Bioorganic & Medicinal Chemistry 20 (2012) 6181–6194

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

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

ANN multiplexing model of drugs effect on macrophages; theoretical and flow cytometry study on the cytotoxicity of the anti-microbial drug G1 in spleen

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

[a] Department of Organic Chemistry, Faculty of Pharmacy, University of Santiago de Compostela, CP15782 Santiago de Compostela, Spain
[b] Center for Research and Advanced Studies in Animal Health, Faculty of Veterinary Medicines and Animal Husbandry, Autonomous University of Mexico State, Km 15.5 Autopista de cuota Toluca-Atlacomulco, Toluca, Estado de México, México 50200, Mexico
[c] Chemical Bioactive Center (CBQ), Central University of Las Villas (UCLV), 50200 Santa Clara, Cuba
[d] Department of Microbiology & Parasitology, Faculty of Pharmacy, University of Santiago de Compostela, CP15782 Santiago de Compostela, Spain

ARTICLE INFO

Article history:
Received 26 April 2012
Revised 11 July 2012
Accepted 13 July 2012
Available online 24 July 2012

Keywords:
CHEMBL
Multi-target versus multiplexing QSAR
Spleen macrophage
Drug cytotoxicity
Artificial Neural Networks
Flow cytometry
Multiplexing assay endpoints
Anti-parasite/anti-microbial drug

ABSTRACT

Multiplexed biological assays provide multiple measurements of cellular parameters in the same test. In this work, 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), Radial Basis Function (RBF), Probabilistic Neural Networks (PNN) and Multi-Layer Perceptrons (MLP). The best model found was the LNN, which correctly classified 8258 out of 9000 (Accuracy = 93.0%) multiplexing assay endpoints of 7903 drugs (including both training and test series). 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 1265 multiplexing assays not measured experimentally (including 152 cytotoxicity assay endpoints). Both experimental and theoretical results point out a low macrophage cytotoxicity of G1. This work breaks new ground for the 'in silico' multiplexing screening of large libraries of compounds. The results obtained are very significant because they complement the immunotoxicology studies of this important anti-microbial/anti-parasite drug.

1. Introduction

The large number of experimental results reported by different groups worldwide has led to the accumulation of huge amounts of information in large databases. Consequently, this determines the need of Machine Learning algorithms to perform data mining of these databases. Artificial Neural Networks (ANNs) are one of the most powerful algorithms that can be used to find predictive models in biosciences. One of the most important classes of theoretical models that can be found using ANNs are the so-called Quantitative Structure–Activity Relationships (QSARs). These ANN-based QSARs can be used to predict multiplexing assay endpoints (results from multiple assays).[1–5] With special emphasis on predictions of large compound libraries against several molecular or cellular targets, one of the most outstanding databases with respect to drug cytotoxicity/biological effects on macrophage cells is CHEMBL. CHEMBL is an Open Data database containing Binding (B), Functional (F), and ADMET (A) information for a large number of drug-like bioactive compounds. These data are manually abstracted from the primary published literature on a regular basis, then further curated and standardized to maximize their quality and utility across a wide range of chemical biology and drug-discovery research problems. Currently, the database contains 5.4 million bioactivity measurements for more than 1 million compounds and 5200 protein targets. Access is available through a web-based interface, data downloads and web services 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). Consequently, the search for computational models to predict the possible results for new drugs in all these assays has become a goal of major importance to reduce experimentation costs. In addition, despite the large number

* Corresponding author.
E-mail address: gonzalezdiazh@gmail.com (H. González-Díaz).

0968-0896/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2012.07.020

of assays described, many drugs have been assayed only for some selected tests. Consequently, predictive models may also become an important tool to carry out an 'in silico' mining of CHEMBL predicting new results for drugs already released.

Mok et al. has recognized that the mining of CHEMBL using different computational tools is a very interesting source of new knowledge.[7] Specifically, Quantitative Structure-Toxicity Relationships (QSTR), or the very QSAR in toxicology, have been widely used to predict toxicity from chemical structure and corresponding physicochemical properties.[8] Unfortunately, almost all 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 different chemoinformatics tools that can be used to calculate molecular descriptors in QSAR. Furthermore, we can select different Machine Learning (ML) algorithms to seek the QSAR model using the molecular descriptors as input. In particular, the chemoinformatics method TOPS-MODE implemented by Estrada et al. [refs] in the Modeslab software has been demonstrated to be successful in both QSAR and mt-QSAR studies.

The cytotoxicity of drugs on the immune system in general (immunotoxicity) and on macrophages in particular is one of the most important properties in CHEMBL. Macrophages play a central role in the immune response by presenting antigens to lymphocytes during the development of specific immunity and by serving as supportive accessory cells to lymphocytes, in which they release soluble factors.[11] Macrophages are the heterogeneous grouping of cells that are derived from monocytes. They have a multitude of functions depending on their final differentiated state. A great deal of test development for toxicologist screening proves that the cytotoxicity test is a screening method that typically uses permanent cell lines for ranking acute toxicities of parent compounds based on the basal cytotoxicity theory chemicals which exert their acute toxic effects by interfering with basic cellular functions that are common to all mammalian cells.[12] In vitro drug cytotoxicity may be variable among different cell lines and a parameter for cell death is the integrity of the cell membrane, which can be measured by the cytoplasmic enzyme activity released by damaged cells.[13] A commonly used source of mouse and rat macrophages is the spleen. In the spleen, the majority of macrophages are located in the Cords of Billroth in the red pulp (red pulp macrophages). They show a high acid phosphatase activity, which can be detected with various pan-macrophage markers such as: F4/80,[14] Mac-I,[15] MOMA-2[16,17] and CD14[18] which should be considered as scavenging cells involved in removing foreign particles from the blood stream. The spleen has at least four different macrophage populations with different phenotypes and different functions.[19] Local proliferation of macrophage precursors seems to be particularly important for white pulp and metallophilic macrophages; however, monocyte influx may account for the other macrophage populations in the spleen.[20]

However, drug macrophage cytotoxicity is one of the properties reported in CHEMBL to be less studied in computational terms. To the best of our knowledge, there are no mt-QSAR models based on TOPS-MODE and ANNs useful to predict multiplexing assay endpoints for drug effects on macrophages. The main objective of the present work is to develop a valid mt-QSAR model combining both ANNs and TOPS-MODE to predict the biological effect of drugs on macrophages in a large set of m_j assay conditions based on CHEMBL data (see Fig. 1). Another important goal is to illustrate the use of the new method in a real-life example. To this end, we have downloaded and calculated TOPS-MODE selected descriptors for the large dataset reported in CHEMBL. Next, we have trained and tested a new ANN model using the STATISTICA software. After

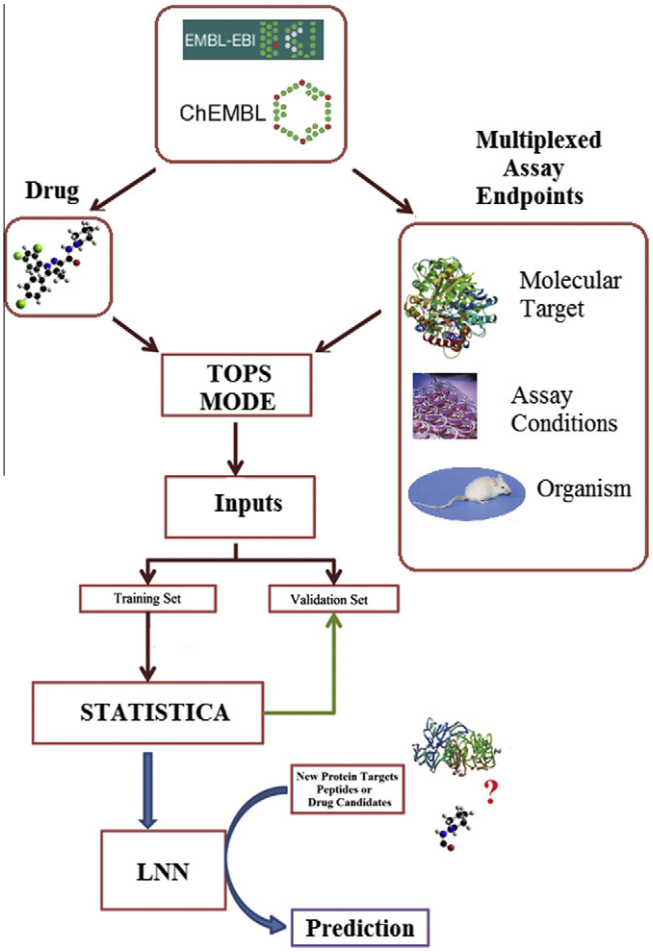

Figure 1. Workflow of the mt-QSAR study.

that, we have been able to report, for the first time, the experimental study of the effect of the drug G1 on Balb/C mouse spleen macrophage population using a flow cytometer. Last, we have used the best ANN model found to predict several assay endpoints for G1, not experimentally determined in this work.

2. Results and discussion

2.1. LNN multiplexing model of drug effect on macrophages

It is well-known that biological outcomes in multiplex cell viability assay for drugs effect on different cellular lineages depend not only on drug structure but also on the set of selected assay conditions (m_j).[21] In this work we have used different ANNs to seek different mt-QSAR models that are able to assign each drug into one out of two possible activity classes given the molecular structure and several multiplex assay conditions m_j. The activity classes are positive outcome/active compound ($C = 1$) or negative outcome/non-active compounds ($C = 0$). These models are expected to give different classification probabilities of the compound for different: organisms (o_t), biological assays (a_u), 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 account the different degrees of accuracy or level of curation (c_l) in the experimental data. We have fit the classifier using different ANNs. The best ANN model found presented the LNN topology; the equation of this model is:

Table 1
Overall results of the ANN classification models

Profile	Sub-set	Train			Statistics[a]	Test		
		Total	Correct	(%)		(%)	Correct	Total
LNN 4:4-1:1	Positive	3219	3022	93.9	Sn	93.1	1015	1090
	Negative	3528	3262	92.5	Sp	92.2	1072	1163
	Overall	6747	6284	93.1	Ac	92.6	2087	2253
MLP 4:4-7-1:1	Positive	3219	3037	94.3	Sn	94.1	1026	1090
	Negative	3528	3273	92.8	Sp	93.2	1084	1163
	Overall	6747	6310	93.5	Ac	93.7	2110	2253
MLP 4:4-10-9-1:1	Positive	3219	3037	94.3	Sn	94.4	1029	1090
	Negative	3528	3273	92.8	Sp	93.5	1087	1163
	Overall	6747	6310	93.5	Ac	93.9	2116	2253
RBF 4:4-1-1:1	Positive	3219	1730	53.7	Sn	51.5	561	1090
	Negative	3528	1842	52.2	Sp	51.9	604	1163
	Overall	6747	3572	52.9	Ac	51.7	1165	2253
PNN 4:4-6747-2-2:1	Positive	3219	191	5.9	Sn	6.5	71	1090
	Negative	3528	3528	100	Sp	100	1163	1163
	Overall	6747	3719	55.1	Ac	54.8	1234	2253

[a] Sensitivity = Sn = Positive Correct/Positive Total; Specificity = Sp = Negative Correct/Negative Total; Accuracy = Ac = Total Correct/Overall Total.

$$S_i(m_j) = -4.361 \cdot v_5 + 0.507 \cdot v_2 - 1.413 \cdot v_4 + 4.203 \cdot v_3 - 4.0372$$
$$N = 6747 \quad S_p = 92.2 \quad S_n = 93.1 \quad A_c = 92.6 \quad A_{roc} = 0.99 \tag{1}$$

$S(m_j) = S(d_i, a_u, c_l, o_t, t_e, s_x)$ 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 selected conditions m_j. We should take into consideration that this 'mt-QSAR' model applies not only to different targets but also to different multiplexing conditions. Consequently, we decide to classify it not as a multiplexing QSAR (mx-QSAR) model and not as one mt-QSAR model. The statistical parameters for the above equation in training are: Number of cases (N), Specificity (Sp), Sensitivity (Sn), Accuracy (Ac), and Area under ROC curve (A_{roc}).[22] The final cut-off for this LNN model is $S_i(m_j)_{cutoff}$ <0.567 = >C_{ij} = 1. It means that the i^{th} drugs with $S_i(m_j)_{cutoff}$ <0.567 are predicted by the model to give a positive outcome in the j^{th} assays carried out under the given set of m_j conditions. This linear equation has presented good results both in training and external validation series, with an overall Accuracy higher than 93% (see Table 1). Please, see detailed results for all cases in training and test set in the Supplementary data 2 (SD2). According to previous reports, Accuracy values higher than 75% are acceptable for QSAR models that derive from linear techniques like LDA and LNN.[23–31]

The reader should be aware that N is not a number of compounds here, but a 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.[32] The above equation was written in a compact form using the components of the Book Code vector $BC_j = v_j$. These v_j are transformed variables obtained after pre-processing the original variables (ov_j) calculated with TOPS-MODE and introduced in STATISTICA. In the LNN algorithm the STATISTICA software used the following transformation v_j = Scale $factor_j$*(ov_j) + Shift $factor_j$. Next, we back-project the above equation over the original space of variables (ov_j) in order to obtain the model in terms of these ov_j introduced in STATISTICA.

$$S_i(m_j) = -4.361 \cdot \left[\frac{{}^{*}\mu_5^i + 0.023}{16.92 \cdot 10^{-6}}\right] + 0.507 \cdot \left[\frac{\Delta\mu_5^i(o) - 0.112}{6.540 \cdot 10^{-6}}\right] - 1.413 \cdot \left[\frac{\Delta\mu_5^i(t) - 0.181}{6.110 \cdot 10^{-6}}\right] + 4.203 \cdot \left[\frac{\Delta\mu_5^i(s) - 0.120}{6.668 \cdot 10^{-6}}\right] - 4.0372 \tag{2}$$

In Table 2 we depict the values of Shift $factor_j$ and Scaling $factor_j$ that are necessary to make this transformation. After making the necessary arithmetic operations, we can eventually obtain the final model in terms of the initial variables ov_j. The output of this model is ${}^{*}S_i(m_j) = S_i(m_j)\cdot 10^{-5} = S(d_i, a_u, c_l, o_t, t_e, s_x)\cdot 10^{-5}$ is a real-valued variable scaled by a factor in order to obtain more manageable values. Both ${}^{*}S_i(m_j)$ and $S_i(m_j)$ score the propensity of the drug to be active in multiplex pharmacological assays of the drug depending on the selected conditions m_j.

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

In Table 2 we also report the results of the sensitivity analysis for the variables present in our LNN model (Ratio values). We can conclude that there is no dominant effect (v_5 = drug structure ratio ≈ v_4 = target ratio ≈ v_2 = organ ratio ≈ v_3 = ratio of the standard type of activity measured ≈ 1). This means that all the chemico-biological effects considered have a similar influence on the final predicted value (all ratios are approximately equal to 1). As follows, we expand the terms of the model in order to facilitate their interpretation in biological terms.

$${}^{*}S_i(m_j) = -2.57 \cdot p(a_u) \cdot p(c_l) \cdot {}^{std}\mu_5^i + 0.775 \cdot ({}^{std}\mu_5^i - \langle {}^{std}\mu_5^i(o_t)\rangle) - 2.31 \cdot ({}^{std}\mu_5^i - \langle {}^{std}\mu_5^i(t_e)\rangle) + 6.30 \cdot ({}^{std}\mu_5^i - \langle {}^{std}\mu_5^i(s_x)\rangle) + 0.4300523 \tag{4}$$

The first parameter ${}^{*}\mu^i{}_5 = p(a)\cdot p(c)\cdot {}^{std}\mu^i{}_5$ codifies the influence of the chemical structure of the compound on the biological activity. It is known that the spectral moment of order 5 codifies information about all types of structural fragments with five or less bonds in the molecule. In addition to the topological information, ${}^{w}\mu^i{}_5$ also codifies 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.[33–37] Consequently,

Table 2
Standard numeric pre-processing, model weights and sensitivity analysis results

ANN	Parameters [a]	Original variables (ov_j)			
		$\Delta\mu^i{}_5(o)$	$\Delta\mu^i{}_5(s)$	$\Delta\mu^i{}_5(t)$	$*\mu^i{}_5$
LNN 4:4-1:1	Model Coefficient	0.775295207	6.302952175	−2.313935024	−2.576654593
	Scale factor	0.000006540	0.000006668	0.000006111	0.000016925
	Shift factor	0.111678663	0.120268995	0.181726246	−0.002313763
	Book codes (v_j)	v_2	v_3	v_4	v_5
	Weights of v_j	0.50706528	4.202934792	−1.413934566	−4.361036926
	Ratio of v_j	1	1.5	1.1	1.6
MLP 4:4-7-1:1	Ratio of v_j	3.3	12.6	1.7	14.7
MLP 4:4-10-9-1:1	Ratio of v_j	5.4	13.5	3.1	21.4
RBF 4:4-1-1:1	Ratio of v_j	1	1	1	1
PNN 4:4-6747-2-2:1	Ratio of v_j	1	1	1	1

[a] Book Codes vector components $BC_i = v_i$ = Scale*(Variable) + Shift, are the variables used to train the model that has been obtained after initial pre-processing.

$^*\mu^i{}_5$ codifies the effect of the drug structure on the biological activity, but depending on the type of assay carried out. In this sense, we have pre-multiplied $\mu^i{}_5$ by the parameters $p(a_u)$ and $p(c_l)$. The parameter p(a) is a probability (*a priori*) that codifies the propensity of an assay to yield positive results. We have defined $p(a_u) = n_1(a_u)/n_{tot}(a_u)$; where $n_1(a_u)$ and $n_{tot}(a_u)$ are the number of positive or total results for the i^{th} pharmacological assay a_i 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. Next, we have defined 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 3 we give some example of assays and their $p(a_u)$ values. In Table SM1 of the online Supplementary data file we list exhaustive values of these parameters.

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 have quantified this effect in terms of the deviation $\Delta\mu^i{}_5(m_j) = {}^{std}\mu^i{}_5 - <{}^{std}\mu^i{}_5(m_j)>$. These deviation terms represent the hypothesis: H_0 the structural dissimilarity between one compound, with respect to the average of all compounds in a group, predicts the final behavior of the compound. For instance, $\Delta\mu^i{}_5(o_t) = {}^{std}\mu^i{}_5 - <{}^{std}\mu^i{}_5(o_t)>$ measures the deviation from the average value $<\mu^i{}_5(t_e)>$ of $\mu^i{}_5$ 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 $<\mu^i{}_5(o_1)> = 18139.7$ and $<\mu^i{}_5(o_2)> = 18149.6$. This type of model able to model/interpret cross-species activity is of major importance in order to reduce assays in humans.[38] By analogy, $\Delta\mu^i{}_5(t_e) = {}^{std}\mu^i{}_5 - <{}^{std}\mu^i{}_5(t_e)>$ is the dissimilarity between the structure of compound i^{th} (expressed by $^{std}\mu^i{}_5$) with respect to all compounds active against the molecular or cellular target t_e. In Table 4 and Table 5 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.

We have also used STATISTICA to train other ANNs, such as: Radial Basis Function (RBF) and Probabilistic Neural Networks (PNN) as well as 3-layer and 4-layer Multi-Layer Perceptron (MLP3) and (MLP4).[39] In Table 1 we show a summary of the results obtained for the different ANN models trained and tested in this work. Please, note that the different ANN profiles are specified using the following notation: N_I: I-H_1-H_2-O: No. Where, NI is the number of input variables and No is the number of output variables. Furthermore, I is the number of neurons in the input layer, H_1 and H_2 are the number of neurons in the first and second hidden layers, and last O is the number of neurons in the output layer. The values of the area under the ROC curve for the LNN model are 0.99 (in

Table 3
Some examples p(a) values for different assays

CHEMBLID of a_u	$p(a_u)$	n_1	n_{tot}	Cutoff	Relation	Type	Units	Assay description
1002955	0.263	54	208	81.81	>	K_i	μM	Inhibition of MMP12
964734	0.263	54	208	81.81	<	IC_{50}	μM	Inhibition of MMP12
924957	0.263	54	208	81.81	<	$\log(1/K_i)$		Inhibition of MMP12
970762	0.743	74	100	21.28	<	IC_{50}	μM	Inhibition of TGH
660813	0.711	68	96	29.92	<	IC_{50}	μM	Inhibitory activity against recombinant human Chemokine receptor type 3 (CCR3) expressed in Chinese hamster ovary cells
1776768	0.821	77	94	55.93	<	ID_{50}	μM	Cytotoxicity against mouse J774 cells
1261026	0.821	77	94	55.93	<	EC_{50}	μM	Cytotoxicity against mouse J774 cells
940865	0.021	1	93	30.1	>	Inhibition	%	Inhibition of CCR1 at 10 μM
1674458	0.430	39	92	73.34	<	TC_{50}	μM	Cytotoxicity against mouse RAW264.7 cells after 24 h by MTT assay
1175699	0.430	39	92	73.34	<	IC_{50}	μM	Cytotoxicity against mouse RAW264.7 cells after 24 h by MTT assay
1657211	0.430	39	92	73.34	<	IC_{50}	μg mL^{-1}	Cytotoxicity against mouse RAW264.7 cells after 24 h by MTT assay
860201	0.224	16	75	19.61	>	K_i	μM	Inhibition of CSF1R
1025517	0.224	16	75	19.61	<	IC_{50}	μM	Inhibition of CSF1R
1664436	0.413	30	74	12.7	>	Inhibition	%	Inhibition of mouse recombinant iNOS at 1 mM after 40 min by colorimetric assay
867926	0.840	62	74	0.22	<	IC_{50}	μM	Inhibition of LPS-induced TNFalpha production in human monocytes
1285558	0.222	15	71	178.34	>	K_i	μM	Inhibition of mouse recombinant iNOS
957262	0.222	15	71	178.34	<	IC_{50}	μM	Inhibition of mouse recombinant iNOS
921708	0.130	8	68	2.04	>	Selectivity ratio		Inhibition of cFms
998565	0.130	8	68	2.04	<	IC_{50}	μM	Inhibition of cFMS

Table 4
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 5
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 $CC50/IC_{50}$	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	E_{max}	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

training and test sets) both very close to 1 (the highest possible value for a perfect classifier) and notably different from 0.5 (the typical value of a random classifier). The Topologies of the different ANN models trained are shown (Fig. 2). All these ANN models were more complex than the LNN model and the values of Ac, Sn, and/or Sp obtained are similar or even lower. We have also validated the model by means of a ROC curve[22] analysis (see Fig. 3).

2.2. Experimental-theoretical study of G1 anti-microbial drug

2.2.1. Experimental results

We have studied a novel series of furylethylene compounds that are isomers of position of the 5-nitrofurylethylene derivatives, whose common characteristic is that the nitro group is attached to the β-carbon by an exocyclic double bond of the ethylenic chain.

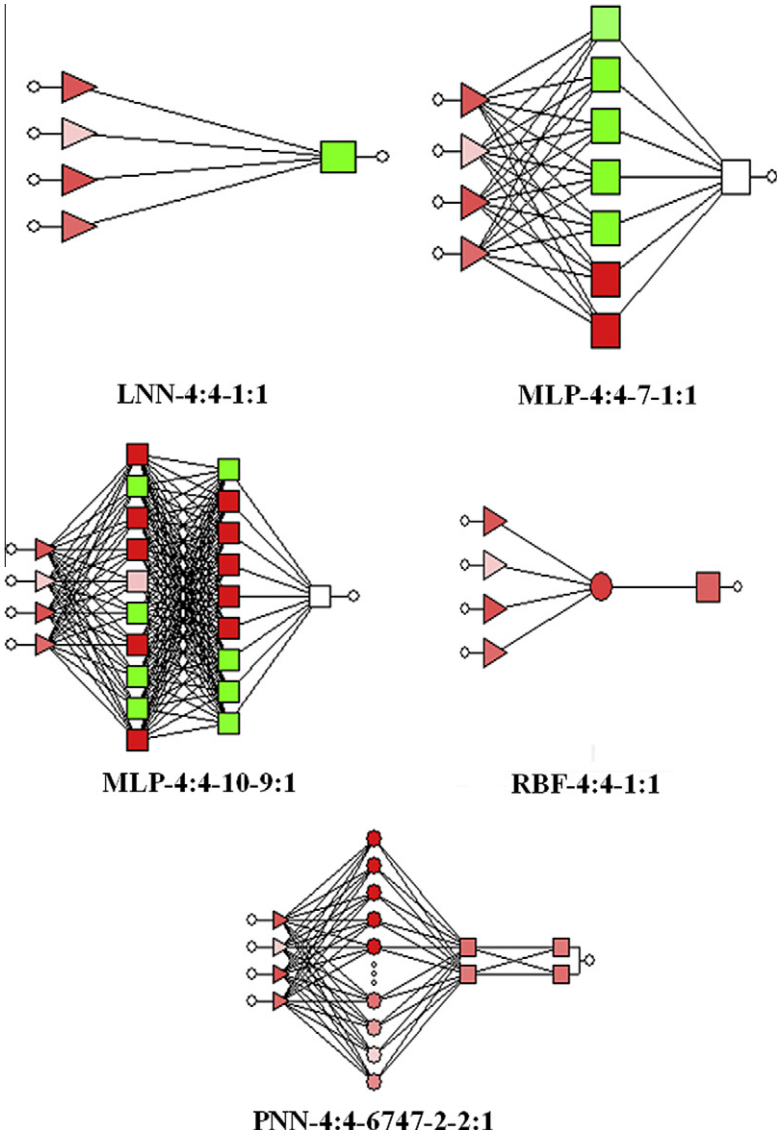

Figure 2. Topology of the different ANN models trained (showing hot-map activation of neurons for assay endpoint number 1).

The G-1 compound is one of the members of a new family of furylethylene derivatives with both anti-bacterial and anti-fungal properties.[40] These derivatives have shown interesting biological properties, with potential uses in human and veterinary medicine as anti-microbial and anti-fungal agents. More recently, anti-parasitic activity has been also reported.[41] The compound was synthesized in the laboratories of the Chemical Bioactives Center at the Central University of Las Villas (UCLV), Cuba. Nitrovinilfuran 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 microbicidal 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.[42]

2.2.2. Cytotoxicity assays

Cytotoxicity assays are widely used by the pharmaceutical industry to screen for cytotoxicity in compound libraries. On the other hand, the quantifying cell viability or cytotoxicity is crucial to understand cancer biology, compound toxicity and cellular response to cytokines and other biological questions. Researchers can either look for cytotoxic compounds, if they are interested in developing a therapy that targets rapidly dividing cancer cells. The endpoint measurements and assays are very used in cytotoxicity tests. Assessing cell membrane integrity is one of the most common ways to measure cell viability and cytotoxic effects. The compounds that have cytotoxic effects often compromise cell membrane integrity. Due to different methods, there have been recent developments which offer increased sensitivity, throughput, and specificity. This requires multiplexing of methods, or methods that are able to distinguish between the different cell states and different endpoint evaluated.[43,44] The term 'multiplexing' simply means gathering more than one set of data from the same sample. The key benefit of multiplexing is that you gain a better understanding of the event you are measuring in the context of another parameter, minimizing faulty interpretation of data or ambiguity from data sets.[45,46] In our study, we have used only the detection of membrane integrity by staining with 7AAD and flow cytometry. Several parameters have been 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.[47] The analyses have been performed with flow cytometry; in order to follow the percentage of live macrophages present in the macrophages

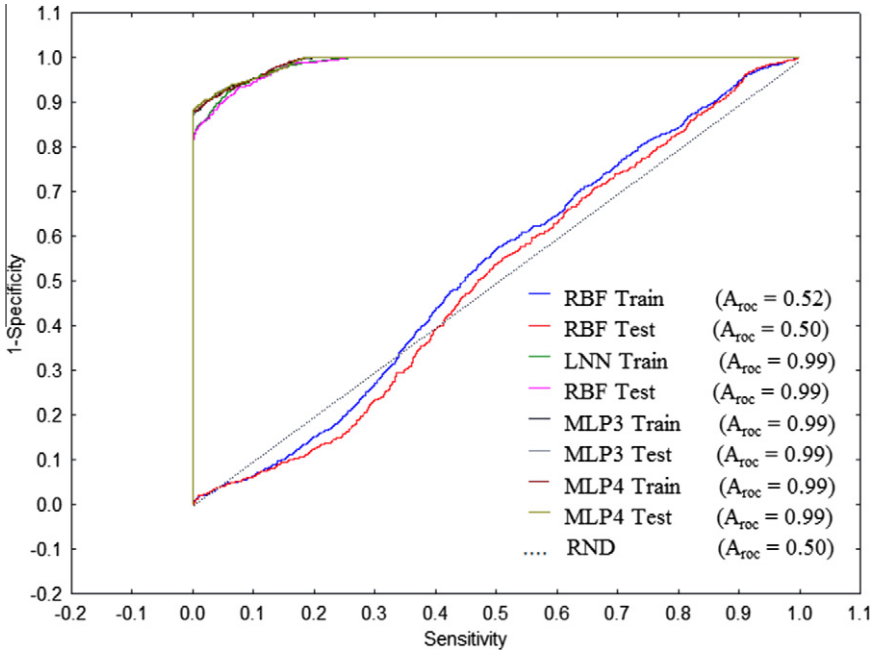

Figure 3. ROC analysis of the ANN models trained.

populations treated with G1 at different concentrations we have observed changes in the viability of the macrophages after 24 hours. The assay shows a significant increase of dead cells, Cytotoxicity (%) = 27.13%, compared to the untreated group (6.4%) and the DMSO group (15.4%) at c_{max} = 10 μg/mL. In other studies with the product in lymphocyte populations, cytotoxicity can be observed at the maximum reported concentration, 15 μg/mL.[48] The percent of cytotoxicity is similar in concentrations 8, 6 and 4 μg/mL (approximately 23%). The treatment of 2 μg/mL results is significant, approximately similar to cytotoxicity of the control vehicle (19.1%) (Fig. 4). Organic solvents, such as DMSO and methanol are widely used as vehicles to solubilize lipophilic test compounds in toxicity testing. The dimethylsulfoxide (DMSO) is usually used to solubilize poorly soluble drugs in permeation assays.[49] Reports of the cytotoxicity of DMSO have been mentioned in several published articles; we recommend using it to 0.5% in cell culture.[50] These results indicate that G1 has low cytotoxicity at this concentration (10 μg/mL) as cytotoxicity (%) <50% which is lower than a threshold value is considered for cytotoxic compounds.[51] We

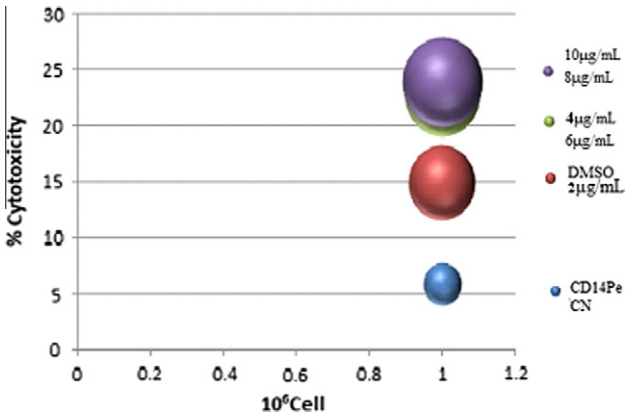

Figure 4. Percent cytotoxicity (%) in macrophages spleen of mouse Balb/C exposed to different concentrations of the anti-microbial G1.

should consider that the use of DMSO as a solvent may increase the cytotoxicity if formulated at inappropriate concentrations.

On the other hand, in the flow cytometer the particles and cells, which are detected in forward scatter (FS) and side scatter (SS) mode, pass through the beam scatter light. The FSC correlates with the cell size and SS depends on the density of the particle/cell (granularity). The forward scatter detector can be used to determine the ratio of dead cells. These cells are more permeable to certain fluorophores.[52] In addition, the flow cytometer can determine the granularity (SS) of the cell population.[53] The difference in side scattering between granulocytes, lymphocytes, monocytes, and other cell types is clear and can be used for measuring their ratios without the need for a stain. 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 versus Side Scatter (FSC/SSC) cluster. 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).[54] Another parameter evaluated in our experiment of cytotoxicity was MFI. These spleen macrophages were exposed to different concentrations of G1. In Fig. 5, we depict a pseudo-color and zebra projection of MFI values over FSC versus SSC plot after exposure of G1 at $c_{max.}$ Fig. 5A shows the total purchase including the population of lymphocytes and macrophages. The highly homogenous macrophages population was defined by the expression of CD14. Then, Fig. 5B shows the region top left (SSC-H^+,FSC-H^-) that represents a 3.48% of the macrophages identified by granularity. The regions of top right (SSC-H^+,FSC-H^+) positive include the population of macrophages as identified by size and granularity (96.2%). The 5C and 5D show the regions (R1 and R2) of macrophages of the total population. The gate represents 16.1% and 26.4% respectively. 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.[55]

In Table 6, we show the average values of MFI in SSC and/or FSC scattering mode, for all samples (Negative Control, DMSO,

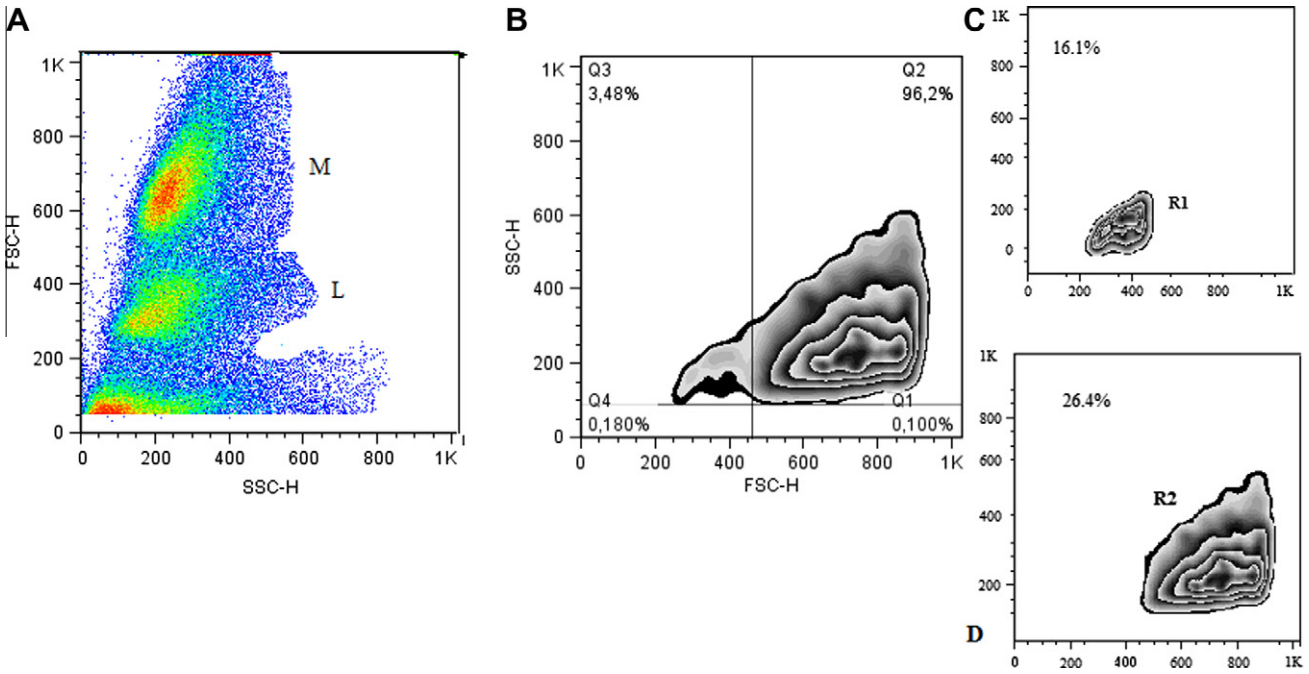

Figure 5. Pseudo color—Zebra of percentage values FSC versus SSC plot after administration of G1 at c_{max}. Total cell spleen population (A) with macrophages (M) and Lymphocytes (L). The parameters are SSC-H,FSC-H subs-set macrophages population (B) SSC-H$^+$,FSC-H$^-$ macrophages population (C) and SSC-H$^+$, FSC-H$^+$ macrophages population (D) analyses were conducted on 10,000 cells in each case.

Table 6
Results of statistical analysis of the parameters of the flow cytometer showed significant differences between different concentrations and control groups

Group1 Conc (μg/mL)	Group2			
	Parameters	NC p	DMSO p	MFI
10	SCC-H+FSC-H+	0.702502	0.128479	407 ± 99.40
	SCC-H+CD14Pe-	0.994129	0.072907	106.2 ± 21.84
	CD14Pe+7AAD+	0.033803[a]	0.117235	24.8 ± 5.36
	CD14Pe+7AAD-	0.028235[a]	0.036801[a]	12.8 ± 3.90
8	SCC-H+FSC-H+	0.436929	0.041706	359.7 ± 114.25
	SCC-H+CD14Pe-	0.942073	0.021946[a]	91.89 ± 22.33
	CD14Pe+7AAD+	0.039352[a]	0.140063	30.6 ± 7.94
	CD14Pe+7AAD-	0.120769	0.184298	14.65 ± 4.46
6	SCC-H+FSC-H+	0.995074	0.605653	431.8 ± 97.0
	SCC-H+CD14Pe-	0.999957	0.169451	91.22 ± 29.83
	CD14Pe+7AAD+	0.107138	0.329536	28.65 ± 8.25
	CD14Pe+7AAD-	0.114631	0.174308	13.65 ± 3.83
4	SCC-H+FSC-H+	1.000000	0.928208	465.9 ± 88.25
	SCC-H+CD14Pe-	0.906891	0.022344[a]	89.93 ± 28.03
	CD14Pe+7AAD+	0.105049	0.324203	27.2 ± 7.80
	CD14Pe+7AAD-	0.115863	0.176316	12.85 ± 4.70
2	SCC-H+FSC-H+	0.996486	0.631624	450.5 ± 92.97
	SCC-H-+CD14Pe-	0.999802	0.141323	95.63 ± 31.62
	CD14Pe+7AAD+	0.181544	0.497724	21.9 ± 10.40
	CD14Pe+7AAD-	0.637972	0.855528	11.65 ± 4.57

[a] Significant differences. NC is control negative and DMSO is control vehicle.

and CD14Pe phenotypic marker macrophages exposed to G1). Table 6 shows the results of a statistical analysis performed on all MFI values obtained from flow cytometer. The SSC-H+, FSC-H+ shows no significant differences at p <0.05. The only significant differences have been observed in CD14Pe+7AAD+ at 10 μg/mL compared to control (CN), which shows that there is some cytotoxicity in macrophages, thus corresponding to the results of cytotoxicity percentage calculated for this population. This is referred to in the preceding paragraphs. Although shown in the same concentration, there are differences for unstained (7-AAD) macrophages. MFI and cell count (Event count) in FSC scattering mode give an idea of the cellular size, while the same parameters, but in SSC scattering mode, measure internal cellular damage.[56] It has been observed that there was a higher damage to the granularity. We have used the STATISTICA software for both means and ANOVA analysis.[22] The ANOVA analysis was carried out applying Tukey's method. We confirm that there no significant differences for treated samples of living macrophages with respect to control groups 6, 4 and 2 μg/mL (see Table 6 and Fig. 6).

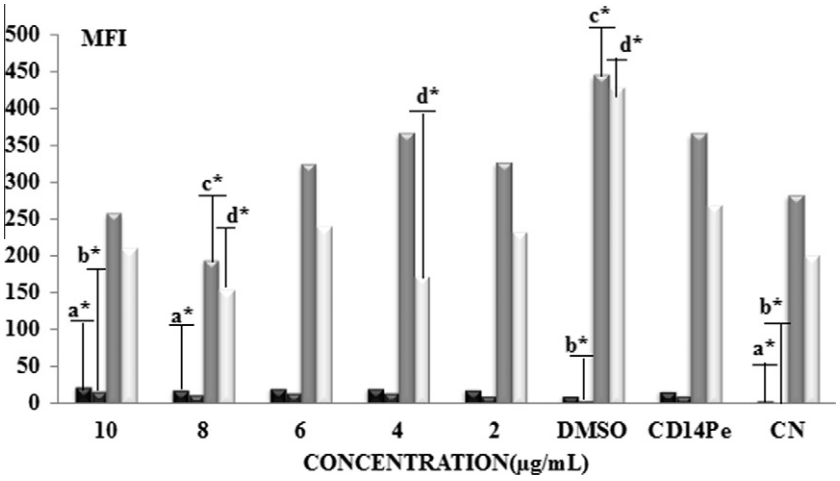

Figure 6. Analysis in flow cytometer of the effect of G1 on primary culture of spleen macrophages which were exposed to different concentrations (10, 8, 6, 4, and 2 µg/mL) for a period of 24 h. The results are expressed as Mean Intensity Fluorescence (MFI) the $N = 6$ animals per group and 1.106 Cell. *Significant changes compared to controls ($p < 0.05$), Control vehicle (DMSO) control antibody (CD14Pe) and negative control (CN). The parameters evaluated are a(**CD14⁺**,7AAD⁻), b(**CD14⁺**,7AAD⁺), c(**SSC-H⁺,**FSC-H⁺) and d(**CD14⁻,** SSC-H⁺).

The number of cells acquired was 10,000 in a range between 2000 and 4000 events. In addition, CD14 PE was used as a macrophages marker in the presence of 7AAD, as described in the Section 3. Altogether, 46.5% of macrophages were marked with CD14Pe and 7AAD (Fig. 6 region Q6) or region R4 with gate 43.4%. MFI average was 24.8. R1 represents the population stained with 7AAD but not macrophages (0.98%) with gate (85.05%). R5 (51.1%) represents dead cells (3.46%) and R6 (30.9%) macrophages marked with CD14Pe but not stained with 7AAD (49%), see Fig. 7.

Finally, EC_{50} calculations using the best fitting feature methodology have been shown below. The results show that the best dose-response curve was the five parameter logistic (1/Ŷ2) for all MFI obtained in flow cytometer. The granularity (SSC-H+) showed an $R^2 = 0.9676$. The Root Mean Square Error (RMSE) was equal to 0 and the $EC_{50} = 16.05$. This study calculated the EC_{50}, observing that these values differ in relation to the date obtained in the flow cytometer (Table 7). The EC_{50} calculated with all parameters read by the flow cytometer was 11.41. Heterogeneous variability or *heteroscedasticity* occurs in most types of observed data. Non-constant or heterogeneous variability often complicates the fitting of non-linear models to observed data. This is especially true for biochemical assays where concentration or dose is the predictor. The non-linear (or linear) least square algorithm assumes that all points have the same variability, so all points influence the curve fit equally.[57] Many methods of regression analysis are based on the assumption of equal variances, but MasterPlex ReaderFit software used to calculate the EC_{50} offering 4 different weighting algorithms to account for heteroscedasticity. The five-parameter logistic regression model 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/Ŷ2, which minimizes residuals (errors) based on relative Mean Fluorescence Intensity MFI/RLU values. The need for a curve model that accommodates asymmetry has been acknowledged by improvements in instrument and laboratory technology. The reason for this is that even modest levels of lack-of-fit error caused by fitting mildly

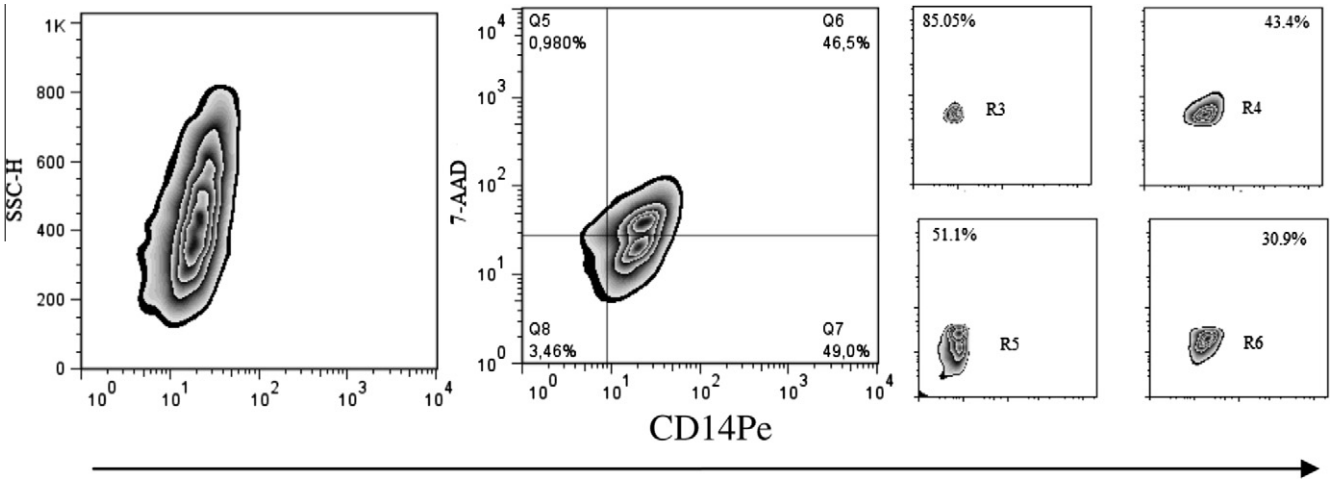

Figure 7. Zebra graph type. Gating strategy for macrophages CD14Pe labeling analysis and stained with 7AAD. Results of flow cytometry for Balb/C mouse spleen macrophages exposed to G1 at 10 µg/mL.

Table 7
Results of dose versus effect EC_{50} curve fitting by different parameters of flow cytometer

Parameters	EC_{50}	Curve fitting	R^2	RMSE	a	b	c	d	e
SSC-H+	16.05	5PL (1/Ŷ2)[a]	0.9676	0	−2481689	13.8884	4.5431	311.5956	0.0351
SSC-H-	9.37	5PL (1/Ŷ2)[a]	0.8694	0	173.6849	1044848	12.8058	252.8945	4.5657
FSC-H+	32.87	5PL (1/Ŷ2)[a]	0.8079	0	52.1223	14.1604	4.7643	584.8720	0.0222
FSC-H-	7.97	5PL (1/Ŷ2)[a]	0.9407	0	−252.9625	3.8792	9.4500	571.3106	0.7831
FL2-H+	0.12	5PL (1/Ŷ2) [a]	0.5120	0	−1.2405	−144877	5.7565	16.4048	0.0110
FL2-H-	n.v	5PL (1/Ŷ2)[a]	−0.2563	0	−70.9054	0.01247	1405.32	210.361	1.6244
FL3-H	1.94	5PL (1/Ŷ2) [a]	0.0107	0	−9.9706	0.0112	00000	124.1797	1.2477
FL3-H-	11.41	5PL (1/Ŷ2)[a]	0.7673	0	−20.5232	−14.896	1.27132	32.6552	107.72

[a] Best fit model, 5PL is five parameters logistic, the parameters are FL2 (CD14Pe), FL3 (7-AAD), SSC-H(side scatter), FSC (size scatter) n.v is not values.

asymmetric data to a symmetric model can dominate the pure error due to random variation in low-noise modern assays. The five-parameter logistic function, which includes a fifth parameter, permits asymmetry to be effectively modeled.[58] Example of the formula for positive Side Scatter analysis (SSC-H+):

$$\left(\frac{1}{MFI_i^2}\right) = -2481689 - \frac{311.5956}{\left(1+\left(\left(\frac{c_i}{4.5431}\right)^{13.8884}\right)\right)^{0.0351}} \quad (5)$$

In conclusion, two of the highest concentrations showed some cytotoxicity but it is noted 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 a 24 h exposure. The EC_{50} values depended on the parameters measured in the flow cytometer. The method used for the analysis of all data obtained in the flow cytometer was 5PL using (1/Ŷ2).

2.2.3. Prediction of G1 cytotoxicity for other assays

Altogether, we have predicted 1265 multiplexing assay endpoints for G1 biological activities, including 152 cytotoxicity assays. Notably, the LLN model has predicted a very low score $S_{G1}(m_j) = 0.29$ for G1 cytotoxicity in assay conditions $m_j = TC_{50}$ <100 μM against human macrophages. Interestingly, in all cases the model has predicted a low score $S_{G1}(m_j) = 0.19$ for G1 cytotoxicity effect against J774 cell line. However, the model has also predicted low score for cytotoxicity in other assays using RAW264.7 cell (Monocytic-macrophage leukemia cells). The same Table 8 shows that G1 presenting a 0.20 score, could inhibit such cells in some specific assay conditions. The J774 and Raw 264.7 macrophage cell lines are well-established model systems in cell biology and immunology.[59]

The predictive score obtained for this compound in the cytotoxicity assay against EL4 cell line was 0.21. EL4 cells have been shown to secrete IL2, CSF, and MAF (macrophage activating factor) into the conditioned medium. Recently, EL4 thymoma cells have been found to be a source of biologically active CD40 ligand. In the macrophage, the primary signal for activation is IFN-γ from Th1-type CD4 T cells.[60] The secondary signal is CD40L (CD154) on the T cell, which binds CD40 on the macrophage cell surface.[61] As a result, the macrophage expresses more CD40 and TNF receptors on its surface, helping to increase the level of activation. The increase in activation results in the induction of potent microbicidal substances in the macrophage, including reactive oxygen species and nitric oxide, leading to the destruction of ingested microbe.[62,63]

On the other hand, Table 9 shows the theoretic prediction of some endpoints for G1 interaction with human protein targets. G1, showing a 0.26 score to act on Dipeptidyl peptidase IV protein is expressed on the surface of most cell types and is associated with the immune system.[64] This drug also shows a 0.16 chance to act on the macrophage migration inhibitory factor (MIF). The MIF plays an important role in the regulation of macrophage function in host defense through the suppression of anti-inflammatory effects of glucocorticoids. Our study predicted the activity of G1 in 618 proteins *Mus musculus* and *Homo sapiens*.

3. Materials and methods

3.1. Computational methods

3.1.1. CHEMBL dataset

This dataset includes $N_d = 7903$ drugs and/or organic compounds previously assayed in different multiplexing assay conditions (m_j). Every drug evaluated in different m_j conditions was assigned 1 out of 2 possible activity classes: active ($C = 1$) or non-active compounds ($C = 0$). One compound may lead to 1 or more statistical cases because it may give different outcomes (statistical cases) for alternative biological assays carried out in diverse sets of multiplex conditions. In this work, we have defined m_j according to the ontology $m_j = >(a_u, c_l, o_t, t_e, s_x)$. A general data set made up of >10,000 multiplexing assay endpoints was downloaded from the public database CHEMBL.[6] In any case, after a careful curation of the dataset, we have retained 9000 multiplexing assay endpoints (statistical cases) after eliminating all cases with missing information or very low representation. The different conditions that may change in the dataset are different: organisms (o_t), biological assays (a_u), molecular or cellular targets (t_e), or standard type of activity measure (s_x). In closing, we have analyzed $N = 9000$ statistical cases consisting of the above mentioned $N_d = 7903$ drugs, which have been assayed each one in at least one out of $N_a = 1418$ possible assays. For each one of these assays, the dataset studied presented for each drug at least one out of $N_s = 46$ standard types of biological activity measures in turn carried out in at least one out of $N_t = 36$ molecular or cellular targets. These values have been reported in CHEMBL as results of experiments carried out on at least 1 out of 3 possible organisms $N_o = 3$ (*Homo sapiens* and *Mus musmuculus*). The values are reported in CHEMBL with three different levels of data curation $N_c = 3$ (expert, intermediate, or autocuration level). Please, see details on the assignation of cases to different classes in the Section 2.

3.1.2. Artificial Neural Network (ANN) models

In order to seek the mt-QSAR model with different ANNs implemented in the STASTICA 6.0 software package.[39] The Linear Neural Network (LNN) model has the following general form:

$$S_i(m_j) = b_0 + b_1 \cdot {}^{*}\mu_5^i + \sum_{j=2}^{4} b_j \cdot \Delta\mu_5^i(m_j) \quad (6)$$

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

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 selected conditions m_j. The statistical parameters which used to corroborate the model

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

C	$S_{G1}(m_j)$[a]	Type	Rel	Cutoff[b]	Units	Assay description (m_j)[c]
Multiplexed Endpoints for G1 cytotoxicity experimentally determined in this work						
0	11.41	EC_{50}	<	10	μg	7ADD mouse peritoneal macrophages after 24 h
0	27.1	Cytotoxicity	>	50	%	7ADD mouse peritoneal macrophages after 24 hrs at 10 μg/mL
0	23.6	Cytotoxicity	>	50	%	7ADD mouse peritoneal macrophages after 24 hrs at 8 μg/mL
0	23.6	Cytotoxicity	>	50	%	7ADD mouse peritoneal macrophages after 24 hrs at 6 μg/mL
0	23.1	Cytotoxicity	>	50	%	7ADD mouse peritoneal macrophages after 24 hrs at 4 μg/mL
0	19.1	Cytotoxicity	>	50	%	7ADD mouse peritoneal macrophages after 24 hrs at 2 μg/mL
1	—	MFI	<	p <0.05 000.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 10 μg/mL
1	—	MFI	<	p <0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 8 μg/mL
0	—	MFI	<	p <0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 6 μg/mL
1	—	MFI	<	p <0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 4 μg/mL
0	—	MFI	<	p <0.05	%	7ADD mouse peritoneal macrophages after 24 hrs at 2 μg/mL
Multiplexed Endpoints for G1 cytotoxicity predicted in this work						
1	0.29	TC_{50}	<	100	μM	Cytotoxicity against human macrophages
1	0.20	IC_{90}	<	13,11083	μM	Cytotoxicity against rat RAW264.7 cells by MTT assay
1	0.15	ED_{50}	<	11,400	μM	Cytotoxicity against mouse WEHI cell line by MTT assay
1	0.13	ED_{50}	<	45,37279	μM	Cytotoxic activity against mouse macrophage cell line J7.COL-100.
1	0.22	ED_{50}	<	17,09063	μM	Cytotoxic activity against mouse macrophage cell line J7.VBL-1
1	0.19	EC_{50}	<	55,9367	μM	Cytotoxicity against mouse J774 cells
1	0.05	EC_{50}	<	9,1875	μM	Cytotoxicity against macrophage cell line (J774)
1	0.21	EC_{50}	<	164,67024	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell survival after 24 hrs by MTT assay
1	0.21	IC_{50}	<	0.374	μM	Cytotoxicity against human EOL1cells
1	0.19	ED_{50}	<	23.83	μg ml-1	Cytotoxicity against mouse RAW264.7 cells by MTT colorimetric assay in presence of LPS
1	0.17	IC_{50}	<	143,79189	μM	Cytotoxicity against J774.1 cell line after 48 hrs
1	0.10	IC_{50}	<	26,72973	μM	Cytotoxicity against mouse J774 cells expressing RANKL signaling by Alamar blue assay
1	0.23	IC_{50}	<	416.67	μg ml-1	Cytotoxicity against mouse J774 cells after 24 hrs
1	0.22	IC_{50}	<	72,78947	μM	Cytotoxicity against mouse J774 macrophages after 48 hrs by MTT assay
1	0.21	IC_{50}	<	736,200	μM	Cytotoxicity against mouse J774 cells assessed as cell viability after 48 hrs by alamar blue assay
1	0.16	Activity	>	99.42	%	Cytotoxicity against mouse J774 macrophage assessed as cell viability at 10 μg/mL after 48 hrs by MTT assay
1	0.09	IC_{50}	<	508,57331	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell viability after 4 hrs by MTT assay
1	0.09	IC_{50}	<	2256,35263	μM	Cytotoxicity against mouse J774 cells after 24 hrs by resazurin assay
1	0.25	IC_{50}	<	565.75	μg ml-1	Cytotoxicity against mouse J774 cells after 24 hrs by MTT assay
1	0.30	Activity	>	98.77	%	Cytotoxicity against mouse J774 cells assessed as cell viability at 0.1 μg/mL after 24 hrs by MTT assay
1	0.15	IC_{50}	<	221,07843	μM	Cytotoxicity against mouse J774 cells after 48 hrs by MTT assay
1	0.10	Activity	>	97.6	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 10 μg/ml by MTT assay
1	0.19	Activity	>	98.4	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 1 μg/ml by MTT assay
1	0.18	IC_{50}	<	236,69231	μM	Cytotoxicity against mouse J774 cells by MTT assay
1	0.20	IC_{90}	<	13,11083	μM	Cytotoxicity against rat RAW264.7 cells by MTT assay
1	0.15	ED_{50}	<	11,400	μM	Cytotoxicity against mouse WEHI cell line by MTT assay
1	0.13	ED_{50}	<	45,37279	μM	Cytotoxic activity against mouse macrophage cell line J7.COL-100.
1	0.22	ED_{50}	<	17,09063	μM	Cytotoxic activity against mouse macrophage cell line J7.VBL-1
1	0.19	EC_{50}	<	55,9367	μM	Cytotoxicity against mouse J774 cells
1	0.05	EC_{50}	<	9,1875	μM	Cytotoxicity against macrophage cell line (J774)
1	0.21	EC_{50}	<	164,67024	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell survival after 24 hrs by MTT assay
1	0.21	IC_{50}	<	374	μM	Cytotoxicity against human EOL1cells
1	0.19	ED_{50}	<	23.83	μg ml-1	Cytotoxicity against mouse RAW264.7 cells by MTT colorimetric assay in presence of LPS
1	0.17	IC_{50}	<	143,79189	μM	Cytotoxicity against J774.1 cell line after 48 hrs
1	0.10	IC_{50}	<	26,72973	μM	Cytotoxicity against mouse J774 cells expressing RANKL signaling by Alamar blue assay
1	0.23	IC_{50}	<	416.67	μg ml-1	Cytotoxicity against mouse J774 cells after 24 hrs
1	0.22	IC_{50}	<	72,78.47	μM	Cytotoxicity against mouse J774 macrophages after 48 hrs by MTT assay
1	0.21	IC_{50}	<	736,200	μM	Cytotoxicity against mouse J774 cells assessed as cell viability after 48 hrs by alamar blue assay
1	0.16	Activity	>	99.42	%	Cytotoxicity against mouse J774 macrophage assessed as cell viability at 10 μg/mL after 48 hrs by MTT assay
1	0.09	IC_{50}	<	508,57331	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell viability after 4 hrs by MTT assay
1	0.09	IC_{50}	<	2256,35263	μM	Cytotoxicity against mouse J774 cells after 24 hrs by resazurin assay
1	0.25	IC_{50}	<	565.75	μg ml-1	Cytotoxicity against mouse J774 cells after 24 hrs by MTT assay
1	0.30	Activity	>	98.77	%	Cytotoxicity against mouse J774 cells assessed as cell viability at 0.1 μg/mL after 24 hrs by MTT assay
1	0.15	IC50	<	221,07843	μM	Cytotoxicity against mouse J774 cells after 48 hrs by MTT assay
1	0.10	Activity	>	97.6	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 10 μg/ml by MTT assay
1	0.19	Activity	>	98.4	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 1 μg/ml by MTT assay
1	0.18	IC_{50}	<	236,69231	μM	Cytotoxicity against mouse J774 cells by MTT assay
1	0.20	IC_{90}	<	13,11083	μM	Cytotoxicity against rat RAW264.7 cells by MTT assay

(*continued on next page*)

Table 8 (*continued*)

C	$S_{G1}(m_j)$[a]	Type	Rel	Cutoff[b]	Units	Assay description (m_j)[c]
1	0.15	ED_{50}	<	11,400	μM	Cytotoxicity against mouse WEHI cell line by MTT assay
1	0.13	ED_{50}	<	45,37279	μM	Cytotoxic activity against mouse macrophage cell line J7.COL-100.
1	0.22	ED_{50}	<	17,09063	μM	Cytotoxic activity against mouse macrophage cell line J7.VBL-1
1	0.19	EC_{50}	<	55,9367	μM	Cytotoxicity against mouse J774 cells
1	0.05	EC_{50}	<	9,1875	μM	Cytotoxicity against macrophage cell line (J774)
1	0.21	EC_{50}	<	164,67024	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell survival after 24 hrs by MTT assay
1	0.21	IC_{50}	<	0.374	μM	Cytotoxicity against human EOL1cells
1	0.19	ED_{50}	<	23.83	μg ml-1	Cytotoxicity against mouse RAW264.7 cells by MTT colorimetric assay in presence of LPS
1	0.17	IC_{50}	<	143,79189	μM	Cytotoxicity against J774.1 cell line after 48 hrs
1	0.10	IC_{50}	<	26,72973	μM	Cytotoxicity against mouse J774 cells expressing RANKL signaling by Alamar blue assay
1	0.23	IC_{50}	<	416.67	μg ml-1	Cytotoxicity against mouse J774 cells after 24 hrs
1	0.22	IC_{50}	<	72,78947	μM	Cytotoxicity against mouse J774 macrophages after 48 hrs by MTT assay
1	0.21	IC_{50}	<	736,200	μM	Cytotoxicity against mouse J774 cells assessed as cell viability after 48 hrs by alamar blue assay
1	0.16	Activity	>	99.42	%	Cytotoxicity against mouse J774 macrophage assessed as cell viability at 10 μg/mL after 48 hrs by MTT assay
1	0.09	IC_{50}	<	508,57331	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell viability after 4 hrs by MTT assay
1	0.09	IC_{50}	<	2256,35263	μM	Cytotoxicity against mouse J774 cells after 24 hrs by resazurin assay
1	0.25	IC_{50}	<	565.75	μg ml-1	Cytotoxicity against mouse J774 cells after 24 hrs by MTT assay
1	0.30	Activity	>	98.77	%	Cytotoxicity against mouse J774 cells assessed as cell viability at 0.1 μg/mL after 24 hrs by MTT assay
1	0.15	IC_{50}	<	221,07843	μM	Cytotoxicity against mouse J774 cells after 48 hrs by MTT assay
1	0.10	Activity	>	97.6	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 10 μg/ml by MTT assay
1	0.19	Activity	>	98.4	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 1 μg/ml by MTT assay
1	0.18	IC_{50}	<	236,69231	μM	Cytotoxicity against mouse J774 cells by MTT assay
1	0.20	IC90	<	13,11083	μM	Cytotoxicity against rat RAW264.7 cells by MTT assay
1	0.15	ED_{50}	<	11,400	μM	Cytotoxicity against mouse WEHI cell line by MTT assay
1	0.13	ED_{50}	<	45,37279	μM	Cytotoxic activity against mouse macrophage cell line J7.COL-100.
1	0.22	ED_{50}	<	17,09063	μM	Cytotoxic activity against mouse macrophage cell line J7.VBL-1
1	0.19	EC_{50}	<	55,9367	μM	Cytotoxicity against mouse J774 cells
1	0.05	EC_{50}	<	91,875	μM	Cytotoxicity against macrophage cell line (J774)
1	0.21	EC_{50}	<	164,67024	μM	Cytotoxicity against mouse RAW264.7 cells assessed as cell survival after 24 hrs by MTT assay
1	0.21	IC_{50}	<	0.374	μM	Cytotoxicity against human EOL1cells
1	0.19	Activity	>	97.6	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 10 μg/ml by MTT assay
1	0.17	Activity	>	98.4	%	Cytotoxicity against mouse J774 cells infected with Mycobacterium bovis BCG assessed as cell viability at 1 μg/ml by MTT assay

[a] $S_{ij}(1)$ is the score calculated with the model using the STATISTICA software.
[b] 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 cell (MMLC). CAM is Cytotoxicity Against Macrophage.
[c] This is a short description, see description for all assays in Supplementary data 1 (SD1).

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

Target name	ID	$S_{G1}(m_j)$	Standard type	Rel	Cutoff	Unit
Dipeptidyl peptidase IV	3883	0.26	ED_{50}	<	10.1	mg kg^{-1}
C–C chemokine receptor type 3	3473	0.19	E_{max}	>	56.25	%
M-CSFR[a]	1844	0.21	$logIC_{50}$	<	161.92	
MIF factor[b]	2085	0.16	IC_{50}	<	65.124	μM
C–C chemokine receptor type 1	2413	0.21	ED_{50}	<	0.00015	μM
Nitric oxide synthase, inducible	3464	0.21	IC_{50}	<	718.83	μM
Matrix metalloproteinase 12	4393	0.21	IC_{50}	<	5.81	μM
ACoe[c]	2265	0.21	IC_{50}	<	8.03	μM
Cyclooxygenase-2	4321	0.10	IC_{50}	<	0.26	μg ml^{-1}
Dipeptidyl peptidase IV	3883	0.13	SA[d]	>	93.33	%

[a] M-CSFR is Macrophage colony stimulating factor receptor.
[b] MIF is Macrophage migration inhibitory factor.
[c] ACoe is Acyl coenzyme A:cholesterol acyltransferase.

were: Number of cases in training (N), overall values of Specificity (Sp), Sensitivity (Sn), Accuracy (Ac), and Area under ROC curve (A_{roc}).[22] In this model, $^{std}\mu_5{}^i$ is the spectral moment of order $k = 5$ calculated with Modeslab. We have used a standard bond distance (std) as entries of the main diagonal of the bond adjacency matrix. The parameter $p(a_u)$ is a probability, calculated *a priori,* with which any drug is expected to give a positive result in the u^{th} assay a_u. The parameter $p(c_l)$ is a probability, calculated *a priori*, of confidence for a given data value into the CHEMBL dataset studied. The structural deviation terms $\Delta\mu^i{}_5(m_j) = {}^{std}\mu^i{}_5 - <{}^{std}\mu^i{}_5(m_j)>$ represent the hypothesis H_0. H_0: the different deviations of the i^{th} drug (d_i) with respect to the average of all positive drugs for different multiplexing

assay conditions (m_j) predicting the final behavior of the compound. See a detailed discussion on terms and m_j conditions in the Section 2. This type of deviation-like mt-QSAR models has been successfully used by other groups to solve different problems.[65] We have also used STASTICA 6.0[39] software package to train and test other ANNs with non-linear forms. The ANNs trained were Multi-Layer Perceptrons (MLPs) with three (MLP3) or four (MLP4) neuron layers, Radial Basis Function (RBF), and Probabilistic Neural Networks (PNN).[39] These non-linear ANNs used as inputs the same parameters as the LNN described above.

3.2. Biology assays

3.2.1. Reagents and antibody

1-5-Bromofur-2-il-2-bromo-2-nitroethene (G1); CAS number 35950-55-1, was kindly supplied from the Chemical Bioactives Centre, Sample purity was 99.93%. G1 was dissolved in dimethyl sulfoxide (DMSO), which was purchased from Sigma–Aldrich Co. (DF, México). Macrophages were stained with phycoerythrin (PE) labeled 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. Anti-CD14 antibody (used to label CD14 receptors) and 7-aminoactinomycin (7AAD) at 5 μg/mL viability solution were purchased from BD (BD Biosciences, México).

3.2.2. 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%; 12 h light/dark cycle) and given *ad libitum* access to food and water. This work agreed with Ethical Principles in Animal Research adopted by México.[66]

3.2.3. Spleen macrophages culture

After sacrifice, spleens were aseptically removed and cells were suspended at a concentration of 1×10^6/mL in RPMI 1640 (BD. México) supplemented with 10% fetal calf serum, 1% and cultured in flat-bottomed 96-well plates. Cells were cultured in triplicates (300 μL/well) and incubated at 37 °C and 5% CO2. Cell suspensions were washed twice by centrifugation. Cell viability (over 95%) was determined using trypan blue exclusion. macrophages numbers were adjusted to 1×10^6 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% CO_2 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.

3.2.4. Determination of cytotoxicity percentage by flow cytometry analysis

In all cases, 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.

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

.The 7AAD molecule is uncharged at neutral pH. It can be excited at 488 nm, albeit giving a less intense signal than PI. However, the longer wavelength of the 7AAD fluorescence makes it particularly suitable for dead cell detection simultaneously with immunofluorescent staining with phycoerythrin (PE). Emission maxima approximately 580 nm-conjugated antibodies because the emission spectra of 7AAD and PE overlap less than the spectra of PI and PE .The excitation and emission wavelength maxima (nm) of the dyes when complexes with double-stranded DNA in aqueous solution 7AAD are 546/647.[67] Briefly, 1×10^6 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. The percent cytotoxicity was determined by the following formula,[68] where Ma means macrophages count, and the symbol (+) indicates a positive answer to CD14Pe and Negative (−) means negative to 7ADD staining for living cells.

Last, a response curve versus concentration (MFI_i vs. c_i) was fit in order to calculate the EC_{50} values using the MasterPlex 2010software, 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 curve adjustment, allowing two fitting curves and optionally interpolating unknown values of the curve. The ReaderFit contains several equations for the model: 4 parameters logistic (4PL), 5 parameters logistic (5PL), quadratic log-logit, log–log or linear and one of 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. A, B, C, D, and E are the parameters of the 5PL-model. The parameter 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 and E is the asymmetry factor ($E \neq 1$ for a non-symmetric curve). MFI values are obtained after exposing the biological sample to one volume of 100 μL of G1 at different c_i values. This equation is represented with a sigmoid curve:

$$MFI_i = A + \frac{D}{(1 + ((X/C)^B)^E} = (MFI_i)_{\min} + \frac{(MFI_i)_{\max}}{(1 + ((\frac{c_i}{EC_{50}})^B))^E} \quad (8)$$

$$\left(\frac{1}{MFI_i^2}\right) = A + \frac{D}{(1 + ((X/C)^B))^E} = \left(\frac{1}{MFI_i^2}\right)_{\min} + \frac{\left(\frac{1}{MFI_i^2}\right)_{\max}}{(1 + ((\frac{c_i}{EC_{50}})^B))^E} \quad (9)$$

3.2.5. 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. Statistical significances were accepted when $P<0.05$. The Tukey test with 95% confidence was applied to compare these means.[22]

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2012.07.020.

References and notes

1. Prado-Prado, F. J.; García-Mera, X.; González-Díaz, H. *Bioorg. Med. Chem.* **2010**, *18*, 2225.
2. González-Díaz, H.; Bonet, I.; Terán, C.; De Clercq, E.; Bello, R.; García, M. M.; Santana, L.; Uriarte, E. *Eur. J. Med. Chem.* **2007**, *42*, 580.
3. González-Díaz, H.; Prado-Prado, F.; Sobarzo-Sánchez, E.; Haddad, M.; Chevalley, S. M.; Valentín, A.; Quetin-Leclercq, J.; Dea-Ayuela, M. A.; Teresa Gómez-Muñoz, M.; Munteanu, C. R.; Torres-Labandeira, J. J.; García-Mera, X.; Tapia, R. A.; Ubeira, F. M. *J. Theor. Biol.* **2011**, *276*, 229.
4. Concu, R.; Dea-Ayuela, M. A.; Pérez-Montoto, L. G.; Prado-Prado, F. J.; Uriarte, E.; Bolas-Fernández, F.; Podda, G.; Pazos, A.; Munteanu, C. R.; Ubeira, F. M.; González-Díaz, H. *Biochim. Biophys. Acta, Proteins Proteomics* **2009**, *1794*, 1784.
5. Barral, P.; Sánchez-Niño, M. D.; van Rooijen, N.; Cerundolo, V.; Batista, F. D. *EMBO J.* **2012**, *31*, 3029.
6. Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. *Nucleic Acids Res.* **2012**, *40*, D1100.
7. Mok, N. Y.; Brenk, R. *J. Chem. Inform. Model.* **2011**, *51*, 2449.
8. Kuzmin, V. E.; Artemenko, A. G.; Gorb, L.; Qasim, M.; Leszczynski, J. *Chemosphere* **2008**, *72*, 1373.
9. Riera-Fernández, I.; Martín-Romalde, R.; Prado-Prado, F. J.; Escobar, M.; Munteanu, C. R.; Concu, R.; Duardo-Sánchez, A.; González-Díaz, H. *Curr. Top. Med. Chem.* **2012**, *12*, 927.
10. Prado-Prado, F.; García-Mera, X.; Escobar, M.; Sobarzo-Sánchez, E.; Yañez, M.; Riera-Fernández, P.; González-Díaz, H. *Eur. J. Med. Chem.* **2011**, *46*, 5838.
11. Kyung-Ran, K.; Son, E.-W.; Rhee, D.-K.; Pyo, S. *Toxicol. In Vitro* **2002**, *16*, 517.
12. Mingoia, R. T.; Nabb, D. L.; Yang, C.-H.; Han, X. *Toxicol. In Vitro* **2007**, *21*, 165.
13. Weyermann, J. L. D.; Zimmer, A. *Int. J. Pharm.* **2005**, *288*, 369.
14. Hamann, J.; Koning, N.; Pouwels, W.; Ulfman, L. H.; van Eijk, M.; Stacey, M.; Lin, H. H.; Gordon, S.; Kwakkenbos, M. J. *Eur. J. Immunol.* **2007**, *37*, 2797.
15. Koay, D. C.; Sartorelli, A. C. *Blood* **1999**, *93*, 3774.
16. Fischer, M. B.; Ruger, B.; Vaculik, C.; Becherer, A.; Wadsak, W.; Yanagida, G.; Losert, U. M.; Chen, J.; Carroll, M. C.; Eibl, M. M. *Eur. J. Immunol.* **2007**, *37*, 2825.
17. Alatery, A.; Basta, S. *J. Immunol. Methods* **2008**, *338*, 47.
18. Marmey, B.; Boix, C.; Barbaroux, J. B.; Dieu-Nosjean, M. C.; Diebold, J.; Audouin, J.; Fridman, W. H.; Mueller, C. G.; Molina, T. J. *Hum. Pathol.* **2006**, *37*, 68.
19. Den Haan, J.M.; Kraal, G. *J. Innate Immun.* in press.
20. Taylor, P. R.; Martínez-Pomares, L.; Stacey, M.; Lin, H. H.; Brown, G. D.; Gordon, S. *Annu. Rev. Immunol.* **2005**, *23*, 901.
21. Gerets, H. H.; Dhalluin, S.; Atienzar, F. A. *Methods Mol. Biol.* **2011**, *740*, 91.
22. Hill, T.; Lewicki, P. STATISTICS Methods and Applications. A Comprehensive Reference for Science, Industry and Data Mining; Tulsa: StatSoft, 2006.
23. Patankar, S. J.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 885.
24. García-García, A.; Gálvez, J.; de Julián-Ortiz, J. V.; García-Domenech, R.; Muñoz, C.; Guna, R.; Borras, R. *J. Antimicrob. Chemother.* **2004**, *53*, 65.
25. Marrero-Ponce, Y.; Castillo-Garit, J. A.; Olazabal, E.; Serrano, H. S.; Morales, A.; Castañedo, N.; Ibarra-Velarde, F.; Huesca-Guillen, A.; Sánchez, A. M.; Torrens, F.; Castro, E. A. *Bioorg. Med. Chem.* **2005**, *13*, 1005.
26. Marrero-Ponce, Y.; Machado-Tugores, Y.; Pereira, D. M.; Escario, J. A.; Barrio, A. G.; Nogal-Ruiz, J. J.; Ochoa, C.; Arán, V. J.; Martínez-Fernández, A. R.; Sánchez, R. N.; Montero-Torres, A.; Torrens, F.; Meneses-Marcel, A. *Curr. Drug Discov. Technol.* **2005**, *2*, 245.
27. Casanola-Martín, G. M.; Marrero-Ponce, Y.; Khan, M. T.; Ather, A.; Sultán, S.; Torrens, F.; Rotondo, R. *Bioorg. Med. Chem.* **2007**, *15*, 1483.
28. Casanola-Martín, G. M.; Marrero-Ponce, Y.; Tareq Hassan Khan, M.; Torrens, F.; Pérez-Jiménez, F.; Rescigno, A. *J. Biomol. Screen.* **2008**, *13*, 1014.
29. Casanola-Martín, G. M.; Marrero-Ponce, Y.; Khan, M. T.; Khan, S. B.; Torrens, F.; Pérez-Jiménez, F.; Rescigno, A.; Abad, C. *Chem. Biol. Drug Des.* **2010**, *76*, 538.
30. Rodríguez-Soca, Y.; Munteanu, C. R.; Dorado, J.; Pazos, A.; Prado-Prado, F. J.; González-Díaz, H. J. *Proteome Res.* **2010**, *9*, 1182.
31. González-Díaz, H.; Muino, L.; Anadón, A. M.; Romaris, F.; Prado-Prado, F. J.; Munteanu, C. R.; Dorado, J.; Pazos Sierra, A.; Mezo, M.; González-Warleta, M.; Garate, T.; Ubeira, F. M. *Mol. BioSyst.* **1938**, *2011*, 7.
32. Martínez-Romero, M.; Vázquez-Naya, J. M.; Rabunal, J. R.; Pita-Fernández, S.; Macenlle, R.; Castro-Alvarino, J.; López-Roses, L.; Ulla, J. L.; Martínez-Calvo, A. V.; Vázquez, S.; Pereira, J.; Porto-Pazos, A. B.; Dorado, J.; Pazos, A.; Munteanu, C. R. *Curr. Drug Metab.* **2010**, *11*, 347.
33. Estrada, E.; Molina, E.; Uriarte, E. *SAR QSAR Environ. Res.* **2001**, *12*, 445.
34. Estrada, E.; Uriarte, E. *SAR QSAR Environ. Res.* **2001**, *12*, 309.
35. Estrada, E.; Vilar, S.; Uriarte, E.; Gutiérrez, Y. J. *Chem. Inf. Comput. Sci.* **2002**, *42*, 1194.
36. Estrada, E.; González-Díaz, H. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 75.
37. Estrada, E.; Uriarte, E.; Gutiérrez, Y.; González-Díaz, H. *SAR QSAR Environ. Res.* **2003**, *14*, 145.
38. Meinel, T.; Schweiger, M. R.; Ludewig, A. H.; Chenna, R.; Krobitsch, S.; Herwig, R. *BMC Genomics* **2011**, *12*, 483.
39. StatSoft.Inc., 2002.
40. Blondeau, J. M.; Castañedo, N.; González, O.; Mendina, R.; Silveira, E. *Int. J. Antimicrob. Agents* **1999**, *11*, 163.
41. Marrero-Ponce, Y.; Garit, J. A.; Olazabal, E.; Serrano, H. S.; Morales, A.; Castañedo, N.; Ibarra-Velarde, F.; Huesca-Guillen, A.; Sánchez, A. M.; Torrens, F.; Castro, E. A. *Bioorg. Med. Chem.* **2005**, *13*, 1005.
42. Pérez Machado, G.; González-Borroto, J. I.; Castañedo, N.; Creus, A.; Marcos, R. *Mutagenesis* **2004**, *19*, 75.
43. Cao LF, K. L.; Tran, V.; Mi, S.; Jensen, M. C.; Blanchard, S.; Kalos, M. *Cytometry A.* **2010**, *77*, 534.
44. Riss, T. L.; Moravec, R. A. *Assay Drug Dev. Technol.* **2004**, *2*, 51.
45. Prowse, A. B.; Wilson, J.; Osborne, G. W.; Gray, P. P.; Wolvetang, E. J. *Stem. Cells Dev.* **2009**, *18*, 1135.
46. Miret, J. J.; Zhang, J.; Min, H.; Lewis, K.; Roth, M.; Charlton, M.; Bauer, P. H. *J. Biomol. Screen* **2005**, *10*, 780.
47. Jacques, N.; Vimond, N.; Conforti, R.; Griscelli, F.; Lecluse, Y.; Laplanche, A.; Malka, D.; Vielh, P.; Farace, F. *J. Immunol. Methods* **2008**, *337*, 132.
48. González-Borroto, J. I.; Pérez-Machado, G.; Creus, A.; Marcos, R. *Mutagenesis* **2005**, *20*, 193.
49. Ingels, F. M.; Augustijns, P. F. *J. Pharm. Sci.* **2003**, *92*, 1545.
50. Coggins, C. R.; Ballantyne, M.; Curvall, M.; Rutqvist, L. E. *Crit. Rev. Toxicol.* **2012**, *42*, 304.
51. OECD, Guidance document on using cytotoxicity tests to estimate starting doses for acute oral systemic toxicity tests. OECD Series on Testing and Assessment 2010, 20, 1.
52. Watson, D. A.; Brown, L. O.; Gaskill, D. F.; Naivar, M.; Graves, S. W.; Doorn, S. K.; Nolan, J. P. *Cytometry A* **2008**, *73*, 119.
53. Rasouli, A.; Sun, C. H.; Basu, R.; Wong, B. J. *Lasers Surg. Med.* **2003**, *32*, 3.
54. Petrunkina, A. M.; Harrison, R. A. *J. Immunol. Methods* **2011**, *368*, 71.
55. McGowan, P.; Nelles, N.; Wimmer, J.; Williams, D.; Wen, J.; Li, M.; Ewton, A.; Curry, C.; Zu, Y.; Sheehan, A.; Chang, C. C. *Am. J. Clin. Pathol.* **2012**, *137*, 665.
56. Gorczyca, W.; Sun, Z. Y.; Cronin, W.; Li, X.; Mau, S.; Tugulea, S. *Methods Cell Biol.* **2011**, *103*, 221.
57. Manivannan, E.; Prasanna, S. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4496.
58. Shin, K. J.; Wall, E.; Zavzavadjian, J. R.; Santat, L. A.; Liu, J.; Hwang, J. I.; Rebres, R.; Roach, T.; Seaman, W.; Simon, M. I.; Fraser, I. D. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13759.
59. Wert, L.; Alakurtti, S.; Corral, M. J.; Sánchez-Fortún, S.; Yli-Kauhaluoma, J.; Alunda, J. M. *J. Antibiot. (Tokyo)* **2011**, *64*, 475.
60. Lee, C. Y.; Wey, S. P.; Liao, M. H.; Hsu, W. L.; Wu, H. Y.; Jan, T. R. *Int. Immunopharmacol.* **2008**, *8*, 732.
61. Yamamoto, K.; Itoh, M.; Okamura, T.; Kimura, M.; Yokoyama, A.; Yoshino, Y.; Makino, M.; Hayakawa, N.; Suzuki, A. *Thyroid* **2012**, *22*, 516.
62. Cole, A. M.; Waring, A. J. *Am. J. Respir. Med.* **2002**, *1*, 249.
63. Ganz, T. *J. Leukoc. Biol.* **2004**, *75*, 34.
64. Lim, S.; Choi, S. H.; Shin, H.; Cho, B. J.; Park, H. S.; Ahn, B. Y.; Kang, S. M.; Yoon, J. W.; Jang, H. C.; Kim, Y. B.; Park, K. S. *PLoS ONE* **2012**, *7*, e35007.
65. Speck-Planche, A.; Kleandrova, V. V.; Luan, F.; Cordeiro, M. N. D. S. *Bioorg. Med. Chem.* **2011**, *19*, 6239.
66. SENACICA (NOM), Especificaciones técnicas para la producción, cuidado y uso de los animales de laboratorio. ZOO, 1999, 1–58.
67. King, M. A. *J. Immunol. Methods* **2000**, *243*, 155.
68. Tario, J. D.; Muirhead, K. A.; Pan, D.; Munson, M. E.; Wallace, P. K. K. *Flow Cytometry Protocols Methods Mol. Biol.* **2011**, *699*, 119.