [6, 14]. In any case, after a careful curation of the dataset we have retained 9,001 multiplexing assay endpoints (statistical cases) after the elimination of all cases with missing information or very low representation. In order to find the mt-QSTR model, we have used the LDA module of the software package STATISTICA 6.0 [15]. This classifier presented good results in both training and external validation series with an overall Accuracy higher than 90%. According to previous reports in the QSTR literature, the Accuracy values higher than 75% are acceptable [16-22]. The reader should be aware that N here is not the number of compounds but the number of statistical cases. See also **Table 1** for comparative purposes. The best equation found was:

$$S_i(m_j) = 0.7488^{*}\pi_5^i - 0.2668\cdot\Delta\pi_5^i(o) - 0.0127\cdot\Delta\pi_5^i(s) + 0.0569\cdot\Delta\pi_5^i(t) - 2.21514 \quad (2)$$

$$N = 6611 \quad R_c = 0.74 \quad \chi^2 = 543821 \quad p < 0.05$$

Another important method for mt-QSAR studies is TOSS-MODE; which was introduced by Estrada *et al.* [23-27] and implemented in the software MODESLAB (renamed as TOPS-MODE). TOPS-MODE have been demonstrated to be successful in both QSAR [28-32] and QSTR models [33-36]. More recently TOPS-MODE have been applied to High-throughput mt-QSAR studies by our group [37] and also Molina & Speck-Planche *et al.* [38]. In a previous paper [39], we have trained and tested an Artificial Neural Network (ANN) model for the first time, in order to perform a multiplexing prediction of drugs effect on macrophage populations. In so doing, we have used the TOPS-MODE approach to calculate drug molecular descriptors and the software STATISTICA to seek different ANN models such as Linear Neural Network (LNN) and Multi-Layer Perceptrons (MLP). The LNN was best model found with Ac = 93.0% (8,258 out of 9,000 assay endpoints) for7903 drugs in training and test series, see also the **Table 1**.

$$S_i(m_j) = -2.57^{*}\mu_5^i + 0.775\Delta\mu_5^i(o) - 2.31\cdot\Delta\mu_5^i(t) + 6.30\cdot\Delta\mu_5^i(s) + 0.4300523 \quad (3)$$

$$N = 6747 \quad S_p = 92.2 \quad S_n = 93.1$$

Each endpoint corresponds to one out of 1418 assays, 36 molecular or cellular targets, 46 standard type measures, in two possible organisms (human and mouse). Secondly, we have determined experimentally, for the first time, the values of EC_{50} = 11.41 μg/mL and Cytotoxicity = 27.1 % for the drug G1 over Balb/C mouse spleen macrophages using flow cytometry. In addition, we have used the LNN model to predict the G1 activity in 1,265 multiplexing assays not measured experimentally (including 152 cytotoxicity assay endpoints). Both experimental and theoretical results point out a low macrophage cytotoxicity of G1 over spleen macrophages.

In all these previous works, we have used different machine learning (LDA, LNN, and MLP) to seek mt-QSARs for CHEMBL cytotoxicity dataset. We determined also the effect of G1 over spleen and thymic macrophages. However, the LDA model using TOPS-MODE and the cytotoxicity of G1 over peritoneal macrophages have not been reported yet. In this work we reported the first mt-QSAR model for drug cytotoxicity using CHEMBL data set, TOPS-MODE descriptors, and LDA technique to seek the model. Next, we report, by the first time, the experimental study of the effect of the drug G1 over Balb/C mouse peritoneal macrophage population using flow cytometry. Last, we carry out the prediction of other multiplexing assay endpoints for G1, not experimentally determined in this work. We also present by the first time a comparative table with the results obtained with different models developed before and in this work. The results obtained are very important because they complement the toxicological studies of this important anti-bacterial, anti-fungal, and anti-parasite drug.

2. Materials and Methods

2.1. Computational methods

In order to seek the High-throughput mt-QSAR model we used the LDA module of the software package STASTICA 6.0 [15]. The model developed presented the general form.

$$S_i(m_j) = b_0 + b_1 \cdot p(a_u) \cdot p(c_l) \cdot {}^{std}\mu_5^i + \sum_{j=2}^{4} b_j \cdot \Delta\mu_5^i(m_j) \quad (4)$$

$$= b_0 + b_1 \cdot p(a_u) \cdot p(c_l) \cdot {}^{std}\mu_5^i + \sum_{j=2}^{4} b_j \cdot \left({}^{std}\mu_5^i - \left\langle {}^{std}\mu_5^i(m_j) \right\rangle\right)$$

Where, $S(m_j) = S(d_i, a_u, c_l, o_t, t_e, s_x)$ is a real-valued variable that scores the propensity of the drug to be active in multiplex pharmacological assays of the drug depending on the conditions selected mj. The statistical parameters used to corroborate the model were: Number of cases (N), Canonical Regression coefficient (R_c), Chi-square statistic (χ^2), and error level (p-level); which have to be < 0.05 [12].

2.2. Biology assays

*2.2.1. Peritoneal macrophages isolation and cell culture***.**

Peritoneal macrophages were obtained from mice euthanized by cervical dislocation. The peritoneal of the animals were surgically exposed using a midline incision. Peritoneal fluid was harvested by injecting 10 mL of ice-cold PBS into the peritoneal cavity followed by syringe aspiration. Cell suspensions were washed twice by centrifugation. Cell viability (over 95%) was determined using trypan blue exclusion. Macrophage numbers were adjusted to 1 × 106 cell/mL and plated 100 µL/ well in 96-well flat-bottomed tissue culture plates (UNIPARTS, Toluca, México). Cells were incubated in RPMI 1640 complete medium containing 10% FBS, and incubated for 24 h at 37 ℃ under 5% CO2 in a humidified chamber. Non-adherent cells were removed by gently washing with PBS and fresh RPMI 1640 complete medium was replaced. The efficiency of macrophage enrichment was monitored by 7AAD assay and routinely exceeded 90%. Cells were equilibrated for 24 h before commencing the experiment.

2.2.2. Reagents, antibody, and determination of cytotoxicity percentage by flow cytometry analysis

The active compound 1-5-Bromofur-2-il-2-bromo-2-nitroethene (G1), CAS number 35950-55-1, was kindly supplied from the CBQ, Sample purity was 99.93%. The percentage of formation of cytotoxicity cells was determined by evaluating 7-Amino-actinomycin D (7AAD) stained preparations of macrophages treated with the dosed chemical (G1) at 10, 8, 6, 4 and 2 µg/mL in 24 h. G1 was dissolved in dimethylsulfoxide (DMSO), Macrophages were stained with phycoerythrin (PE), labelled monoclonal antibodies according to the manufacturers' instructions. Flow cytometry was performed using a FACalibur cytometer (Becton Dickinson, México). Thereafter, FACS data were analyzed with FlowJo 7.6.5 software. Both, anti-CD14 antibody (used to label CD14 receptor) and 7 –aminoactinomycin (7-AAD) at 5 µg/mL viability solution were purchased from BD (BD Biosciences, México). In all cases, dimethyl sulfoxide (DMSO); which was purchased in turn from Sigma–Aldrich Co. (DF, México). was used as the diluting solvent, and dosage solutions were prepared immediately prior to testing. Incubations were carried out in triplicate; solvent controls were run with each experiment. Conditions of maintenance, treatment, and procedures carry out with animals, Female Balb/C mice, have been published before [39], follow the national normative [40], please see a detailed description in the supplementary material file (SM1) of this work.

3. Results and Discussion

3.1. Multiplexing model of drug effect over macrophage

3.1.1. Model training & validation

It is well known that biological outcomes in multiplex cell viability assay for drugs effect over different cellular lineages depend not only on drug structure but also on the set of assay conditions selected (mj) [42]. In this work we developed a simple High-throughput mt-QSAR model with only four variables able to assign each drug to 1 out of 2 possible activity classes: active (C = 1) or non-active compounds (C = 0); given the molecular structure and several multiplex assay conditions mj. This model is expected to give different classification probabilities of the compound for different: organisms (ot), biological assays (au), molecular or cellular targets (t_e), or standard type of activity measure (s_x). It is also desirable to use an algorithm that takes into consideration the different degrees of accuracy or level of curation (c_l) in the experimental data. We fit the classifier using LDA. The best equation found was:

$$S_i(m_j) = 5.8261 \cdot {}^{*}\mu_5^i - 0.2617\Delta\mu_5^i(o) - 2.2122 \cdot \Delta\mu_5^i(t) + 0.6819 \cdot \Delta\mu_5^i(s) - 4.0372 \quad (5)$$

$$N = 6747 \qquad R_c = 0.7 \qquad \chi^2 = 4571.16 \qquad p < 0.05$$

$S(m_j) = S(d_i, a_u, c_l, o_t, t_e, s_x)$ is a real-valued variable that scores the propensity of the drug to be active in multiplex pharmacological assays of the drug depending on the conditions selected mj. The statistical

parameters for the above equation are: Number of cases (N), Canonical Regression coefficient (R_c), Chi-square statistic (χ^2), and error level (p-level); which have to be < 0.05 [12]. The different parameters in the equation were introduced to codify specific information that is known to be determinant in the final value of biological activity. This discriminant function presented good results both in training and external validation series with overall Accuracy higher than 90%. According to previous reports in the QSAR literature [16-22] values Accuracy higher than 75% are acceptable. All the statistical data of this model are resumed in **Table 1**. The reader should be aware that N here is not number of compounds but number of statistical cases. One compound may lead to 1 or more statistical cases because it may give different outcomes for alternative biological assays carried out in diverse sets of multiplex conditions defined by the ontology m_j => (a_u, c_l, o_t, t_e, s_x). This type of ontology introduced here allows us to clearly define the multiplex conditions for one assay in our dataset following the same line of thinking used for other ontology-like datasets in the literature [43]. The above equation was written in a compact form. At follow we expand the equation n order to better explain the meaning of the different parameters:

$$S_i(m_j) = 5.8261 \cdot p(a_u) \cdot p(c_l) \cdot {}^{std}\mu_5^i - 0.2617 \cdot \left({}^{std}\mu_5^i - \left\langle {}^{std}\mu_5^i(o_t) \right\rangle \right) - 2.2122 \cdot \left({}^{std}\mu_5^i - \left\langle {}^{std}\mu_5^i(t_e) \right\rangle \right) \quad (6)$$
$$+ 0.6819 \cdot \left({}^{std}\mu_5^i - \left\langle {}^{std}\mu_5^i(s_x) \right\rangle \right) - 4.0372$$
$$N = 6747 \qquad R_c = 0.7 \qquad \chi^2 = 4571.16 \qquad p < 0.05$$

The first parameter $*\mu^i_5 = p(a) \cdot p(c) \cdot {}^{std}\mu^i_5$ codify the influence of the chemical structure of the compound over the biological activity. It is known that the spectral moment of order 5 codify information about all types of structural fragments with five or less bonds in the molecule. In addition to the topological information ${}^{w}\mu^i_5$ codify also information about the physicochemical properties of the atoms and bonds in the molecule. It depends on the type of atomic or bond weights w_{ij} used. In our equation we set w_{ij} equal to the values of standard bond distance (std) in order to incorporate geometrical information [33-35, 44, 45]. Consequently, $*\mu^i_5$ codify the effect of the structure of the drug over the biological activity but depending on the type of assay carry out. In this sense, we pre-multiplied μi5 by the parameters $p(a_u)$ and p(cl). The parameter p(a) is a probability (*a priori*) that codify the propensity of one assay to yield positive results. We defined $p(a_u) = n_l(a_u)/n_{tot}(a_u)$; where $n_1(a_u)$ and $n_{tot}(a_u)$ are the number of positive or total results for the i^{th} pharmacological assay ai in the ChEMBL dataset studied, respectively. The parameter $p(c_l)$ is a probability (*a priori*) of confidence for a given data value into the ChEMBL dataset studied. We defined p(c) as follow p(c) = 1, 0.75, or 0.5 for data values reported as being curated at expert, intermediate, or auto-curation level respectively. In **Table 2** we give some example of assays and their $p(a_u)$ values. In the Table SM1 of the online supplementary material file we list exhaustive values of these parameters.

Table 2 comes about here

The other three terms in the equation express the structural dissimilarity between one specific compound and a group of active compounds that have been assayed in specific multiplex conditions defined by the sub-ontology m_j => (o_t, t_e, s_x). We quantify this effect in terms of the deviation $\Delta\mu^i_5(m_j) = {}^{std}\mu^i_5 - \langle {}^{std}\mu^i_5(m_j) \rangle$. This deviation terms represent the hypothesis: H0 the structural dissimilarity between one compound with respect to the average of all compounds in a group predict the final behavior of the compound. For instance, $\Delta\mu^i_5(o_t) = {}^{std}\mu^i_5 - \langle {}^{std}\mu^i_5(o_t) \rangle$ measure the deviation from the average value $\langle \mu^i_5(o_t) \rangle$ of μ^i_5 for all active compounds (C = 1) assayed in the organism o_t => t = 1, 2 for Human or Mouse, respectively. The two possible values for this parameter are $\langle \mu^i_5(o_1) \rangle = 18139.7$, and $\langle \mu^i_5(o_2) \rangle = 18149.6$. This type of model able to model/interpret cross-species activity is of the major importance in order to reduce assays in humans [46]. By analogy, $\Delta\mu^i_5(t_e) = {}^{std}\mu^i_5 - \langle {}^{std}\mu i5(t_e) \rangle$ is the dissimilarity between the structure of compound i^{th} (expressed by stdμi5) with respect to all compounds active against the molecular or cellular target t_e. In **Table 3** and **Table 4** we give the values of $\langle \mu^i_5(t_e) \rangle$ and $\langle \mu^i_5(s_x) \rangle$ for the different targets or standard measure types respectively.

Please insert both Table 3 and Table 4 near here

3.1.2. Domain of application of the model

A QSAR model is only valid within its calibration domain or domain of applicability (DA), and new objects must therefore be assessed as belonging to this domain before the model is applied [47]. The valid DA can easily be defined with the LDA model, as outlined in previous works [48]. In this data set, a total of only 355 out 9000 total objects (statistical cases) fall outside of the DA. This DA may be geometrically defined as the rectangular area inside the 5% confidence bound for the ± 2 residuals interval and the

leverage limit of h = 3·p'/N = 3·(Nv + 1)/N = 3·(4 + 1)/6746 = 0.00223. Where, Nv is the number of variables in the model and N the number of cases used to train it. The DA can be visually illustrated in the so called Williams' graph (see **Figure 1**) [49]. All of the remaining 8645 objects (96.1% of the data set) fall within the valid DA. We found similar error for both train and prediction sub-sets with 6747 and 2253 objects (6747 + 2253 = 9000) respectively. Interestingly, 93.8% of drugs tested in some macrophage cytotoxicity assay lie within the DA as well. Similar behavior was found for other sub-sets of objects (see **Table 5**).

Figure 1 comes about here

In order to predict the classification of one compound one have to substitute in the High-throughput mt-QSAR model in first instance the structural parameter of the compound μ^i_5 is not sufficient to obtain different outputs for the same compound assayed in diverse conditions. In addition, we have to substitute the parameters characteristics of the given assay conditions $p(a_u)$, $p(c_l)$, $<\mu^i_5(o_t)>$, $<\mu^i_5(t_e)>$, and $<\mu^i_5(s_x)>$.

The models is expected to be more accurate for those m_j based on the more representative as possible number of cases (N_j); taking into consideration the influence of N_j in multiplex assays [50]. In **Tables 2, 3,** and **4** we report values of these parameters. In total we analyzed N_a = 1418 assays, Nt = 36 molecular or cellular targets, N_s = 46 standard types of biological activity measures. Considering that we have determined this values independently our High-throughput mt-QSAR model is able to predict a huge number of combinations of biological assay conditions m_j. However, we strongly recommend using the model only for those m_j with at least 10 known cases. The number N_j of m_j that fulfill this stronger requisite are: Na = 437 assays, N_t = 22 targets, Ns = 20. The max number of outputs with this constrain $S_{max} = N_a \times N_t \times N_s \times N_o$ = 437 x 22 x 20 x 2 = 384,560 multiplex conditions m_j. Notably, N_o = 2 is the number of organisms susceptible to be studied with this model - Human (*Homo sapiens*) and Murine (*Mus musmuculus*). Consequently, our model is expected to be successful in the predictive extrapolation of experimental data from Murine species to Human.

Table 5 comes about here

3.2. Experimental-Theoretic Study of G1 anti-microbial drug

3.2.1. Experimental results

The compound G-1 is one of the members of a new family of furylethylene derivatives with both anti-bacterial and anti-fungal properties [51]. More recently anti-parasite activity has been also reported [52]. The compound was synthesized in the laboratories of the Chemical Bioactives Center (CBQ) at the Universidad Central de Las Villas (UCLV), Cuba. Nitrovinilfurans compounds are widely used in medicine, industry and agriculture Interest in the study of these compounds has increased in recent years due to the potent microcidal activity shown by compounds with this type of chemical structure Nitrofurans constitute an important group of chemicals with antimicrobial properties that are currently used in human and veterinary medicine [53] ENREF 14.

3.2.2. Cytotoxicity assays

The cytotoxicity is defined as the response of toxicity of a compound on the cell. The kinetic cell viability measurement provides the temporal information as to when a drug of interest induces its cytotoxic effect [54]. Quantifying cell viability or cytotoxicity is crucial for understanding cancer biology, compound toxicity and cellular response to cytokines and other biological questions [55].The specific method used will greatly influence the interpretation of the data [56, 57]. In our study we used only the detection of membrane integrity by staining with 7AAD and flow cytometry. Several parameters were analyzed for dramatic views on the cytotoxicity of the drug. Viability dye 7AAD is routinely used in four-color flow cytometry assays, and therefore its use in conjunction with fixation should be carefully evaluated [58]. The analyses with flow cytometry were performed; in order to follow the percentage of live macrophages present in the macrophages populations treated with G1 at different concentrations we observed changes in the viability of the macrophages after 24 hours. The assay shows a significant increase of dead cells, Cytotoxicity (%) = 23.6%, compared to the group untreated (2.85 %) and the DMSO group (3.23%) at c_{max} = 10 µg/mL. The treatment of 6 and 8 µg/mL results in a dose-dependent significant increase in cytotoxicity (16.5%) and (19.4%) respectively (**Figure 2**). The percent of cytotoxicity is similar in concentrations 2 and 4 µg/mL (approximately 10%). It is noted further that there is an increased cytotoxicity in a dose-dependent this phenomenon has been reported in several studies using other drugs [59]. These resulted indicate slight toxicity of G1 (10 pg / mL) because the percentage of cytotoxicity calculated was 23.6% <50%. Furthermore

the estimated EC50 for this product was 21.82. [60]. In other studies with the product in lymphocyte populations the concentration 15μg/mL was observed cytotoxicity [61].

Figure 2 comes here

Identification of 'viable' or 'healthy' cells by light-scatter (a common practice as perceived in a core laboratory) is purely empirical, and relies on the shape of the Forward Scatter *vs.* Side Scatter (FSC / SSC) cluster. Essentially, gating is set on the cloud-like distribution of cells with low to medium side-scatter, excluding cells with low forward scatter and high side-scatter. Sometimes this procedure provides a remarkable correlation between the percentage of excluded cells and the percentage of dead cells as identified by a viability stain such as 7-aminoactinomycin D (7-AAD) or propidium iodide (PI) [62]. Secondly, we investigated the MFI on highly homogenous macrophages populations defined by the expression of CD14, obtained from the peritoneal macrophage of healthy mouse. These macrophages were exposed to different concentrations of G1 with DMSO. In **Figure 3,** we depict a pseudo-color smooth projection of mean Intensity Fluorescence (MFI). This Figure represents plot FSC vs. SSC after exposure of G1 at 10μg/mL. In Figure 3A shows a 11.4% of the cell population of the total acquisition. The figure 3B has shown the cell population alone. The 3C and 3D shows the regions (R2 and R3) of macrophages labeled with CD14Pe (98.6%) of the total population. This figure shows the similarity of the dispersion values using foward (FSC) and side scatter (SSC). Cytotoxicity studies were used both forward scatter and side scatter because has shown high correlation [63]. This same methodology was used to represent the regions (R4, R5) stained with 7AAD this cell population represents 53.3% of total (3E; 3F).The SSC-H +, FSC-H + shows no significant differences at $p < 0.05$ compared to control. On the other hand no significant differences are observed in CD14Pe+7AAD+ at 10μg/mL compared to control (CN) which shows that there is some cytotoxicity in macrophages thus corresponds to the results of cytotoxicity percentage calculated for this population. It is known from literature that the forward light scatter versus side scatter 90th is a measure of cell size and cell granularity respectively, the Latter Being dependent upon the presence of intracellular structures that change the refractive index of light [64]. The number of labeled cells with 7AAD indicates a slight cytotoxicity G1 but the actual calculation of cytotoxicity was 23.6%. Statistical analysis also confirms that there are slight cytotoxicity since no significant differences between the treated and control.

Figure 3 comes here

In **Table 6**, we show the average values of Mean fluorescence intensity (MFI) in SSC and/or FSC scattering mode, for all samples (Negative Control, DMSO, and CD14Pe phenotypic marker macrophages exposed to G1).for all concentration of product G1. MFI and cell count (Event count) in FSC scattering mode give an idea of cellular size, while the same parameters but in SSC scattering mode measure internal cellular damage [65]. The events average in the dose 10 μg/mL was 1034. Moreover, the respective averages of MIF in quadrant 2 (Q2) for FSC and SSC are 550.83 ± 103 and 313.83 ± 94.4 (see **Figure 3**). In conclusion it was observed events classified by size and granularity for this concentration.

At a given a concentration, each experiment was carried out two times (repeated two times) using different animals (three animals) and the measure obtained for each animal was replicated three times (see materials and methods). **Table 6** we show the averages of repetitions. We used the software STATISTICA for both means and ANOVA analysis [66]. The results show not significant differences ($p \leq 0.05$) between the mean values of MFI for G1-treated samples at different concentrations (2 - 10 μg/mL) with respects to the negative control (NC) and DMSO groups. In particular, there are not significant differences between the mean values of MFI for G1-treated samples labeled with anti-CD14Pe and stained with 7AAD (living macrophages) with respects to both control groups (see **Table 6**). The ANOVA analysis was carry out applying Tukey's method. We confirm that there not significant differences for treated samples of living macrophages with respect to control groups. The numbers of cells are in a range between 500 and 1500 events in general. The Figure 4 shows significant differences between groups.

Table 6 and Figure 4 comes about here

In addition, CD14 PE was used as a macrophages marker in the presence of 7AAD; as described in the Materials and methods section. In total, 52.7% of macrophages were marked with CD14Pe and 7AAD. The MFI average was 32.55 ± 9.3 and 130.35 ± 29.4 respectively. It means that more than 45% of macrophage were still alive after treatment with G1 at the higher concentration c_{max} = 10 μg/mL. In **Figure 5**, we show two parameters CD14Pe (FL2) and 7AAD (FL3) of the population of macrophages at this concentration. These additional results are consistent with the previous paragraph.

Figure 5 comes here

Finally, EC_{50} calculations using different methodologies have been shown below. The results show that the best dose-response curve was the Five parameter logistic ($1/Y^2$) with an $R^2 = 0.956$. The Root Mean Square Error (RMSE = 0) and $EC_{50} = 21.82$. This study calculated the EC_{50} being observed that these values differ with relation to the methodology applied (**Tabla 7**). Fitting nonlinear models to observed data is often complicated by non-constant or heterogeneous variability. Heterogeneous variability or *heteroscedasticity* occurs in most types of observed data. This is especially true for biochemical assays where concentration or dose is the predictor. The best curve fit is reached when the curve is pulled as close as possible to each data point without breaking the actual curve model. The nonlinear least square algorithm accomplishes this task. The nonlinear (or linear) least square algorithm assumes that all points have the same variability, so all points influence the curve fit equally [67]. ENREF 49 The nature of the data entails a variation of the dependent variable that changes over the data is known as heterocedasticity. Many methods of regression analysis is based on the assumption of equal variances, but MasterPlex ReaderFit software used to calculate the EC_{50} offers 4 different weighting algorithms to account for heterocedasticity. The five parameter logist is the optimal model equation and weighting algorithm with different parameters (Root Mean Square Error (RMSE), R-Square, and Standard Deviation of % Recovery). One way to counterbalance no constant variability is to make them constant again. To accomplish this, weights are assigned to each standard sample data point.These weights are designed to approximate the way measurement errors are distributed. By applying weighting, points on the lower part of the curve are given more influence on the curve again. One of algorithms of assigning weights: is $1/Y^2$ – Minimizes residuals (errors) based on relative Mean Fluorescence Intensity and Relative light Unit, (MFI/RLU) values. Many functions have been tried as curve models for immunoassays, but few of them possess all of these properties. The need for a curve model that accommodates asymmetry has been necessitated by improvements in instrument and laboratory technology. The development of sandwich assays led to dose–response curves that tend to be more asymmetric than earlier types of assays. Additionally, because of improvements in signal-to-noise ratios, asymmetry is an issue even for assays whose dose–response relationships are not as highly asymmetric. The reason for this is that even modest levels of lack-of-fit error caused by fitting mildly asymmetric data to a symmetric model can dominate the pure error due to random variation in low-noise modern assays. For symmetric immunoassay and bioassay data, it can be argued that no curve model has been as successful as the four-parameter logistic function. Despite its utility, the 4PL function is generally not an adequate curve model for much of the asymmetric response data commonly observed in immunoassay and bioassay applications. The five-parameter logistic function, which includes a fifth parameter, permits asymmetry to be effectively modeled [68]. The formula for analysis is:

$$\left(\frac{1}{MFI_i^{\,2}}\right) = 47.85819 - \frac{130.509}{\left(1+\left(\left(\frac{c_i}{21.82}\right)^{-7.17925}\right)\right)^{0.51856}} \qquad (7)$$

In conclusion two of the highest concentrations showed some cytotoxicity but I note that the EC_{50} is above the concentrations used in our study. In general, the cytotoxicity EC_{50} values for each compound were lower after 24h exposure. The best method used for the analysis was the 5PL using ($1/Y^2$). The **Figure 6** displays MasterPlex program used to calculate the EC_{50}. It observes where the parameters that makes up the formula of equation 5PL.

Figure 6 come here

This study evaluated the cytotoxicity by calculating the percentage of cytotoxicity and EC50 by Equation 5 PL. The best equation showed a R = 0.95. A comparison between the MFI of the groups treated with the negative control for parameters that reads the flow cytometer. The evaluated product showed slight cytotoxicity

3.2.3. Prediction of G1 cytotoxicity for other assays

In total we predicted 1,265 multiplexing assay endpoints for G1 biological activities. Notably the model predict very low probability (0.28) for G1 cytotoxicity (cutoff of $TC_{50} < 100$ μM) against human macrophages. The model also predicts only 7 positive endpoints for G1 out of 1,251 cytotoxicity assays (0.56% of probability of cytotoxicity in multiple assays), see **Table 8**. Interestingly, the predictive

probability obtained for this compound in the cytotoxicity assay against WEHI cell line was 0.84. WEHI cell line is a biological model for leukemia and has been used to test anti-carcinogenic activity [69].

Several predictions were conducted in J774 macrophages cell line (170 assays). In all cases the model predicts low probability of G1 to present cytotoxicity effect against J774 macrophages. Macrophages are highly motile cells capable of chemotaxis and pathogen engulfment [70]. J774 and Raw 264.7 macrophage cell lines; which are well-established model systems in cell biology and immunology. The resistance of passive J774 cells to expansion of their surface areas is about one order of magnitude higher than that of human neutrophils [71]. The J774 has been used [72] to assess drugs anti-parasitic activity against diverse parasite species such as Plasmodium parasites, *Trypanosoma brucei brucei*, and *Leishmania mexicana mexicana*. Other research reported this cell line to assess anti-leishmanial activity of compounds against both the promastigote and intracellular amastigote stages of *Leishmania infantum* and *L. donovani* [73]. This is of a great importance if we know that the G1 have been demonstrated experimentally to be active against bacteria and parasites [74].

Table 8 comes about here

Some of these positive results in predictive tests included the evaluation of the cytotoxicity in RAW264.7 (Monocytic-macrophage leukemia cells) cell lines. The same **Table 8** shows that the G1 could inhibit with 89% of probability such cells in some specific assay conditions. However, the model predicts low probabilities of cytotoxicity in other assays using RAW264.7 cell. The RAW264.7 cell line was derived about 30 years ago from a tumor developing cells in a BAB/14 mouse, a BALB/c IgH congenic strain, inoculated with Abelson murine leukemia virus (MuLV), a defective transforming virus containing the v-abl tyrosine kinase oncogene, and replication-competent Moloney (Mo-MuLV) that served as helper virus [75] In addition, because of ease of cell propagation, high efficiency for DNA transfection, sensitivity to RNA interference, possession of receptors for many relevant ligands, and other properties, RAW264.7 has been chosen by the Alliance for Cellular Signaling as the primary experimental system for their large-scale study of signaling pathways [68, 76].

4. Conclusion

mt-QSAR techniques are become an important tool for prediction of cytotoxicity and High-throughput Screening (HTS) of drugs to rationalize drug discovery process. The results obtained through this new theoretically mt-QSAR methodology coincides with the experimental data G1 drug, in this sense; we can conclude that our model applies both to predict the biological activity of a drug as its cytotoxicity. This work adds a new tool to the existing pool of few methods useful for multi-target HTS of ChEMBL and other libraries of compounds towards drug discovery.

Acknowledgements

All authors acknowledge the project DCS-UQROO PIFI (P/PIFI-2012-23MSU0140Z-09 DCS).

Appendix A. Supplementary data

Supplementary data to this article can be found online

References

1. Mingoia R.T., N.D.L., Ching-Hui Yang, Xing Han., (2007). Primary culture of rat hepatocytes in 96-well plates: Effects of extracellular matrix conWguration on cytochrome P450 enzyme activity and inducibility, and its application in in vitro cytotoxicity screening. Toxicol. in Vitro *21*, 165–173.
2. Weyermann J., L.D., Zimmer Andreas., (2005). A practical note on the use of cytotoxicity assays. Int. J. Pharm. *288*, 369–376.
3. Tripathi, A., and Sodhi, A. (2009). Growth hormone-induced production of cytokines in murine peritoneal macrophages in vitro: role of JAK/STAT, PI3K, PKC and MAP kinases. Immunobiology *214*, 430-440.
4. Cunnick Jess., K.P., Young Jin Cho, Groffen J, Heisterkamp Nora., (2006). Use of bone marrow-derived macrophages to model murine innate immune responses. Journal of Immunological Methods., *311*, 96–105.

5. Barnett J. B. and Brundage Kathleen M. (2010). Evaluating Macrophages in Immunotoxicity Testing. In Immunotoxicity Testing Methods and Protocols .Methods in Molecular Biology Volume 598.
6. Gaulton, A., Bellis, L.J., Bento, A.P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., et al. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. *40*, D1100-1107.
7. Mok, N.Y., and Brenk, R. (2011). Mining the ChEMBL database: an efficient chemoinformatics workflow for assembling an ion channel-focused screening library. Journal of chemical information and modeling *51*, 2449-2454.
8. Kuzmin V.E., M.E.N., Artemenko A.G., Gorb L., Qasim M., Leszczynski J.., (2008). The effect of nitroaromatics' composition on their toxicity in vivo: Novel, efficient non-additive 1D QSAR analysis. Chemosphere *72*, 1373–1380.
9. Riera-Fernandez, I., Martin-Romalde, R., Prado-Prado, F.J., Escobar, M., Munteanu, C.R., Concu, R., Duardo-Sanchez, A., and Gonzalez-Diaz, H. (2012). From QSAR models of Drugs to Complex Networks: State-of-Art Review and Introduction of New Markov-Spectral Moments Indices. Curr. Top. Med. Chem.
10. Prado-Prado, F., Garcia-Mera, X., Escobar, M., Sobarzo-Sanchez, E., Yanez, M., Riera-Fernandez, P., and Gonzalez-Diaz, H. (2011). 2D MI-DRAGON: a new predictor for protein-ligands interactions and theoretic-experimental studies of US FDA drug-target network, oxoisoaporphine inhibitors for MAO-A and human parasite proteins. Eur. J. Med. Chem. *46*, 5838-5851.
11. Tenorio-Borroto, E., Garcia-Mera, X., Penuelas-Rivas, C.G., Vasquez-Chagoyan, J.C., Prado-Prado, F.J., Castanedo, N., and Gonzalez-Diaz, H. (2013). Entropy Model For Multiplex Drug-Target Interaction Endpoints Of Drug Immunotoxicity. Curr Top Med Chem.
12. Van Waterbeemd, H. (1995). Chemometric methods in molecular design, Volume 2, (New York: Wiley-VCH).
13. Tenorio-Borroto, E., Penuelas-Rivas, C.G., Vasquez-Chagoyan, J.C., Prado-Pradoa, F.J., Garcia-Mera, X., and Gonzalez-Diaz, H. (2012). Immunotoxicity, flow cytometry, and chemoinformatics: review, bibliometric analysis, and new QSAR model of drug effects over macrophages. Curr Top Med Chem *12*, 1815-1833.
14. Heikamp, K., and Bajorath, J. (2011). Large-scale similarity search profiling of ChEMBL compound data sets. Journal of chemical information and modeling *51*, 1831-1839.
15. StatSoft.Inc. (2002). STATISTICA (data analysis software system), version 6.0, www.statsoft.com.Statsoft, Inc., 6.0 Edition.
16. Patankar, S.J., and Jurs, P.C. (2003). Classification of inhibitors of protein tyrosine phosphatase 1B using molecular structure based descriptors. J. Chem. Inf. Comput. Sci. *43*, 885-899.
17. Garcia-Garcia, A., Galvez, J., de Julian-Ortiz, J.V., Garcia-Domenech, R., Munoz, C., Guna, R., and Borras, R. (2004). New agents active against Mycobacterium avium complex selected by molecular topology: a virtual screening method. J. Antimicrob. Chemother. *53*, 65-73.
18. Marrero-Ponce, Y., Castillo-Garit, J.A., Olazabal, E., Serrano, H.S., Morales, A., Castanedo, N., Ibarra-Velarde, F., Huesca-Guillen, A., Sanchez, A.M., Torrens, F., et al. (2005). Atom, atom-type and total molecular linear indices as a promising approach for bioorganic and medicinal chemistry: theoretical and experimental assessment of a novel method for virtual screening and rational design of new lead anthelmintic. Bioorg. Med. Chem. *13*, 1005-1020.
19. Marrero-Ponce, Y., Machado-Tugores, Y., Pereira, D.M., Escario, J.A., Barrio, A.G., Nogal-Ruiz, J.J., Ochoa, C., Aran, V.J., Martinez-Fernandez, A.R., Sanchez, R.N., et al. (2005). A computer-based approach to the rational discovery of new trichomonacidal drugs by atom-type linear indices. Curr Drug Discov Technol *2*, 245-265.
20. Casanola-Martin, G.M., Marrero-Ponce, Y., Khan, M.T., Ather, A., Sultan, S., Torrens, F., and Rotondo, R. (2007). TOMOCOMD-CARDD descriptors-based virtual screening of tyrosinase inhibitors: evaluation of different classification model combinations using bond-based linear indices. Bioorg. Med. Chem. *15*, 1483-1503.
21. Casanola-Martin, G.M., Marrero-Ponce, Y., Tareq Hassan Khan, M., Torrens, F., Perez-Gimenez, F., and Rescigno, A. (2008). Atom- and bond-based 2D TOMOCOMD-CARDD approach and ligand-

based virtual screening for the drug discovery of new tyrosinase inhibitors. J Biomol Screen *13*, 1014-1024.

22. Casanola-Martin, G.M., Marrero-Ponce, Y., Khan, M.T., Khan, S.B., Torrens, F., Perez-Jimenez, F., Rescigno, A., and Abad, C. (2010). Bond-based 2D quadratic fingerprints in QSAR studies: virtual and in vitro tyrosinase inhibitory activity elucidation. Chemical biology & drug design *76*, 538-545.
23. Estrada, E., Uriarte, E., Montero, A., Teijeira, M., Santana, L., and De Clercq, E. (2000). A novel approach for the virtual screening and rational design of anticancer compounds. J Med Chem *43*, 1975-1985.
24. Estrada, E., and Peña, A. (2000). In silico studies for the rational discovery of anticonvulsant compounds. Bioorg Med Chem *8*, 2755-2770.
25. Estrada, E., Gutierrez, Y., and González, H. (2000). Modeling diamagnetic and magnetooptic properties of organic compounds with the TOSS-MODE approach. J Chem Inf Comput Sci *40*, 1386-1399.
26. Estrada, E. (2000). On the topological sub-structural molecular design (TOSS-MODE) in QSPR/QSAR and drug design research. SAR QSAR Environ Res *11*, 55-73.
27. Estrada, E., and Uriarte, E. (2001). Recent advances on the role of topological indices in drug discovery research. Curr Med Chem *8*, 1573-1588.
28. Estrada, E., Vilar, S., Uriarte, E., and Gutierrez, Y. (2002). In silico studies toward the discovery of new anti-HIV nucleoside compounds with the use of TOPS-MODE and 2D/3D connectivity indices. 1. Pyrimidyl derivatives. J Chem Inf Comput Sci *42*, 1194-1203.
29. Estrada, E., Quincoces, J.A., and Patlewicz, G. (2004). Creating molecular diversity from antioxidants in Brazilian propolis. Combination of TOPS-MODE QSAR and virtual structure generation. Mol Divers *8*, 21-33.
30. Estrada, E., Uriarte, E., Molina, E., Simón-Manso, Y., and Milne, G.W. (2006). An integrated in silico analysis of drug-binding to human serum albumin. J Chem Inf Model *46*, 2709-2724.
31. Estrada, E., Molina, E., Nodarse, D., and Uriarte, E. (2010). Structural contributions of substrates to their binding to P-Glycoprotein. A TOPS-MODE approach. Curr Pharm Des *16*, 2676-2709.
32. Pisco, L., Kordian, M., Peseke, K., Feist, H., Michalik, D., Estrada, E., Carvalho, J., Hamilton, G., Rando, D., and Quincoces, J. (2006). Synthesis of compounds with antiproliferative activity as analogues of prenylated natural products existing in Brazilian propolis. Eur J Med Chem *41*, 401-407.
33. Estrada, E., and Uriarte, E. (2001). Quantitative structure--toxicity relationships using TOPS-MODE. 1. Nitrobenzene toxicity to Tetrahymena pyriformis. SAR QSAR Environ Res *12*, 309-324.
34. Estrada, E., Molina, E., and Uriarte, E. (2001). Quantitative structure-toxicity relationships using TOPS-MODE. 2. Neurotoxicity of a non-congeneric series of solvents. SAR QSAR Environ Res *12*, 445-459.
35. Estrada, E., Uriarte, E., Gutierrez, Y., and González-Díaz, H. (2003). Quantitative structure-toxicity relationships using TOPS-MODE. 3. Structural factors influencing the permeability of commercial solvents through living human skin. SAR QSAR Environ Res *14*, 145-163.
36. Estrada, E., Patlewicz, G., and Gutierrez, Y. (2004). From knowledge generation to knowledge archive. A general strategy using TOPS-MODE with DEREK to formulate new alerts for skin sensitization. J Chem Inf Comput Sci *44*, 688-698.
37. Marzaro, G., Chilin, A., Guiotto, A., Uriarte, E., Brun, P., Castagliuolo, I., Tonus, F., and Gonzalez-Diaz, H. (2011). Using the TOPS-MODE approach to fit multi-target QSAR models for tyrosine kinases inhibitors. European Journal of Medicinal Chemistry *46*, 2185-2192.
38. Molina, E., Sobarzo-Sánchez, E., Speck-Planche, A., and Matos, M.J. (2012). Monoamino Oxidase A: an interesting pharmacological target for the development of multi-target QSAR. Mini Rev Med Chem.
39. Tenorio-Borroto, E., Penuelas Rivas, C.G., Vasquez Chagoyan, J.C., Castanedo, N., Prado-Prado, F.J., Garcia-Mera, X., and Gonzalez-Diaz, H. (2012). ANN multiplexing model of drugs effect on macrophages; theoretical and flow cytometry study on the cytotoxicity of the anti-microbial drug G1 in spleen. Bioorg Med Chem *20*, 6181-6194.

40. NOM, S. (1999). Especificaciones técnicas para la producción, cuidado y uso de los animales de laboratorio. SENACICA *ZOO*, 1-58.
41. Tario J. D.., M.K.A., Pan Dalin., Munson Mark E.,, and K, a.W.P. (2011). Tracking Immune Cell Proliferation and Cytotoxic Potential Using Flow Cytometry. In Flow Cytometry Protocols, Methods in Molecular Biology.,, Volume 699, T.S.H.a.R.G.H. (eds.), ed.
42. Gerets, H.H., Dhalluin, S., and Atienzar, F.A. (2011). Multiplexing cell viability assays. Methods Mol. Biol. *740*, 91-101.
43. Martinez-Romero, M., Vazquez-Naya, J.M., Rabunal, J.R., Pita-Fernandez, S., Macenlle, R., Castro-Alvarino, J., Lopez-Roses, L., Ulla, J.L., Martinez-Calvo, A.V., Vazquez, S., et al. (2010). Artificial intelligence techniques for colorectal cancer drug metabolism: ontology and complex network. Curr Drug Metab *11*, 347-368.
44. Estrada, E., Vilar, S., Uriarte, E., and Gutierrez, Y. (2002). In silico studies toward the discovery of new anti-HIV nucleoside compounds with the use of TOPS-MODE and 2D/3D connectivity indices. 1. Pyrimidyl derivatives. J. Chem. Inf. Comput. Sci. *42*, 1194-1203.
45. Estrada, E., and González-Díaz, H. (2003). What are the limits of applicability for graph theoretic descriptors in QSPR/QSAR? Modeling dipole moments of aromatic compounds with TOPS-MODE descriptors. J. Chem. Inf. Comput. Sci. *43*, 75-84.
46. Meinel, T., Schweiger, M.R., Ludewig, A.H., Chenna, R., Krobitsch, S., and Herwig, R. (2011). Ortho2ExpressMatrix--a web server that interprets cross-species gene expression data by gene family information. BMC Genomics *12*, 483.
47. Oberg, T. (2004). A QSAR for baseline toxicity: validation, domain of application, and prediction. Chem. Res. Toxicol. *17*, 1630-1637.
48. Gonzalez-Diaz, H., Vilar, S., Santana, L., Podda, G., and Uriarte, E. (2007). On the applicability of QSAR for recognition of miRNA bioorganic structures at early stages of organism and cell development: embryo and stem cells. Bioorg. Med. Chem. *15*, 2544-2550.
49. Papa, E., and Gramatica, P. (2008). Externally validated QSPR modelling of VOC tropospheric oxidation by NO3 radicals. SAR QSAR Environ Res *19*, 655-668.
50. Atienzar, F.A., Tilmant, K., Gerets, H.H., Toussaint, G., Speeckaert, S., Hanon, E., Depelchin, O., and Dhalluin, S. (2011). The use of real-time cell analyzer technology in drug discovery: defining optimal cell culture conditions and assay reproducibility with different adherent cellular models. J. Biomol. Screen. *16*, 575-587.
51. Blondeau, J.M., Castanedo, N., Gonzalez, O., Mendina, R., and Silveira, E. (1999). In vitro evaluation of G1: a novel antimicrobial compound. Int. J. Antimicrob. Agents *11*, 163-166.
52. Marrero-Ponce Y, C.-G.J., Olazabal E, Serrano HS, Morales A, Castañedo N, Ibarra-Velarde F, Huesca-Guillen A, Sánchez AM, Torrens F, Castro EA (2005). Atom, atom-type and total molecular linear indices as a promising approach for bioorganic and medicinal chemistry: theoretical and experimental assessment of a novel method for virtual screening and rational design of new lead anthelmintic. Bioorg Med Chem. *13*, 1005-1020.
53. Perez Machado Giselle., G.B.J.I., Castañedo N., Creus A., and Marcos R., (2004). In vitro genotoxicity testing of the furylethylene derivative UC-245 in human cells. Mutagenesis., *19*, 75-80.
54. Colombo P, G.K., Iatropoulos M, Brughera M., (2001). Toxicological testing of cytotoxic drugs (review). Int J Oncol. *19*, 1021-1028.
55. Skowron, J., and Zapor, L. (2004). Cytotoxicity of resorcinol under short- and long-term exposure in vitro. Int J Occup Saf Ergon *10*, 147-156.
56. Cao LF, K.L., Tran V, Mi S, Jensen MC, Blanchard S, Kalos M. (2010). Development and application of a multiplexable flow cytometry-based assay to quantify cell-mediated cytolysis. Cytometry A. *77*, 534-545.
57. Riss TL, M.R. (2004). Use of multiple assay endpoints to investigate the effects of incubation time, dose of toxin, and plating density in cell-based cytotoxicity assays. Assay Drug Dev Technol. *2*, 51-62.
58. Jacques Nathalie , V.N., Conforti Rosa, Griscelli Franck , Lecluse Yann., Laplanche Agnes., Malka David., Vielh Philippe., Farace Françoise., (2008). Quantification of circulating mature endothelial

cells using a whole blood four-color flow cytometric assay. Journal of Immunological Methods., *337*, 132–143.

59. Savaşan S, B.S., Ozdemir O, Hamre M, Asselin B, Pullen J, Ravindranath Y. (2005). Evaluation of cytotoxicity by flow cytometric drug sensitivity assay in childhood T-cell acute lymphoblastic leukemia. Leuk Lymphoma. *46*, 833-840.
60. OECD, E.D. (2010). Guidance document on using cytotoxicity tests to estimate starting doses for acute oral systemic toxicity tests. OECD Series on Testing and Assessment *20*, 1-54.
61. González Borroto JI, P.M.G., Creus A, Marcos R. (2005). Comparative genotoxic evaluation of 2-furylethylenes and 5-nitrofurans by using the comet assay in TK6 cells. Mutagenesis. *20*, 193-197.
62. Petrunkina A.M., H.R.A.P. (2011). Mathematical analysis of mis-estimation of cell subsets in flow cytometry: Viability staining revisited. Journal of Immunological Methods., *368*, 71–79.
63. Veselá R, D.L., Pytlík R, Rychtrmocová H, Marečková H, Trněný M. (2011). The evaluation of survival and proliferation of lymphocytes in autologous mixed leukocyte reaction with dendritic cells. The comparison of incorporation of (3)H-thymidine and differential gating method. Cell Immunol. *271*, 78-84. .
64. McGowan P, N.N., Wimmer J, Williams D, Wen J, Li M, Ewton A, Curry C, Zu Y, Sheehan A, Chang CC. (2012). Differentiating Between Burkitt Lymphoma and CD10+ Diffuse Large B-Cell Lymphoma: The Role of Commonly Used Flow Cytometry Cell Markers and the Application of a Multiparameter Scoring System. Am J Clin Pathol, *137*, 665-670.
65. Gorczyca W, S.Z., Cronin W, Li X, Mau S, Tugulea S. (2011). Immunophenotypic pattern of myeloid populations by flow cytometry analysis. Methods Cell Biol. *103*, 221-266.
66. Hill, T., and Lewicki, P. (2006). STATISTICS Methods and Applications. A Comprehensive Reference for Science, Industry and Data Mining, Volume 1, (Tulsa: StatSoft).
67. Manivannan, E., and Prasanna, S. (2005). First QSAR report on FSH receptor antagonistic activity: quantitative investigations on physico-chemical and structural features among 6-amino-4-phenyltetrahydroquinoline derivatives. Bioorg. Med. Chem. Lett. *15*, 4496-4501.
68. Shin KJ, W.E., Zavzavadjian JR, Santat LA, Liu J, Hwang JI, Rebres R, Roach T, Seaman W, Simon MI, Fraser ID. (2006). A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. Proc Natl Acad Sci USA *103*, 13759-13764.
69. Lin CC, K.C., Lee MH, Hsu SC, Huang AC, Tang NY, Lin JP, Yang JS, Lu CC, Chiang JH, Chueh FS, Chung JG (2011). Extract of Hedyotis diffusa Willd influences murine leukemia WEHI-3 cells in vivo as well as promoting T- and B-cell proliferation in leukemic mice. In Vivo. *25*, 633-640.
70. Costa Lima S, R.V., Garrido J, Borges F, Kong Thoo Lin P, Cordeiro da Silva A. (2012). In vitro evaluation of bisnaphthalimidopropyl derivatives loaded into pegylated nanoparticles against Leishmania infantum protozoa. Int J Antimicrob Agents.
71. Lam J., H.M., Dembo M., and Heinrich V., (2009). Baseline Mechanical Characterization of J774 Macrophages. Biophysical Journal Volume., *96*, 248–254.
72. Ganfon H, B.J., Tchinda AT, Gbaguidi F, Gbenou J, Moudachirou M, Frédérich M, Quetin-Leclercq J. (2012). Antiparasitic activities of two sesquiterpenic lactones isolated from Acanthospermum hispidum D.C. J Ethnopharmacol.
73. Wert L, A.S., Corral MJ, Sánchez-Fortún S, Yli-Kauhaluoma J, Alunda JM. (2011). Toxicity of betulin derivatives and in vitro effect on promastigotes and amastigotes of Leishmania infantum and L. donovani. J Antibiot (Tokyo). *64*, 475-481.
74. Marrero-Ponce Y, C.-G.J., Olazabal E, Serrano HS, Morales A, Castañedo N, Ibarra-Velarde F, Huesca-Guillen A, Jorge E, del Valle A, Torrens F, Castro EA. (2004). TOMOCOMD-CARDD, a novel approach for computer-aided 'rational' drug design: I. Theoretical and experimental assessment of a promising method for computational screening and in silico design of new anthelmintic compounds. J Comput Aided Mol Des. *18*, 615-634.
75. Raschke WC, B.S., Ralph P, Nakoinz I. (1978). Functional macrophage cell lines transformed by Abelson leukemia virus. Cell. *15*, 261-267.
76. Park HY, K.G., Hyun JW, Hwang HJ, Kim ND, Kim BW, Choi YH. (2012). 7,8-Dihydroxyflavone exhibits anti-inflammatory properties by downregulating the NF-κB and MAPK signaling pathways in lipopolysaccharide-treated RAW264.7 cells. Int J Mol Med. .

TABLES TO BE INSERTED IN THE TEXT

Table 1. Overall results of the classification model and comparison with othe mt-QSTR models

Descriptors	Technique	Set	Stat. [a]	%	Sub-set	n_0	n_1	Ref.
Spectral Moments	LDA	t	Sp	97.4	n_0	3438	90	This work
			Sn	85.6	n_1	464	2755	
			Ac	91.8	total			
		cv	Sp	97.9	n_0	1138	25	
			Sn	85.0	n_1	163	927	
			Ac	91.7	total			
	LNN	t	Sp	92.5	n_0	3262	266	[39]
			Sn	93.9	n_1	197	3022	
			Ac	93.1	total			
		cv	Sn	93.1	n_1	1015	75	
			Sp	92.2	n_0	91	1072	
			Ac	92.6	total			
	MLP	t	Sn	94.3	n_1	3037	182	
			Sp	92.8	n_0	255	3273	
			Ac	93.5	total			
		cv	Sn	94.1	n_1	1026	64	
			Sp	93.2	n_0	79	1084	
			Ac	93.7	total			
Stochastic moments	LDA	t	Sp	94.5	n_0	3325	195	[13]
			Sn	82.9	n_1	554	2684	
			Ac	88.9	total			
		cv	Sp	95.3	n_0	1116	55	
			Sn	83.3	n_1	179	893	
			Ac	89.6	total			
Entropies	LDA	t	Sp	96.5	n_0	3398	122	[11]
			Sn	85.1	n_1	482	2756	
			Ac	91.0	total			
		cv	Sp	96.2	n_0	1127	44	
			Sn	85.6	n_1	154	918	
			Ac	91.1	total			

[a] Sensitivity = Sn% = 100 * n_1(correct)/n_1(total), Specificity = Sp% = 100 * n_0(correct)/n_0(total), and Accuracy = Ac = 100 * n(correct)/n(total) = 100 * [n_1(correct) + n_0(correct)]/n(total) for different m_j = organism, endpoint types, or targets.

Table 2. Some examples p(a) values for different assays

ID of a_u [i]	$p(a_u)$	n_1	n_{tot}	Cutoff	Relation	Type	Units	Assay Description
1002955	0.263	54	208	81814.3	>	K_i	nM	Inhibition of MMP12
964734	0.263	54	208	81814.3	<	IC_{50}	nM	Inhibition of MMP12
924957	0.263	54	208	81814.3	<	$\log(1/K_i)$		Inhibition of MMP12
970762	0.743	74	100	21281.7	<	IC50	nM	Inhibition of TGH
660813	0.711	68	96	2992.4	<	IC_{50}	nM	Inhibitory activity against recombinant human Chemokine receptor type 3 (CCR3) expressed in chinese hamster ovary cells
1776768	0.821	77	94	55936.7	<	ID_{50}	nM	Cytotoxicity against mouse J774 cells
1261026	0.821	77	94	55936.7	<	EC_{50}	nM	Cytotoxicity against mouse J774 cells
940865	0.021	1	93	30.1	>	Inhibition	%	Inhibition of CCR1 at 10 uM
1674458	0.430	39	92	73346.2	<	TC_{50}	uM	Cytotoxicity against mouse RAW264.7 cells after 24 hrs by MTT assay
1175699	0.430	39	92	73346.2	<	IC_{50}	nM	Cytotoxicity against mouse RAW264.7 cells after 24 hrs by MTT assay
1657211	0.430	39	92	73346.2	<	IC_{50}	ug mL-1	Cytotoxicity against mouse RAW264.7 cells after 24 hrs by MTT assay
860201	0.224	16	75	1961.4	>	K_i	nM	Inhibition of CSF1R
1025517	0.224	16	75	1961.4	<	IC_{50}	nM	Inhibition of CSF1R
1664436	0.413	30	74	12.7	>	Inhibition	%	Inhibition of mouse recombinant iNOS at 1 mM after 40 mins by colorimetric assay
867926	0.840	62	74	229.4	<	IC_{50}	nM	Inhibition of LPS-induced TNFalpha production in human monocytes
1285558	0.222	15	71	178344.1	>	K_i	nM	Inhibition of mouse recombinant iNOS
957262	0.222	15	71	178344.1	<	IC_{50}	nM	Inhibition of mouse recombinant iNOS
921708	0.130	8	68	2046.7	>	Selectivity ratio		Inhibition of cFms
998565	0.130	8	68	2046.7	<	IC_{50}	nM	Inhibition of cFMS

[i] ChEMBL ID for the assay a_u

Table 3. Values of $<\mu^{i}_{5}(t_e)>$ for all molecular or cellular targets studied

t_e	Target name	$<\mu^{i}_{5}(t_e)>$	n_1	n_{total}
1	RAW264.7 (Monocytic-macrophage leukemia cells)	20060.3	1630	3376
2	C-C chemokine receptor type 3	23855.95	700	1185
3	J774 (Macrophage cells)	16309.1	601	1001
4	Cyclooxygenase-2	14102.64	558	1061
5	C-C chemokine receptor type 1	20864.24	440	825
6	Nitric oxide synthase, inducible	13168.54	420	1082
7	J774.A1 (Macrophage cells)	22090.65	375	694
8	MCSFreceptor	16919.75	343	752
9	Matrix metalloproteinase 12	13775.03	280	566
10	Acyl coenzyme A:cholesterol acyltransferase	12571.06	271	486
11	Macrophage migration inhibitory factor	11088.6	257	461
12	Macrophage-stimulating protein receptor	16863.76	74	159
13	Monocytes	12711.64	68	84
14	Dipeptidyl peptidase IV	16785.55	50	116
15	EL4 (Thymoma cells)	24952.45	48	128
16	Interleukin-8	15053.32	40	107
17	Interleukin-5	21910.88	28	74
18	C-C motif chemokine 5	32323.55	27	34
19	Macrophage colony-stimulating factor 1 receptor	25464.67	21	29
20	RAC-alpha serine/threonine-protein kinase	13962.93	12	38
21	Serine/threonine-protein kinase TAO3	21791.05	12	26
22	PMNL (Polymorphonuclear leukocytes)	18763.72	12	15
23	Macrophages	42504.04	9	24
24	Monocytes (Monocytic cells)	18649.45	7	15
25	Scavenger receptor type A	48625.99	6	21
26	eosinophils (Eosinophils)	15608.45	6	11
27	WEHI (Macrophages)	13590.42	5	8
28	Human macrophage cell line	16607.23	4	6
29	EOL1 (Eosinophilic cells)	8907.19	2	6
30	Granulocyte colony stimulating factor receptor	23935.49	1	2
31	Macrophage scavenger receptor types I and II	11803.07	1	2
32	Macrophage metalloelastase	16403.66	1	2

Table 4. Values of $<\mu^i_5(s)>$ for different standard type measures of biological activity

s_x	Standard Type	$<\mu^i_5(s)>$	n_1	n_{tot}	s_x	Standard Type	$<\mu^i_5(s)>$	n_1	n_{tot}
1	IC_{50}	960.45	3641	6070	23	Ratio EC_{50}	494.89	8	23
2	Inhibition	810.03	809	1997	24	TC_{50}	1203.74	7	15
3	Activity	892.48	721	1615	25	Ratio CC_{50}/IC_{50}	734.58	7	12
4	K_i	748.39	355	1045	26	NO formation	318.8	6	19
5	EC_{50}	1123.73	176	218	27	TD_{50}	619.19	6	11
6	CC_{50}	964.87	143	205	28	Ratio IC_{50}	955.38	5	34
7	Selectivity	761.4	83	240	29	Emax	1299.51	4	10
8	ED_{50}	1208.73	42	75	30	LD_{50}	809.01	4	5
9	ID_{50}	727.11	39	67	31	Count	390.21	4	6
10	K_d	1016.32	37	92	32	Initial rates	354.7	4	12
11	Ratio	743.76	21	92	33	SI	811.54	4	12
12	GI_{50}	607.89	19	60	34	$MNTD_{70}$	557.23	3	12
13	Efficacy	994.28	16	33	35	Specific activity	1051.26	3	6
14	K_m	759.45	15	57	36	Selectivity index	772.11	3	6
15	Selectivity ratio	869.72	12	41	37	k_{cat}	501.13	2	11
16	FC	3483.18	12	20	38	IC_{90}	1421.46	2	3
17	NOHA	271.55	12	37	39	RBA	1078.46	2	7
18	$MNTD_{90}$	524.19	10	12	40	Ratio K_i	970.76	2	3
19	Fold change	558.31	10	35	41	K_b	1214.19	1	3
20	Residual activity	1167.52	9	18	42	pIC_{50}	809.49	1	1
21	LC_{50}	841.72	8	22	43	K_{inact}	423.27	1	2
22	Survival	642.2	8	18	44	Cytotoxicity	366.49	1	3

n_1=number of active(C=1) cases for standart type ,n ($_{Total)}$= Total cases for standart type

Table 5. Results of the study of Domain of Applicability (DA) for the model

Endpoints Sub-set[a]	DA count	Total Sub-set count	DA %
Train	6481	6747	96.1
CV	2164	2253	96.0
Positive effect	4134	4309	95.9
Negative effect	4511	4691	96.2
Cytotoxicity	1174	1251	93.8
Human	3405	3506	97.1
Mouse	5234	5485	95.4
IC_{50}	4313	4494	96.0
EC_{50}	166	180	92.2
All	8645	9000	96.1

[a] Positive effect indicates that C = 1, this sub-set includes all cytotoxicity endpoints together with other biological effects.

Table 6. Effect on cytotoxicity for G1-treated samples at different concentrations *vs.* control groups

Group 1 Conc of G1 µg/mL	Macrophage Cytometry Parameter	Group 2							
		NC				DMSO			
		$Mean_1$	$Mean_2$	t	p	$Mean_1$	$Mean_2$	t	p
10	MFI	502.9	548.8	-0.57	0.59	502.9	480.1	0.35	0.73
	MFI SSC	416.0	432.0	-0.11	0.92	416.0	426.0	-0.10	0.92
	MFI SFC	669.2	665.5	0.09	0.93	669.2	604.8	1.44	0.18
	Anti-CD14PE	38.6	21.2	2.06	0.08	38.6	29.9	1.94	0.08
	7AAD	157.7	166.0	-0.18	0.86	157.7	174.8	-0.55	0.59
	Anti-CD14PE + 7AAD	91.5	93.6	-0.07	**0.95**	91.5	108.1	-0.99	0.35
8	MFI	516.9	548.8	-0.44	0.68	516.9	480.1	0.60	0.56
	MFI SSC	409.0	432.0	-0.17	0.87	409.0	426.0	-0.18	0.86
	MFI SFC	688.5	665.5	0.76	0.48	688.5	604.8	1.99	0.07
	Anti-CD14PE	41.2	21.2	2.65	0.04	41.2	29.9	2.82	0.02
	7AAD	153.2	166.0	-0.33	0.75	153.2	174.8	-0.78	0.45
	Anti-CD14PE + 7AAD	89.5	93.6	-0.17	0.87	89.5	108.1	-1.31	0.22
6	MFI	537.7	548.8	-0.14	0.89	537.7	480.1	0.89	0.39
	MFI SSC	444.0	432.0	0.08	0.94	444.0	426.0	0.17	0.86
	MFI SFC	705.3	665.5	0.84	0.43	705.3	604.8	2.16	0.06
	Anti-CD14PE	38.9	21.2	2.27	0.06	38.9	29.9	2.16	0.06
	7AAD	175.2	166.0	0.19	0.85	175.2	174.8	0.01	0.99
	Anti-CD14PE + 7AAD	116.2	93.6	0.85	0.43	116.2	108.1	0.53	0.61
4	MFI	498.8	548.8	-0.69	0.52	498.8	480.1	0.30	0.77
	MFI SSC	474.2	432.0	0.30	0.78	474.2	426.0	0.49	0.63
	MFI SFC	594.5	665.5	-0.56	0.60	594.5	604.8	-0.13	0.90
	Anti-CD14PE	42.8	21.2	2.40	0.05	42.8	29.9	2.67	0.02
	7AAD	214.5	166.0	0.89	0.41	214.5	174.8	1.14	0.28
	Anti-CD14PE + 7AAD	138.7	93.6	1.25	0.26	138.7	108.1	1.50	0.17
2	MFI	497.2	548.8	-0.69	0.52	497.2	480.1	0.27	0.79
	MFI SSC	398.5	432.0	-0.25	0.81	398.5	426.0	-0.29	0.78
	MFI SFC	668.2	665.5	0.08	0.94	668.2	604.8	1.48	0.17
	Anti-CD14PE	40.7	21.2	2.41	0.05	40.7	29.9	2.50	0.03
	7AAD	199.7	166.0	0.80	0.45	199.7	174.8	0.85	0.41
	Anti-CD14PE + 7AAD	124.8	93.6	1.18	0.28	124.8	108.1	1.09	0.30

$Mean_1$=mean group 1; mean $_2$= mean group2

Table 7. Results of Dosis *vs.* Effect EC_{50} curve fitting by different algorithms

Curve Fitting	R2	RMSE	a	b	c	d	e
5PL ($1/Y^2$) *	0.9588	0.0000	47.85819	-7.1792.	21.8259	130.5090	0.5185
5PL (1/Y)	0.1833	5.8115	47.7578	-12.029	18.2210	-76.5320	0.3032
4PL (1/Y)	0.9583	1.2316	-214.1114	-3.7650	25.8656	47.8645	
Log-Log	0.6222	2.3871	0.0948	1.1722			
Quadratic (1/Y)	0.9526	0.9381	-0.1505	0.9424	46.5202		
Linear (1/Y)	0.8154	1.4961	-0.8616	50.7310			

* Best fit model, 5PL is Five Parameters Logistic, 4PL is Five Parameters Logistic

Table 8. Theoretic-experimental determination of some endpoints for G1 cytotoxicity in multiplexing assays

C	p(1)	Type	Rel	Cutoff	Units	Assay Description b
Endpoints for G1 cytotoxicity experimentally determined in this work						
0	0	EC_{50}	<	21.58/10	µg	7ADD mouse peritoneal macrophages after 24h
0	0	Cytotoxicity	>	23.6 / 50	%	7ADD mouse peritoneal macrophages after 24 hrs at 10
0	0	Cytotoxicity	>	19.4 / 50	%	7ADD mouse peritoneal macrophages after 24 hrs at 8 µg/mL
0	0	Cytotoxicity	>	16.5 / 50	%	7ADD mouse peritoneal macrophages after 24 hrs at 6 µg/mL
0	0	Cytotoxicity	>	9.6 / 50	%	7ADD mouse peritoneal macrophages after 24 hrs at 4 µg/mL
0	0	Cytotoxicity	>	9.9 / 50	%	7ADD mouse peritoneal macrophages after 24 hrs at 2 µg/mL
0	0	MFI	<	p < 0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 10
0	0	MFI	<	p < 0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 8 µg/mL
0	0	MFI	<	p < 0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 6 µg/mL
0	0	MFI	<	p < 0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 4 µg/mL
0	0	MFI	<	p < 0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 2 µg/mL
Predicted multiplexing endpoints for G1 cytotoxicity						
0	**0.28**	TC_{50}	<	100	µM	Cytotoxicity against human macrophages
1	**0.84**	ED_{50}	<	11.4	µM	CAM WEHI cell line by MTT assay
1	**0.63**	IC_{50}	<	33.4	µM	Cytotoxicity against WEHI cell lines.
1	**0.72**	EC_{50}	<	164.7	uM	CAM RAW264.7 cells (MMLC) assessed as cell survival
1	**0.69**	EC_{50}	<	23.83	ug ml-1	CAM RAW264.7 cells (MMLC) by MTT colorimetric assay
1	**0.89**	IC_{90}	<	13.1	µM	Cytotoxicity against rat RAW264.7 cells by MTT assay
1	**0.78**	EC_{50}	<	10.2	µM	In vitro cytotoxicity against J774.2 cells after 72 h incubation.
1	**0.60**	IC_{50}	<	26.7	µM	CAM J774 cells expressing RANKL signaling by Alamar
1	**0.63**	IC_{50}	<	143.8	µM	Cytotoxicity against J774.1 cell line after 48 hrs
1	**0.74**	EC_{50}	<	9.19	uM	Cytotoxicity against macrophage cell line (J774)
0	0.33	CC_{50}	<	133.23	µM	CAM RAW264.7 cells after 48 hrs by MTT assay
0	0.37	CC_{50}	<	28.87	uM	CAM J774A1 cells assessed as cell viability after 72 hrs by
0	0.37	CC_{50}	<	50	ug.ml-1	Cytotoxicity against human J774A1
0	0.40	CC_{50}	<	79.03	ug.ml-1	CAM J774A1 cells after 48 hrs by Geimsa staining method
0	0.44	CC_{50}	<	76.79	ug ml-1	CAM J774A1 cells after 24 to 72 hrs by MTT assay
0	0.44	CC_{50}	<	57.64	ug ml-1	CAM J774A1 cells after 72 hrs by MTT assay
0	0.45	CC_{50}	<	42.49	ug ml-1	CAM J774A1 cells
0	0.46	CC_{50}	<	43.1	ug ml-1	CAM J774A1 cells by MTT assay
0	0.39	IC_{50}	<	90.93	uM	Cytotoxicity against LPS-stimulated mouse RAW264.7 cells
0	0.40	IC_{50}	<	68.76	ug ml-1	CAM J774A1 by MTT assay
0	0.42	IC_{50}	<	13.34	ug ml-1	CAM J774A1 cells by rapid colorimetric assay
0	0.42	IC_{50}	<	236.7	uM	Cytotoxicity against LPS-stimulated mouse RAW264.7 cells
0	0.42	IC_{50}	<	30.94	uM	Cytotoxicity in mouse RAW264.7 cells
0	0.43	IC_{50}	<	6.76	ug ml-1	*In vitro* cytotoxicity of compound against EL4. mouse
0	0.43	IC_{50}	<	73.35	uM	CAM RAW264.7 cells after 24 hrs by MTT assay
0	0.43	IC_{50}	<	231.62	uM	CAM J774 cells after 24 hrs by by resazurin reduction test
0	0.44	IC_{50}	<	23.8	ug ml-1	CAM J774A1 cells after 72 hrs by cell-titer assay
0	0.44	IC_{50}	<	48.22	uM	CAM RAW264.7 cells assessed as cell viability after 24 hrs
0	0.45	IC_{50}	<	103.9	uM	Cytotoxicity against LPS-stimulated mouse RAW264.7 cells
0	0.45	IC_{50}	<	10	uM	CAM RAW264.7 cells assessed as cell viability after 24 hrs
0	0.45	IC_{50}	<	6.3	ug ml-1	Cytotoxicity against human EL4 cells
0	0.46	IC_{50}	<	57	uM	CAM RAW264.7 cells assessed as reduction in cell viability
0	0.46	IC_{50}	<	175	ug ml-1	CAM J774A1 cells after 24 hrs by trypan blue exclusion assay
0	0.46	IC_{50}	<	42.72	uM	CAM RAW264.7 cells by MTT assay
0	0.47	IC_{50}	<	207.5	uM	CAM macrophage RAW264.7 cells after 48 hrs by MTT
0	0.49	IC_{50}	<	12.17	ug ml-1	CAM RAW264.7 cells after 2 days by MTT assay
0	0.50	IC_{50}	<	27.5	uM	CAM RAW264.7 cells after 72 hrs by resazurin assay

0	0.25	Inhibition	>	17.3	%	CAM J774A1 cells assessed as reduction in metabolic activity
0	0.37	LC_{50}	<	12.28	ug ml-1	CAM RAW264.7 cells
0	0.09	$MNTD_{70}$	<	88.51	uM	CAM RAW264.7 cells assessed as maximum non-toxic dose
0	0.12	Survival	>	68.14	%	CAM RAW264.7 cells assessed as cell survival rate at 10 uM
0	0.12	Survival	>	94	%	CAM RAW264.7 cells at 21 uM
0	0.33	Activity	>	100	%	CAM J774 macrophage assessed as cell viability at 1 ug/mL
0	0.35	Activity	>	65.3	%	CAM J774A1 cells assessed as macrophage number at 40 to
0	0.35	Activity	>	64.3	%	CAM J774A1 cells assessed as macrophage number at 2.3
0	0.35	Activity	>	23.27	%	CAM J744A.1 cells at 0.5 ug/mL by MTT assay
0	0.35	Activity	>	99.5	%	CAM J774A1 cells assessed as macrophage number at 4
0	0.35	Activity	>	5	%	CAM J774A1 cells assessed as dead cells at 10 uM by Sytox
0	0.35	Activity	>	94	%	CAM J774A1 cells assessed as live viable cells at 10 uM by
0	0.35	Activity	>	93	%	CAM J774A1 cells assessed as live cells at 50 uM by calcein
0	0.35	Activity	>	58.9	%	CAM J774A1 cells assessed as live cells at 100 uM by calcein
0	0.35	Activity	>	94.4	%	CAM J774A1 cells assessed as macrophage number at 10.5
0	0.35	Activity	>	73.4	%	CAM J774A1 cells assessed as macrophage number at 3
0	0.35	Activity	>	16	%	CAM J744A.1 cell assessed as survival rate at 10 uM
0	0.35	Activity	>	12.3	%	CAM J774A1 cells assessed as macrophage number at 3.2
0	0.35	Activity	>	55.9	%	CAM J774A1 cells assessed as macrophage number at 1.7
0	0.35	Activity	>	50	%	CAM J774 macrophage at 40 uM
0	0.36	Activity	>	100	%	Cytotoxicity against murine J774 cells (MC) assessed as cell
0	0.36	Activity	>	100	%	Cytotoxicity against murine J774 cells assessed as cell
0	0.36	Activity	>	100	%	Cytotoxicity against murine J774 cells assessed as cell
0	0.39	Activity	>	11.35	%	CAM J744A.1 cells at 0.005 ug/mL by MTT assay
0	0.39	Activity	>	12.35	%	CAM J744A.1 cells at 0.05 ug/mL by MTT assay
0	0.41	Activity	>	70.83	%	CAM J744A.1 cell assessed as survival rate
0	0.43	Activity	>	0.95	%	CAM J774 cells at 100 uM relative to 5-
0	0.43	Activity	>	42.21	%	Unspecific cytotoxicity against murine J774 macrophages at
0	0.43	Activity	>	41.5	%	CAM J774 cells assessed as cell viability at 100 ug/ml by
0	0.44	Activity	>	50	%	CAM J774 macrophage at 400 uM
0	0.44	Activity	>	88.2	%	CAM J774 cells infected with Mycobacterium bovis BCG
0	0.44	Activity	>	0.76	%	CAM J774 macrophages at 2.1 uM after 24 hrs by resazurin
0	0.44	Activity	>	4.89	%	CAM J774 macrophages at 8.6 uM after 24 hrs by resazurin
0	0.44	Activity	>	18.88	%	CAM J774 macrophages at 21.7 uM after 24 hrs by resazurin
0	0.44	Activity	>	51.8	%	CAM J774 cells at 400 uM after 48 hrs by MTT assay
0	0.44	Activity	>	2.11	%	CAM J774 macrophages at 4.3 uM after 24 hrs by resazurin
0	0.45	Activity	>	51.17	%	CAM J774 cells assessed as cell viability at 100 ug/mL after
0	0.45	Activity	>	95.34	%	CAM RAW264.7 cells assessed as cell viability at 1 uM after
0	0.45	Activity	>	48.86	%	CAM J774 cells assessed as cell viability at 10 ug/ml by MTT
0	0.46	Activity	>	95.42	%	CAM RAW264.7 cells assessed as cell viability at 100 uM
0	0.46	Activity	>	75.95	%	Cytotoxicity against Mycobacterium bovis Bacillus Calmette-
0	0.47	Activity	>	81.68	%	Cytotoxicity against Mycobacterium bovis Bacillus Calmette-
0	0.47	Activity	>	89	%	Cytotoxicity against Mycobacterium bovis Bacillus Calmette-
0	0.47	Activity	>	93.57	%	CAM J774 cells assessed as cell viability at 1 ug/ml by MTT
0	0.33	IC_{50}	<	6	uM	Cytotoxicity against EL-4 cell line (mouse thymoma cells)
0	0.39	IC_{50}	<	21.41	uM	Inhibition of Bacillus anthracis lethal toxin-induced
1	0.54	IC_{50}	<	221.07	uM	CAM J774 cells after 48 hrs by MTT assay
1	0.54	IC_{50}	<	2256.3	uM	CAM J774 cells after 24 hrs by resazurin assay
1	0.54	IC_{50}	<	565.75	ug ml-1	CAM J774 cells after 24 hrs by MTT assay
1	0.55	IC_{50}	<	508.6	uM	CAM RAW264.7 cells assessed as cell viability after 4 hrs by
1	0.56	IC_{50}	<	25.51	uM	In vitro CAM J774 macrophages.
1	0.56	IC_{50}	<	736.2	uM	CAM J774 cells assessed as cell viability after 48 hrs by

1	0.57	IC_{50}	<	72.79	uM	CAM J774 macrophages after 48 hrs by MTT assay
1	0.58	IC_{50}	<	0.24	uM	Inhibitory concentration required for cytotoxicity in J774.2
1	0.58	IC_{50}	<	416.67	ug ml-1	CAM J774 cells after 24 hrs

[a] Cutoff used was the threshold value recommended by REACH for this assay (in experimental outcomes) or the average value for all compounds in ChEMBL for this assay (in predicted outcomes). The J774 cell lines are Macrophage Cells (MC) and RAW264.7 is a murine macrophage-like cells (MMLC). CAM is Cytotoxicity Against Macrophage.

Table 9. Theoretic prediction of some endpoints for G1 interaction with human protein targets

Target name	ID	p(1)	Res.	Lev.	Type	Rel	Cutoff	Units
C-C chemokine receptor type 3	3473	0.78	-0.63	0.002	E_{max}	>	56.25	%
C-C chemokine receptor type 3	3473	0.63	-0.53	0.002	K_b	>	0.01	uM
C-C chemokine receptor type 1	2413	0.71	-0.58	0.001	ED_{50}	<	0.00.1	uM
Matrix metalloproteinase 12	4393	0.66	-0.55	0.001	IC_{50}	<	5.8	uM
Interleukin-8	2157	0.64	-0.54	0.000	IC_{50}	<	5.3	uM
MSR I and II	5811	0.63	-0.53	0.001	IC_{50}	<	38.3	μM
Acyl-CoA: cholesterol Acyltransferase	2265	0.64	-0.54	0.001	IC_{50}	<	21.3	μM
MSPR	2689	0.62	-0.52	0.001	Activity	>	78.67	%
MSPR	2689	0.60	-0.52	0.000	K_d	>	9.74	uM
MCSFreceptor	1844	0.62	-0.52	0.001	Activity	>	100.33	%
MCSFreceptor	1844	0.61	-0.52	0.000	IC_{50}	<	1.38	μM
MMIF	2085	0.70	-0.57	0.001	IC_{50}	<	65.1	μM
MMIF	2085	0.60	-0.51	0.001	Activity	>	36	%

[a] MMIF is Macrophage Migration Inhibitory Factor, MCSF is Macrophage Colony Stimulating Factor, MSPR is Macrophage-Stimulating Protein Receptor, MSR is Macrophage Scavenger Receptor I and II

FIGURES TO BE INSERTED IN THE TEXT

Figure 1. Analysis of the Domain of Application of the model

Figure 2. Dose-Response of cytotoxicity in Balb/C mouse peritoneal macrophages marked with CD14Pe/7AAD exposed to different concentrations of G1

Figure 3. Pseudo-color smooth projection of MFI values over FSC *vs.* SSC plot after administration of G1 at cmax. Total cell populations (A) or Macrophages population (C.B.D.E.F)

Figure 4. Effect of G1 on Balb/C mouse macrophages culture primary and were exposed to different concentrations ((10, 8, 6, 4, and 2 µg/mL) for a period of 24 hours The results are expressed as Mean Intensity Fluorescence (MFI) of control values N = 6 animals per group and 1.106 Cell; NS= Not Statistically significant differences $\boldsymbol{p} \leq 0.05$ for the same groups. Dark Green=macrophages labeled with CD14PE and stained with 7AAD (Dead), Dark Purple = macrophages labeled CD14 PE (live)

Figure 5. Results of flow cytometry for Balb/C mouse peritoneal macrophages exposed to G1 at 10 µg/mL

Figure 6. Masterplex interface illustrating MFI *vs.* conc. effect of G1 in Balb/c mouse peritoneal macrophages **(A)** and Sigmoidal curve representative of the 5PL model **(B)**.

2.2. Biology assays

2.2.1. Reagents and antibody

1-5-Bromofur-2-il-2-bromo-2-nitroethene (G1); CAS number 35950-55-1, was kindly supplied from the CBQ, Sample purity was 99.93%. G1 was dissolved in dimethylsulfoxide (DMSO), which was purchased in turn from Sigma–Aldrich Co. (DF, México). Macrophages were stained with phycoerythrin (PE), labelled monoclonal antibodies according to the manufacturers' instructions. Flow cytometry was performed using a FACalibur cytometer (Becton Dickinson, México). Thereafter, FACS data were analyzed with FlowJo 7.6.5 software. Both, anti-CD14 antibody (used to label CD14 receptor) and 7 –aminoactinomycin (7-AAD) at 5 μg/mL viability solution were purchased from BD (BD Biosciences, México).

2.2.2. Animals.

Female Balb/C mice weighing 18–20 g were purchased from the UNAM-Harlan laboratories (DF, México). All animals (n=6) were allowed to acclimate to our laboratory facilities for at least 7 days before their inclusion in an experiment. They were housed in standard laboratory conditions (22 3 °C; relative humidity 50–55%; 12h light/dark cycle) and given *ad libitum* access to food and water. This work agreed with Ethical Principles in Animal Research adopted by México [40].

2.2.3. Peritoneal macrophages isolation and cell culture.

Peritoneal macrophages were obtained from mice euthanized by cervical dislocation. The peritoneal of the animals were surgically exposed using a midline incision. Peritoneal fluid was harvested by injecting 10 mL of ice-cold PBS into the peritoneal cavity followed by syringe aspiration. Cell suspensions were washed twice by centrifugation. Cell viability (over 95%) was determined using trypan blue exclusion. Macrophage numbers were adjusted to 1 × 106 cell/mL and plated 100 μL/ well in 96-well flat-bottomed tissue culture plates (UNIPARTS, Toluca, México). Cells were incubated in RPMI 1640 complete medium containing 10% FBS, and incubated for 24 h at 37 °C under 5% CO2 in a humidified chamber. Non-adherent cells were removed by gently washing with PBS and fresh RPMI 1640 complete medium was replaced. The efficiency of macrophage enrichment was monitored by 7AAD assay and routinely exceeded 90%. Cells were equilibrated for 24 h before commencing the experiment.

2.2.4. Determination of cytotoxicity percentage by flow cytometry analysis

In all cases, dimethyl sulfoxide (DMSO) was used as the diluting solvent, and dosage solutions were prepared immediately prior to testing. Incubations were carried out in triplicate; solvent controls were run with each experiment. The percentage of formation of cytotoxicity cells was determined by evaluating 7-Amino-actinomycin D (7AAD) stained preparations of macrophages treated with the dosed chemical (G1) at 10, 8, 6, 4 and 2 μg/mL in 24 h.

$$\text{Cytotoxicity}(\%) = 100 \cdot (\text{Ma}^{*} - 7\text{AAD}^{*})/(\text{Total event Ma}) \qquad (5)$$

Where, Ma* = Positive Macrophages labeled CD14PE, 7AAD*= Positive 7AAD (Dead macrophages), Total event Ma = Total macrophages labeled and unlabeled CD14 with CD14. Briefly, 1X 106 cells were washed twice with 1 mL ice-cold PBS. Cytotoxicity was determined using flow cytometry with a FACSCalibur cytometer (Becton Dickinson, USA) equipped with an argon-ion laser at 488 nm wavelength. Tubes 21 and 22, isotypic controls and tubes with antibodies alone were used to adjust PMT and fluorescence compensation. Fluorescence compensations were also occasionally adjusted with Compbeads (BD Biosciences) by determining the median of both positive and negative populations. Percent cytotoxicity was determined by the following formula

[41], where Ma mean macrophages count, the symbol * indicates a positive answer to CD14Pe and Negative mean negative to 7ADD staining for living cells.

Last, was fitted a response curve *vs.* concentration (MFI_i *vs.* c_i) in order to calculate the EC_{50} values using the software MasterPlex 2010, 2.0.0.73 created for the MiariBio group (*www.miraibio.com*). The MasterPlex includes Readerfit to calculate the EC_{50} and adjust the curve. ReaderFit is a free online application for adjustment of the curve that allows two fitting curves and optionally interpolates unknown values of the curve. The ReaderFit contain several equations for the model: 4 parameters logistic (4PL), 5 parameters logistic (5PL), quadratic log-logit, log-log or linear and one out four optional weighting algorithms: 1/Y, $1/Y^2$, 1/X and $1/X^2$ to minimize the error. In our case, Y variable contains the different Mean Fluorescence Intensity (MFI_i) response values and X the different concentrations (c_i) for different samples. The parameters of 5PL model are: A, B, C, D, and E. A is the MFI value for the minimum asymptote. B is the Hill slope. C is the concentration at the inflection point. D is the MFI for the maximum asymptote. E is the asymmetry factor ($E \neq 1$ for a non-symmetric curve). MFI is the. MFI values are obtained after exposition of the biological sample to one volume of 100 μL of G1 at different c_i values. This equation is represented through a sigmoid curve:

$$MFI_i = A + \frac{D}{\left(1+\left(\left(X/C\right)^B\right)\right)^E} = \left(MFI_i\right)_{\min} + \frac{\left(MFI_i\right)_{\max}}{\left(1+\left(\left(\frac{c_i}{EC_{50}}\right)^B\right)\right)^E} \quad (6)$$

$$\left(\frac{1}{MFI_i^{\,2}}\right) = A + \frac{D}{\left(1+\left(\left(X/C\right)^B\right)\right)^E} = \left(\frac{1}{MFI_i^{\,2}}\right)_{\min} + \frac{\left(\frac{1}{MFI_i^{\,2}}\right)_{\max}}{\left(1+\left(\left(\frac{c_i}{EC_{50}}\right)^B\right)\right)^E} \quad (7)$$

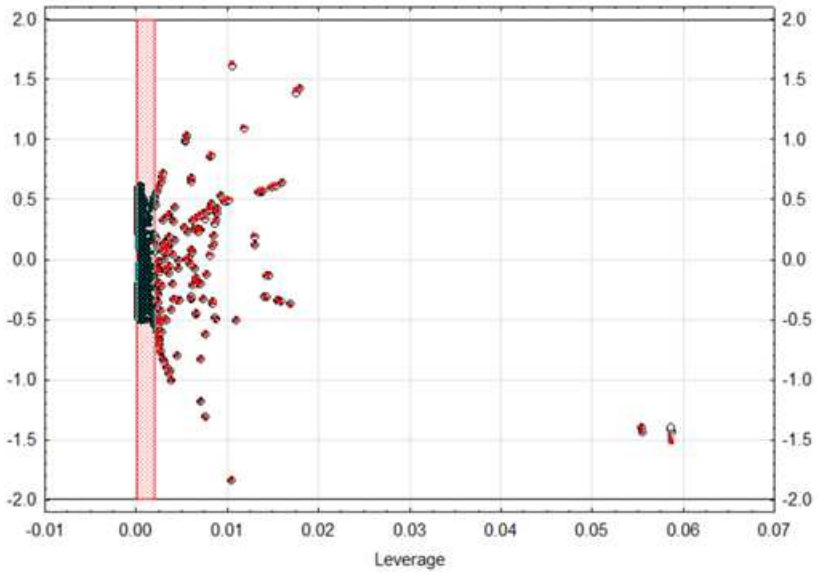
Where, MFI = Mean Fluorescence Intensity, A= is the MFI/RLU value for the minimum asymptote

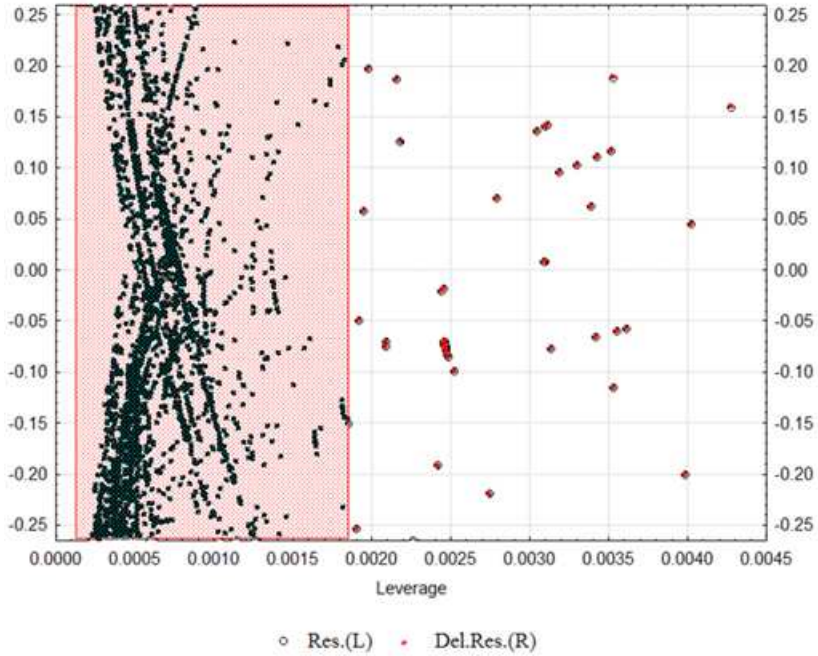
B = is the hill slope, C = EC_{50} is the concentration at the inflection point, D is the MFI/RLU value for the maximum asymptote, and E is the asymmetry factor.

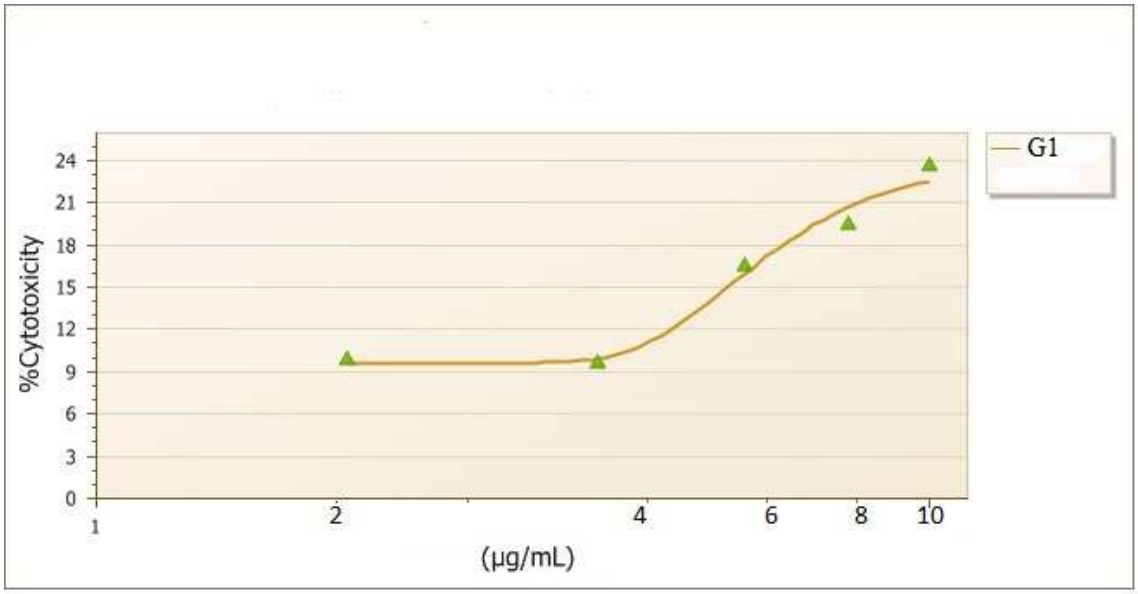
2.2.5. Statistical Analysis of experimental assays

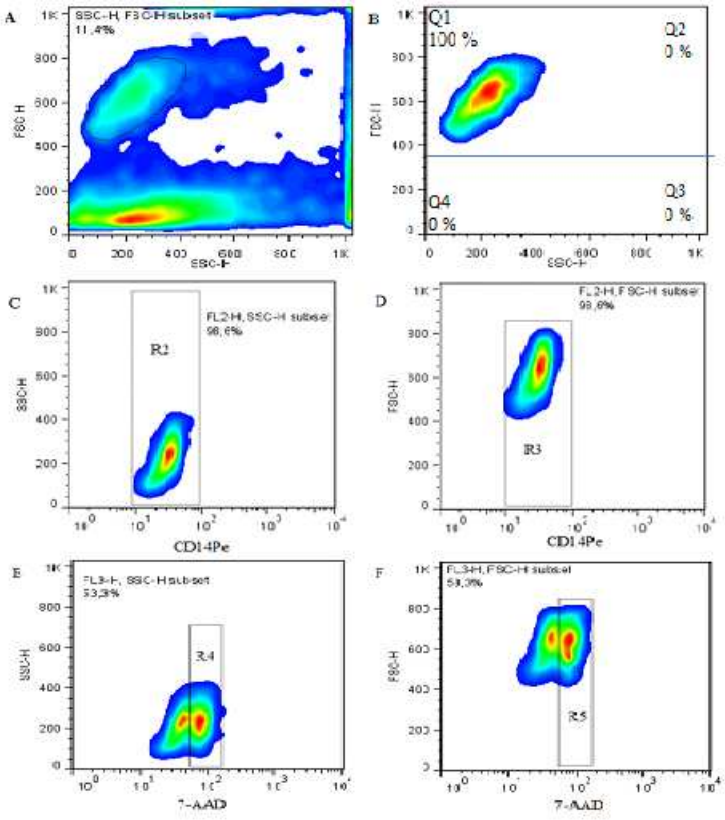
Data were analyzed using Statistica 6.0 software. Significant differences between treatments were determined by analysis of variance (ANOVA), followed by t test. Statistic significances were accepted when $P < 0.05$. The Tukey test with 95% confidence was applied to compare the means.

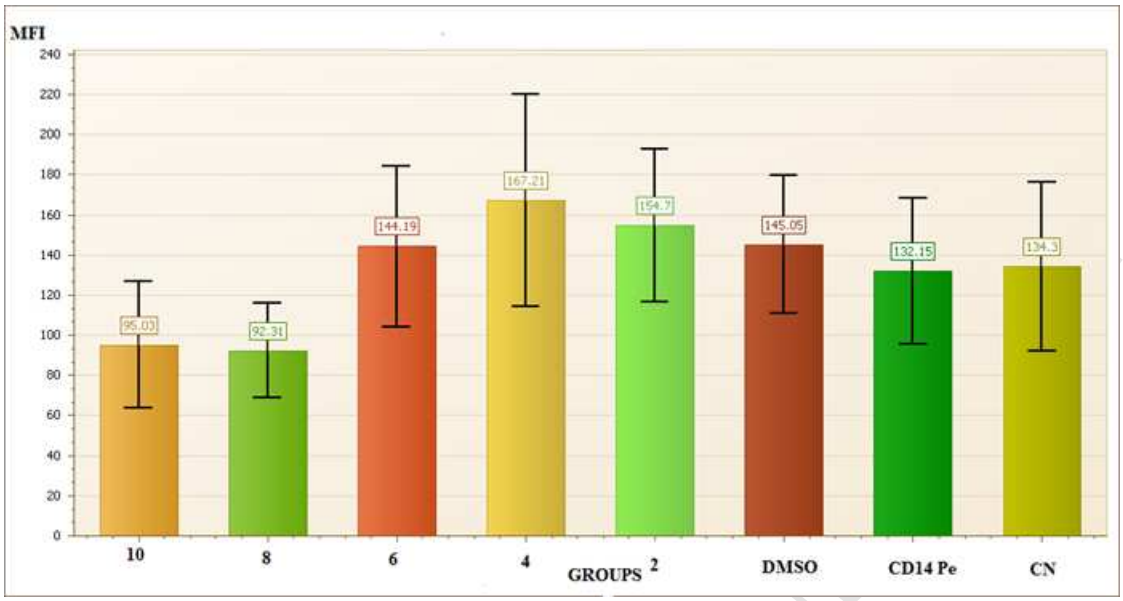
Note: the references cited here appear in the main body of the paper.

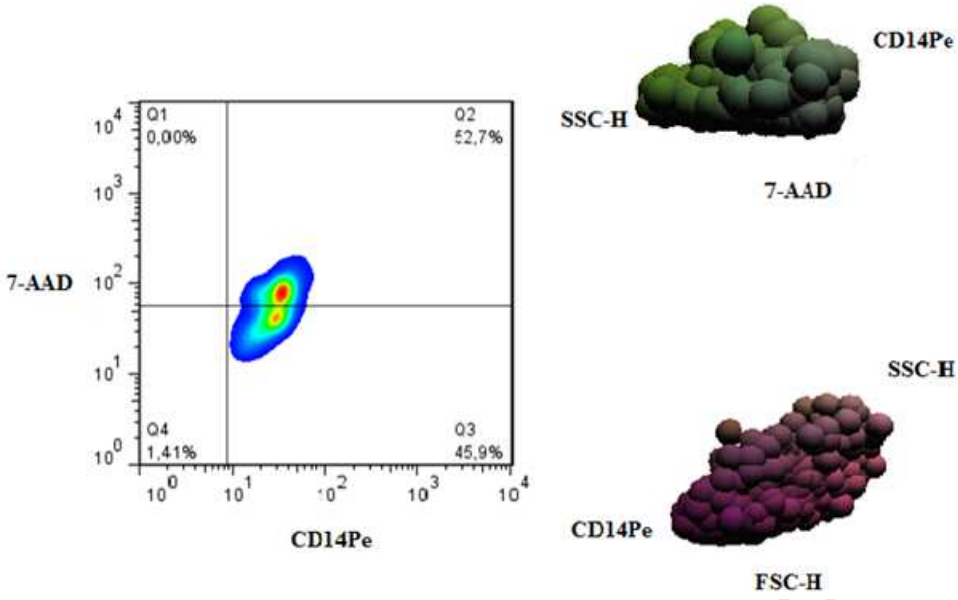

2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-0.01
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
Leverage

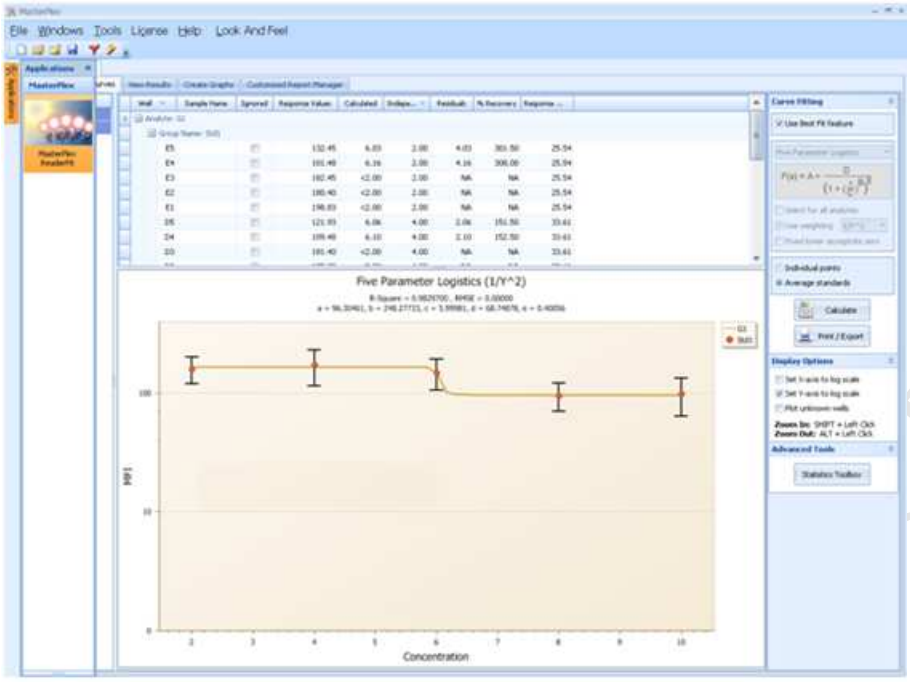

0.25
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.20
-0.25
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
0.0035
0.0040
0.0045
Leverage
Res.(L)
Del.Res.(R)



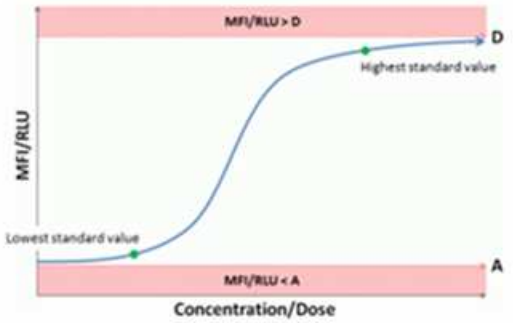







File Windows Tools License Help Look And Feel
Five Parameter Logistics (1/Y^2)
MFI
Concentration

(A)


MFI/RLU > D
D
Highest standard value
MFI/RLU
Lowest standard value
A
MFI/RLU < A
Concentration/Dose

(B)