

Experimental-Theoretic Approach to Drug-Lymphocyte Interactome Networks with Flow Cytometry and Spectral Moments Perturbation Theory

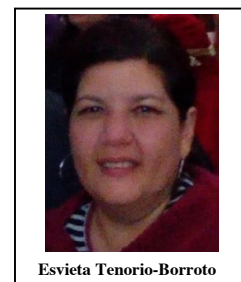
Esvieta Tenorio-Borroto^{a,b}, Fabiola Rivera Ramírez^c, Juan Carlos Vázquez Chagoyán^a, Roberto Montes de Oca Jiménez^a, Xerardo García-Mera^b and Humberto González-Díaz^{d,e,*}

^aCenter for Research and Advanced Studies in Animal Health, Faculty of Veterinary Medicines and Animal Husbandry, Autonomous University of Mexico State (UAEM), Toluca, 50200, Estado de México, México; ^bDepartment of Organic Chemistry, University of Santiago de Compostela (USC), 15782, Santiago de Compostela, Spain; ^cUniversidad Tecnológica del Valle de Toluca 52050 Santa maria de Atarrasquillo Lerma, México; ^dDepartment of Organic Chemistry II, University of Basque Country UPV/EHU, 48940, Leioa, Spain; ^eIKERBASQUE, Basque Foundation for Science, 48011, Bilbao, Spain

ARTICLE HISTORY

Received: March 10, 2016
Accepted: August 5, 2016

Abstract: We can combine experimental techniques like Flow Cytometry Analysis (FCA) with Chemoinformatics methods to predict the complex networks of interactions between organic compounds and targets in the immune system. In this work, we determined experimentally the values of $EC_{50} = 17.82 \mu\text{g/mL}$ and Cytotoxicity = 20.6 % for the antimicrobial / anti-parasite drug Dermofural over Balb/C CD9 lymphocytes using flow cytometry. After that, we developed a new Perturbation-theory model for Drug-Cell Target Interactome in Lymphocytes based on dispersion-polarization moments of drug structure. The models correctly classifies 34591 out of 42715 (Accuracy = 80.9%) cases of perturbations in assay endpoints of 11492 drugs (including both train and validation series). Each endpoint correspond to one out of 2616 assays, 38 molecular and cellular targets, 77 standard type measures, in four possible (human and rodents).



Keywords: Electronic charge, Atomic Polarizability, Dipole Moment, Spectral moments, Perturbation theory, Complex networks, ChEMBL, Drug immunotoxicity, Flow cytometry, CD19 Lymphocytes.

INTRODUCTION

We can combine multi-target networks models, perturbations theory, and flow cytometry to carry out a more complete study of the positive effects and/or cytotoxicity of drugs over different cell populations of the immune system. Specifically, Lymphocytes are also one of the more important cellular lineages of the immune system and play a central role in the immune response [1]. The main objective of the present work is to develop a valid High-throughput mt-QSAR model for predicting the immunotoxic effect of drugs over lymphocytes in a large set of m_j assay conditions. Another important goal is to illustrate the use of the new method in a real-life example. For it, we are going to download and calculate TOPS-MODE selected descriptors for the large dataset reported in ChEMBL. Next, we shall fit and validate a new High-throughput mt-QSAR Linear Discriminant Analysis (LDA) model using the software STATISTICA. After that, we report, by the first time, the experimental study of the effect of the drug Dermofural over Balb/C mouse thymic lymphocytes T helpers (CD4+) and T cytotoxic (CD8+) population using flow cytometry. Finally, we conduct the prediction of other endpoints multiplexing assay for Dermofural, not determined experimentally in this work. Also predicts the activity of some proteins that act in the immune response.

REVIEW OF SPECTRAL MOMENTS ANALYSIS FOR MULTI-OUTPUT QSPR MODELING

We can use different type of moments to quantify the structure of molecules in order to predict their interaction with complex

soft-matter systems (proteins, cells, *etc.*) present in living organisms. The spectral moments (μ_k) of the atoms adjacency matrix \mathbf{A} are very important for the definition of graph-spectrum-based molecular structural descriptors. In this context, the spectrum of the graph G is the set of the eigenvalues λ_i of \mathbf{A} . The formula for the calculation of μ_k is the following:

$$\mu_k = \text{Tr}(\mathbf{A}^k) = \sum_{i=0}^n \lambda_i^k \quad (1)$$

Specifically, Babić, Graovac, and Gutman [2] reported a relationship for a descriptor called the energy $E(G)$ of a graph. $E(G)$ is based on the new concept of the resonance energy in conjugated hydrocarbons introduced by Jiang and Zhang [3]. The Jiang and Zhang [3] model involves an expansion of the π -electron energy in terms of spectral moments μ_k , which are equal to numbers of closed walks in a molecular graph [4]. Babić, Graovac, and Gutman defined $E(G)$ as follows:

$$E(G) = \sum_{k=0}^n \lambda_i \quad (2)$$

Estrada *et al.* introduced the method TOSS-MODE [5, 6], and later renamed it as the TOPS-MODE method [7-9]. Both versions are useful to quantify the structure of molecules using the spectral moments of the bond (edge) adjacency matrix $\mathbf{B}(w_t)$ [5-9]. Estrada and Peña [10, 11] used this method to detect critical fragments for sedative/hypnotic compounds present in the Merck Index. Both $\mathbf{B}(w_t)$ and $\mu_k(w_t)$ depends on bond weighting functions w_t , which in turn depend on atomic ${}^a w_t$ or ${}^b w_t$ bond physicochemical properties of type t . Atomic mass and bond standard distances are examples of different types of properties [7-9].

*Address correspondence to this author at the Department of Organic Chemistry II, University of the Basque Country (UPV/EHU), 48940, Bilbao, Spain; E-mail: humberto.gonzalezdiaz@ehu.es

$$\mu_k(w_t) = \text{Tr} [\mathbf{B}(w_t)^k] \quad (3)$$

In any case, almost all predictive studies based on numerical molecular descriptors of the chemical structure of organic compounds (D_i) are unable to perform multi-target predictions. In fact, classic predictive methods for molecule-target interaction use only to predict the biological effect of single molecules over one target or system [12, 13]. This situation applies also to almost all previous models based on spectral moments like μ_k values.

However, new multi-target methods can predict properties of drugs, macromolecular structures, and macroscopic entities like populations taking into consideration changes in characteristics of the system. In doing so, González-Díaz et al. have used Box-Jenkins moving average operators to describe multi-target situations [14-17]. These operators are the molecular version of the classic Moving Average (MA) component of the Box-Jenkins ARIMA models in time series [18]. We see these characteristics of the system as sets of physical and/or bio-molecular boundary conditions c_j . Some of these boundary conditions may be c_1 = drug bio-molecular targets, c_2 = species of organisms, c_3 = specific assay protocols, c_4 = temperature or time ranges, c_5 = solvents, etc. We calculated the Box-Jenkins operators as the deviation of the molecular structure descriptor D_i of one compound from the average value $\langle D_i(c_j) \rangle$. This is the average value for the structural descriptors of a sub-set of molecules measured under the same condition c_j .

$$\Delta D_k = D_k - \langle D_k(c_j) \rangle \quad (4)$$

The method has a high potential for the study of complex bio-molecular systems in the frontiers between Biophysics and Biochemistry. One interesting case appears when the molecular descriptors D_k of type k^{th} for the i^{th} drug are the spectral moments of the adjacency matrix $D_k = \mu_k$. Very recently, Luan *et al.* [19] used Box-Jenkins operators of spectral moments to predict the interaction of chemical compounds with 148 possible molecular or cellular targets in 11 model organisms (including human). Cytotoxicity assays are screening methods that uses different permanent cell lines for ranking acute toxicities of parent compounds [20, 21]. We can use different types of molecular parameters to predict the interaction of drugs with different targets in immune system. Tenorio *et al.* [22] also used spectral moments to find a multi-target model for drug immunotoxicity over 36 molecular or cellular targets in macrophages. In this particular case of spectral moments, we can write MA operators as follow:

$$\Delta \mu_k(w_t) = \mu_k(w_t) - \langle \mu_k(w_t) \rangle_j \quad (5)$$

EXPERIMENTAL STUDY OF THE EFFECT DERMOFURAL OVER CD19 LYMPHOCYTE TARGETS

The measurement of cell viability is an important goal in cytotoxicity studies [23-26]. In our study, we examined the populations of CD19 lymphocytes using flow cytometry. The analyses with flow cytometry were performed in order to follow the percentage of live lymphocytes present in the lymphocytes populations treated with Dermofural at different concentrations we observed changes in the viability of the lymphocytes after 24 hours. The assay shows a significant increase of dead cells, Cytotoxicity (%) = 20.6%, compared to the group untreated (1.5 %) and the DMSO group (3.1%) at $c_{\text{max}} = 10 \mu\text{g/mL}$. The treatment of 6 and 8 $\mu\text{g/mL}$ results in a dose-dependent significant increase in cytotoxicity (14.5%) and (17.4%) respectively (Fig. 1).

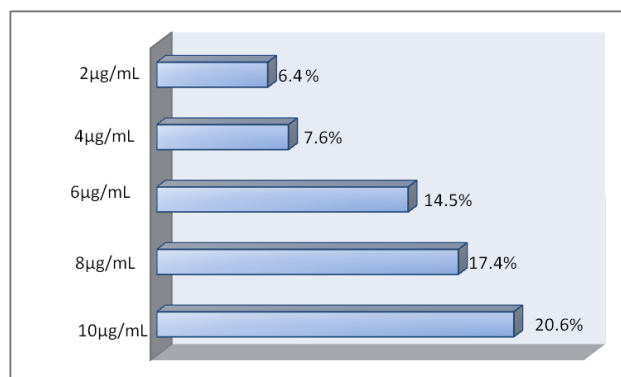


Fig. (1). Percentage of cytotoxicity in Balb/C mouse thymic B-Lymphocytes marked with CD19bioAPC/7AAD exposed to different concentrations of Dermofural.

The percent of cytotoxicity is similar in concentrations 2 and 4 $\mu\text{g/mL}$ (approximately 7%). In studies [27] with these derivatives have worked with other ranges of cytotoxicity (1.25 to 30 $\mu\text{g/mL}$). These results indicate that has low cytotoxicity Dermofural at this concentration (10 $\mu\text{g/mL}$) for cytotoxicity (%) < 50% is smaller than a threshold value is considered to cytotoxic compounds [28]. In general, there is no significant difference between different concentrations of DMSO group Dermofural. In Table 1, we show the results of statistical analysis of the parameters obtained of the flow cytometer, using the values for p and significant differences for $p < 0.05$ between different concentrations and control groups (Negative Control, DMSO), and then the average values of MFI in SSC and/or FSC scattering mode, for all samples. We confirm that there were significant differences for treated samples of living B lymphocytes with respect to control groups.

We also investigated the MFI on highly homogenous lymphocytes populations defined by the expression of CD19bioAPC obtained from the thymus of healthy mouse. In Fig. (2), we depict results for thymus B Lymphocytes of Balb/ C mouse analysis exposed to compound Dermofural at 10 $\mu\text{g/mL}$. This corroborates that the B-lineage cells in the thymus account for less than 1% of total lymphocytes.

PT-QSPR MODEL FOR DRUG-LYMPHOCYTE TARGET INTERACTOME

We introduced one special case of QSPR-perturbation theory model combining the spectral moments of a molecular graph $\mu_k(Q_t)$ and the approach to multiple input-output perturbations in bio-molecular systems reported by Gonzalez-Díaz *et al.* [29, 30]. Please, see details about the method in these previous works. Here we used this previous method to develop the first QSTR-Perturbation model to predict interactions of organic compounds with molecular and cellular targets on lymphocyte cells (see Fig. 3). We downloaded the database from ChEML <https://www.ebi.ac.uk/chembl/db> [31] with >20,000 assays of drugs related somehow to lymphocytes. The best QSPR-Perturbation model of first order found here with LDA was:

$$S(\epsilon_{ij})_{\text{new}} = 0.1578 \cdot f(\epsilon_{ij})_{\text{ref}} + 33.5734 \cdot \Delta\mu(Q_0)_{\lambda_0} + 1.5541 \cdot \Delta\mu(Q_0)_{\lambda_1} - 33.4696 \cdot \Delta\mu(Q_0)_{\lambda_2} \quad (6)$$

$$- 34.4871 \cdot \Delta\mu(Q_1)_{\lambda_0} + 33.5374 \cdot \Delta\mu(Q_1)_{\lambda_2} + 33.5374 \cdot \Delta\Delta\mu(D)_{\lambda_2}$$

$$- 2.0914 \cdot \Delta\mu(Q_2)_{\lambda_1} + 1.4713 \cdot \Delta\mu(Q_3)_{\lambda_1} - 0.0693$$

$$N = 42715 \quad F = 2470.742 \quad U = 0.15 \quad p < 0.005$$

The first input term $f(\epsilon_{ij})_{\text{ref}}$ is the scoring function f of the efficiency of the known drug-target interaction process ϵ_{ij} (known solution). The term $f(\epsilon_{ij})_{\text{ref}} = 1$ if we already know from previous

Table 1. Results of statistical analysis of the parameters of the flow cytometer showed significant differences between different concentrations and control groups.

Group1	Parameters	Group2		MFI
		NC	DMSO	
Conc (µg/mL)	SCC-H+ FSC-H+	p	p	
	SCC-H+ CD3FITC+			
10	SCC-H+ 7-AAD+	0.819773	0.967863	233.2±38.38
	SCC-H+ CD19bio-APC+	0.208782	0.508540	196.28±31.79
	CD3FITC+ 7AAD+	0.130045	0.919263	105.5±17.82
	CD3FITC+CD19bioAPC+	0.083138	0.837216	37.77±37.77
	CD19bioAPC+7AAD+	0.061232	0.532019	114.96±38.58
	CD19bioAPC+7AAD-	0.035743 ^a	0.079477	59.18±21.09
8	SCC-H+ FSC-H+	0.574506	0.618669	480.7±43.34
	SCC-H+ CD3FITC+	0.009396 ^a	0.266945	165.3±26.80
	SCC-H+ 7-AAD+	0.804872	0.962317	207.7±35.39
	SCC-H+ CD19bio-APC+	0.080902	0.237483	197.6±122.9
	CD3FITC+ 7AAD+	0.046746 ^a	0.633704	110.1±18.43
	CD3FITC+CD19bioAPC+	0.007041 ^a	0.172750	46.01±46.01
6	CD19bioAPC+7AAD+	0.174550	0.835663	109.44±65.7
	CD19bioAPC+7AAD-	0.126061	0.256566	48.68±10.73
	SCC-H+ FSC-H+	0.271318	0.306663	358.43±31.67
	SCC-H+ CD3FITC+	0.023197 ^a	0.660648	212.41±15.51
	SCC-H+ 7-AAD+	0.816395	0.976136	151.2±22.92
	SCC-H+ CD19bio-APC+	0.100570	0.323846	131.83±18.23
4	CD3FITC+ 7AAD+	0.086465	0.917597	76.42±11.36
	CD3FITC+CD19bioAPC+	0.023535 ^a	0.583148	31.85±31.85
	CD19bioAPC+7AAD+	0.030059 ^a	0.418862	80.70±13.96
	CD19bioAPC+7AAD-	0.017667 ^a	0.044506 ^a	34.94±7.87
	SCC-H+ FSC-H+	0.421978	0.468035	362.12±30.54
	SCC-H+ CD3FITC+	0.024891 ^a	0.683118	121.69±16.33
4	SCC-H+ 7-AAD+	0.894592	0.993257	164.9±22.86
	SCC-H+ CD19bio-APC+	0.755928	0.988748	91.53±11.82
	CD3FITC+ 7AAD+	0.032853 ^a	0.656848	68.41±10.94
	CD3FITC+CD19bioAPC+	0.049169 ^a	0.810848	28.13±28.13
	CD19bioAPC+7AAD+	0.034993 ^a	0.463901	63.18±22.05
	CD19bioAPC+7AAD-	0.016652 ^a	0.041979 ^a	31.54±7.23
4	SCC-H+ FSC-H+	0.626131	0.675905	361.65±34.73
	SCC-H+ CD3FITC+	0.021791 ^a	0.640636	125.03±18.13

^aSignificant differences. NC is control negative and DMSO is control vehicle.

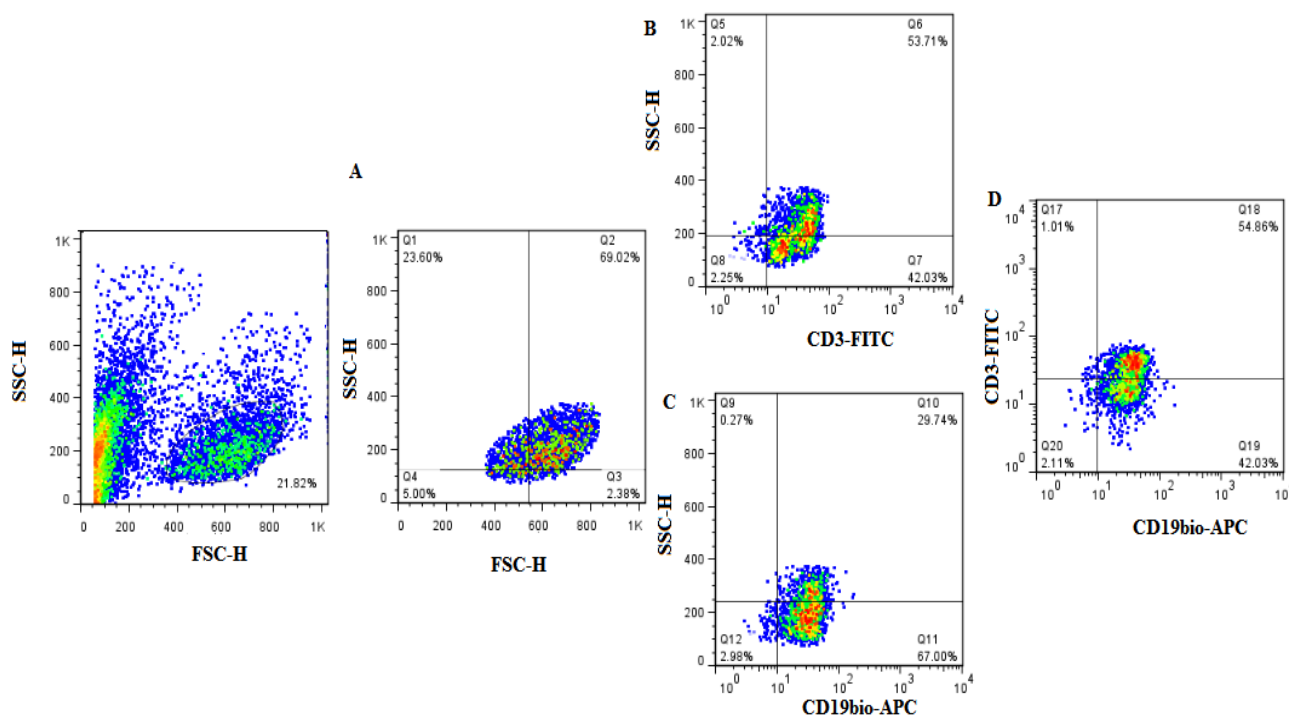


Fig. (2). Graphic type Pseudo-color of gating strategy for thymus B Lymphocytes of Balb/ C mouse analysis exposed to compound Dermofural at 10µg/mL. Example Total population Cell SSC-H & FSC-H (A) lymphocytes marked with CD3-FITC (B) CD19bio-APC (C) and Both (D).

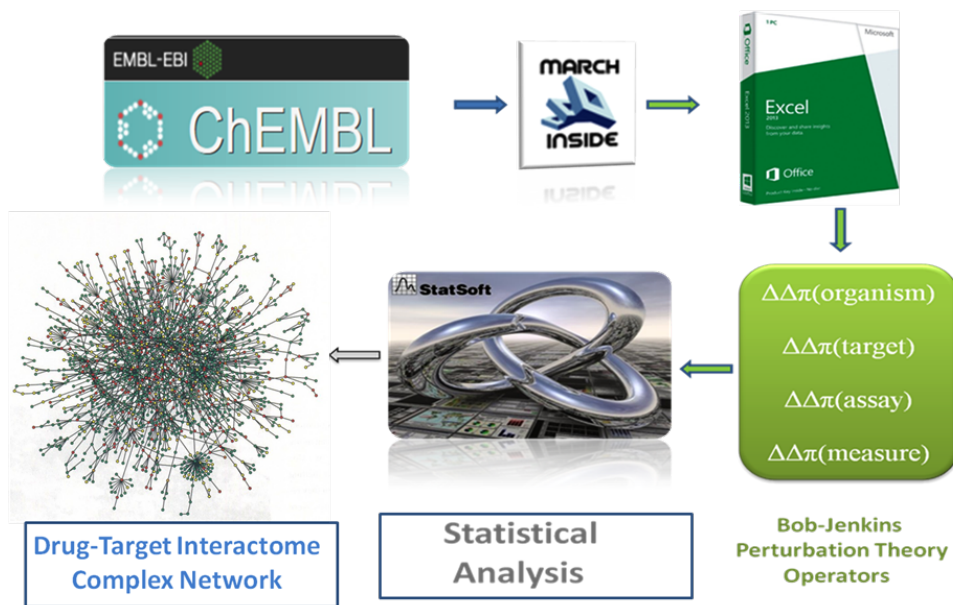


Fig. (3). Workflow for multi-target prediction of drug-target interactome networks.

experiments that the compound i^{th} interact with the cellular target in the boundary conditions c_j , $f(\epsilon_{ij})_{ref} = 0$ otherwise. The atomic-bond polarization properties are the $Q_0 =$ atomic charge (q), $Q_1 =$ Bond dipole moment (D), and $Q_3 =$ atomic polarizability (α). The conditions of assay are the following $c_0 =$ chemical structure, $c_1 =$ biological property, $c_2 =$ organism, and $c_3 =$ molecular or cellular target.

In Table 2, we depict the classification matrix and the values of A_c , S_p , and S_n for training and validation series. This values are

very good taking into consideration the results published for other problems with similar methodologies based on average models [32, 33] or perturbation models [29].

CONCLUSION

QSPR-perturbation theory using spectral moments have a high potential to develop different methods for the study of complex bio-systems. Develop general models based on spectral moments to predict the effect of multiple input-output perturbations in complex systems it is possible. This new QSPR-Perturbation model can be

Table 2. Results of QSPR-Perturbation model of drug-target interactome developed in this work.

Order	Observed	Predicted	Values	Predicted Sets		
n	sets	Statistics	%	n	0	1
1	0	Sp	80.0	24814	19861	4953
	1	Sn	82.3	17901	3171	14730
	Train (t)	Ac	81.0	42715		
	0	Sp	80.3	8233	6608	1625
	1	Sn	82.8	6007	1034	4973
	Validation (v)	Ac	81.3	14240		
	Total t + v	Ac	81.1	56955		

^aThe positive (1) and negative control cases (0) were assigned as follows $L_{ij} = 1 \Rightarrow z_{ij} > 0 \Rightarrow v_j > \text{Avg}(v_j)$ otherwise $L_{ij} = 0$.

used to study properties like interaction of drugs with multiple targets or very different conditions, useful to predict changes on the ability of organic compounds to form complex interaction networks with molecular and cellular lymphocyte targets such as Dermofural.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers Web site along with the published article.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

The authors thank the sponsorship of grant DCS-UQROO PROFOCIE (P/PROFOCIE-2014-23MSU0140Z-09 DCS) and research project CTQ2013-41229-P funded by MINECO, Spain.

REFERENCES

- [1] Klein AB, Witonsky SG, Ahmed SA, *et al.* Impact of different cell isolation techniques on lymphocyte viability and function. *J Immunoassay Immunochem* 2006; 27: 61-76.
- [2] Babic D, Graovac A, Gutman I. On a resonance energy model based on expansion in terms of acyclic moments: Exact results. *Theoretica chimica acta* 2013; 79: 403-411.
- [3] Y. J. H. Z. Stability and reactivities based on moment analysis. *Theoretica chimica acta* 2013; 75: 279-297.
- [4] Gutman I, Rucker C, Rucker G. On walks in molecular graphs. *J Chem Inf Comput Sci* 2001; 41: 739-45.
- [5] Estrada E. On the topological sub-structural molecular design (TOSS-MODE) in QSPR/QSAR and drug design research. *SAR QSAR Environ Res* 2000; 11: 55-73.
- [6] Estrada E, Gutierrez Y, González H. Modeling diamagnetic and magneto-optic properties of organic compounds with the TOSS-MODE approach. *J Chem Inf Comput Sci* 2000; 40: 1386-1399.
- [7] Estrada E, Uriarte E. Quantitative structure-toxicity relationships using TOPS-MODE. 1. Nitrobenzene toxicity to *Tetrahymena pyriformis*. *SAR QSAR Environ Res* 2001; 12: 309-24.
- [8] Estrada E, Molina E, Uriarte E. Quantitative structure-toxicity relationships using TOPS-MODE. 2. Neurotoxicity of a non-congeneric series of solvents. *SAR QSAR Environ Res* 2001; 12: 445-59.
- [9] Estrada E, Uriarte E, Gutierrez Y, Gonzalez H. Quantitative structure-toxicity relationships using TOPS-MODE. 3. Structural factors influencing the permeability of commercial solvents through living human skin. *SAR QSAR Environ Res* 2003; 14: 145-63.
- [10] Estrada E, Peña A, García-Domenech R. Designing sedative/hypnotic compounds from a novel substructural graph-theoretical approach. *J Comput Aided Mol Des* 1998; 12: 583-95.
- [11] Estrada E, Peña A. In silico studies for the rational discovery of anticonvulsant compounds. *Bioorg Med Chem* 2000; 8: 2755-70.
- [12] Gonzalez-Díaz H, Vilar S, Santana L, Uriarte E. Medicinal chemistry and bioinformatics - Current trends in drugs discovery with networks topological indices. *Current Topics in Medicinal Chemistry* 2007; 7: 1015-1029.
- [13] Gonzalez-Díaz H, Prado-Prado F, Ubeira FM. Predicting antimicrobial drugs and targets with the MARCH-INSIDE approach. *Curr Top Med Chem* 2008; 8: 1676-90.
- [14] Duardo-Sanchez A, Munteanu CR, Riera-Fernández P, López-Díaz A, Pazos A, González-Díaz H. Modelling complex metabolic reactions, ecological systems, and financial-legal networks with MI-ANN models based on Markov-Wiener centralities. *J Chemical Information and Modelling* 2013: submitted.
- [15] Gonzalez-Díaz H. Bioinformatics and Quantitative Structure-Property Relationship (QSPR) Models. *Current Bioinformatics* 2013; 8: 387-389.
- [16] Gonzalez-Díaz H, Riera-Fernandez P, Pazos A, Munteanu CR. The Rucker-Markov invariants of complex Bio-Systems: Applications in Parasitology and Neuroinformatics. *Biosystems* 2013; 111: 199-207.
- [17] Riera-Fernandez P, Munteanu CR, Martin-Romalde R, Duardo-Sanchez A, Gonzalez-Díaz H. Markov-Randic Indices for QSPR Re-Evaluation of Metabolic, Parasite-Host, Fasciolosis Spreading, Brain Cortex and Legal-Social Complex Networks. *Current Bioinformatics* 2013; 8: 401-415.
- [18] Box GEP, Jenkins GM. Time series analysis. Holden-Day 1970.
- [19] Luan F, Cordeiro MN, Alonso N, *et al.* TOPS-MODE model of multiplexing neuroprotective effects of drugs and experimental-theoretic study of new 1,3-rasagiline derivatives potentially useful in neurodegenerative diseases. *Bioorg Med Chem* 2013; 21: 1870-9.
- [20] Mingoia RT, ND, Ching-Hui Yang, Xing Han. 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. *Toxicology in Vitro* 2007; 21: 165-173.
- [21] Weyermann J. LD, Zimmer Andreas., A practical note on the use of cytotoxicity assays. *International J Pharmaceutics* 2005; 288: 369-376.
- [22] Tenorio-Borroto E, Penuelas Rivas CG, Vasquez Chagoyan JC, *et al.* ANN multiplexing model of drugs effect on macrophages; theoretical and flow cytometry study on the cytotoxicity of the antimicrobial drug G1 in spleen. *Bioorg Med Chem* 2012; 20: 6181-94.
- [23] Cao LF KL, Tran V, Mi S, Jensen MC, Blanchard S, Kalos M. Development and application of a multiplexable flow cytometry-based assay to quantify cell-mediated cytolysis. *Cytometry A* 2010; 77: 534-45.
- [24] Riss TL MR. 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* 2004; 2: 51-62.

- [25] Marqués-Gallego P, den Dulk H, Backendorf C, Brouwer J, Redijk J, Burke JF. Accurate non-invasive image-based cytotoxicity assays for cultured cells. *BMC Biotechnol* 2010; 10: 43.
- [26] Sekhon BK, Roubin RH, Tan A, Chan WK, Sze DM. High-throughput screening platform for anticancer therapeutic drug cytotoxicity. *Assay Drug Dev Technol* 2008; 6: 711-21.
- [27] Perez Machado Giselle. GBJI, Castañedo N, Creus A, Marcos R. In vitro genotoxicity testing of the furylethylene derivative UC-245 in human cells. *Mutagenesis* 2004; 19: 75-80.
- [28] OECD ED. Guidance document on using cytotoxicity tests to estimate starting doses for acute oral systemic toxicity tests. *OECD Series on Testing and Assessment* 2010; 20: 1-54.
- [29] Gonzalez-Diaz H, Arrasate S, Gomez-San Juan A, *et al.* New theory for multiple input-output perturbations in complex molecular systems. 1. Linear QSPR electronegativity models in physical, organic, and medicinal chemistry. *Curr Top Med Chem* 2013.
- [30] González-Díaz H, Pérez-Montoto LG, Ubeira FM. Model for vaccine design by prediction of B-epitopes of IEDB given perturbations in peptide sequence, in vivo process, experimental techniques, and source or host organisms. *Clin Develop Immunol* 2014; 2014: 1-20.
- [31] Gaulton A, Bellis LJ, Bento AP, *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research* 2012; 40: D1100-7.
- [32] Speck-Planche A, Kleandrova VV, Luan F, Cordeiro MN. In silico discovery and virtual screening of multi-target inhibitors for proteins in *Mycobacterium tuberculosis*. *Comb Chem High Throughput Screen* 2012; 15: 666-73.
- [33] Speck-Planche A, Kleandrova VV, Luan F, Cordeiro MN. Chemoinformatics in anti-cancer chemotherapy: multi-target QSAR model for the in silico discovery of anti-breast cancer agents. *Eur J Pharm Sci* 2012; 47: 273-9.