

Hypertension: A mt-QSAR Model for Seeking New Drugs for the Hypertension Treatment using Multiple Conditions.

Francisco Javier Prado Prado *, Jan-carlo Miguel Díaz-González, Victoria Anahi Noh-Mayo, Francisco Javier Aguirre-Crespo

Biomedical Sciences Department, Health Sciences Division, University of Quintana Roo, UQROO, 77039, Mexico.

Abstract. Hypertension is a multifactorial disease in which blood vessels are extensively exposed to a higher voltage than usual, this tension endures more strain on the heart leading to greater cardiac output to pump blood to the body. Hypertension is classified by the World Health Organization (WHO) as one of the main risk factors for disability and premature death in the world population. WHO has strengthened various health services around the world, listing the groups of basic medicines for high blood pressure such as: angiotensin-converting enzyme inhibitors, thiazide diuretics, beta blockers, long-acting calcium channel blockers, among other groups for drug treatment to the population with this condition. The discovery of new drugs with better activity and less toxicity for the treatment of Hypertension is a goal of the major importance. In this sense, theoretical models as QSAR can be useful to discover new drugs for hypertension treatment. For this reason, we developed a new multi-target-QSAR (mt-QSAR) model to discover new drugs. A public databases ChEMBL contain Big Data sets of multi-target assays of inhibitors of a group of receptors with special relevance in Hypertension was used. However, almost all the computational models known focus in only one target or receptor. In this work, Beta-2 adrenergic receptor, Adrenergic receptor beta, Type-1 angiotensin II receptor, Angiotensin-converting enzyme, Beta-adrenergic receptor, Cytochrome P450 11B2 and Renin were used as receptor inputs in the model. An ANN is our statistical analysis. In that way, we used as input Topological Indices, in specific Wiener, Barabasi and Harary indices calculated by Dragon software. These operators quantify the deviations of the structure of one drug from the expected values for all drugs assayed in different boundary conditions such as type of receptor, type of assay, type of target, target mapping. Overall training performance was 90%. Overall Validation predictability performance was 90%.

Keywords: Hypertension, mt-QSAR, Artificial Neural Network

***corresponding author:** PRADO PRADO, F; (fenoll@hotmail.com or fprado@uqroo.edu.mx), Biomedical Sciences Department, Health Sciences Division; University of Quintana Roo (UQROO).

1. INTRODUCTION

Hypertension is a multifactorial disease in which blood vessels are extensively exposed to a higher voltage than usual, this tension endures more strain on the heart leading to greater cardiac output to pump blood to the body. Hypertension is classified by the World Health Organization (WHO) as one of the main risk factors for disability and premature death in the world population. WHO has strengthened various health services around the world, listing the groups of basic medicines for high blood pressure such as: angiotensin-converting enzyme inhibitors, thiazide diuretics, beta blockers, long-acting calcium channel blockers, among other groups drug treatment (1, 2). As said before, the field in the research of better hypertension treatments is getting bigger as time passes. Bioinformatics is an area that has been widely used in these researches, making it as one of the most important fields. Basing on Bioinformatics, there exists a multi-target model (mt) where by a big and heterogeneous database of several compounds is designed. In this model, the compounds are categorized as being active or inactive. Form these molecules, a portion of them were obtain in order to calculate there was a positive relationship (3).

In this study, some proteins that were found to be related to hypertension and used in this mt-QSAR are Beta-2 adrenergic receptor, Adrenergic receptor beta, Type-1 angiotensin II receptor, Angiotensin-converting enzyme, Beta-adrenergic receptor, Cytochrome P450 11B2 and Renin were used as receptors inputs in the model. The design of new enzyme inhibitors for the treatment of hypertension creates a main objective. From our point of view, QSAR techniques may be very helpful in this case. Unfortunately, some QSAR techniques predict new

outcomes only for one specific condition. We can avoid this by developing a new Multi-target/Multiplexing QSAR models. These approaches are useful to process very large collections of compounds assayed against multiple molecular or cellular targets under different assay conditions (c_j) as is the case of ChEMBL (4, 5). QSAR models can foretell the results of the assay of different drugs for multiple targets. 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 multiple conditions. In this work, we defined c_j according to the ontology $r_t \Rightarrow (a_u, c_j, r_t, t_e, s_x)$. The different conditions that may change in the dataset are: receptors (r_t), biological assays (a_u), molecular or cellular targets (t_e), or standard type of activity measure (s_x). Notably, multi-target QSAR models are able to predict the results of the assay of different drugs for multiple targets. However, QSAR models cannot foretell diverse results for a given sequence of targets when changed under a set of definite assay, organism, targets and assay type conditions for each target. Fortunately, the new mt-QSAR is not only useful for different targets but also to different multiplexing assay conditions (c_j) for all targets. Definitely, we have stated the first QSAR model for multiplexing assays of anti-Alzheimer, anti-parasitic, anti-fungi, and anti-bacterial activity (6-15).

QSAR methods are related on the use of molecular descriptors, which are known as mathematical sequence that codify useful chemical data and enable associations between statistical and biological properties (16, 17). A diverse topological indices (TIs) of molecular graphs (G) can be used to speed up the procedure of codification of the molecular arrangement of drugs in Bioinformatics studies. By the first time we used TI molecular descriptors developed by Wiener, Barabasi and Harary indices to develop one multi-target/multiplexing QSAR model for inhibitors of 7 different enzymes relates to hypertension.

2. MATERIALS AND METHODS

2.1 Nonlinear classifiers

A database from ChEMBL database (18) containing assayed anti-hypertension drugs was used to conduct this experiment. The DRAGON software 4.0 (19) was utilized and provided 1664 descriptors classified as zero- (0D) one- (1D), two- (2D) and three-dimensional (3D) descriptors depending on the fact that they are computed from the chemical formula, substructure list representation, molecular graph or geometrical representation of the molecule, respectively (20). In this research, the following descriptors were calculated: Wiener, Barabasi and Harary indices. The data was processed with different Artificial Neural Networks (ANNs) using the STATISTICA 6.0 software (21) looking for the better model to predict anti-hypertension activity. Five types of ANNs were used, namely, Probabilistic Neural Network (PNN), Radial Basis Function (RBF), Three Layers Perceptron (MLP-3), and Four Layer Perceptron (MLP-4) and Linear (LNN) (22-27).

A very simple type of ANN called Four Layer Perceptron (MLP-4) can be used to fit this discriminant function. The model deals with the classification of a compound set with or without affinity of different receptors. A dummy variable Affinity Class (AC) was used as input to codify the affinity. This variable indicates either high (AC = 1) or low (AC = 0) affinity of the drug of the receptor. S(DTP)pred or DTP affinity predicted score is the output of the model and it is a continuous dimensionless score that sorts compounds from low to high affinity to the target coinciding DTPs with higher values of S(DTP)pred and nDTPs with lowest values. A Forward Stepwise algorithm was used for a variable selection (28).

Let $k\chi(G)$ be drugs molecular descriptors and $k\xi(R)$ receptor or drug target descriptors for different drugs (d) with different receptor; the group attempted to develop a simple linear classifier of mt-QSAR type with the general formula:

$$S(DTP)_{pred} = \sum_{k=0}^5 b_i (G_k)^k \chi(G) + b_0 \quad (1)$$

The quality of models was assessed with different statistical parameters like Specificity (see Equation 2), Sensitivity (see Equation 3), Accuracy (see Equation 4) and ROC curve (Receiver Operating Characteristic curve) which is a graphical plot of the sensitivity, or true positives, vs. (1-specificity), or false positives,

$$\text{specificity} = \frac{NTN}{NTN + NFP} \quad (2) \quad \text{sensitivity} = \frac{NTP}{NTP + NFN} \quad (3)$$

$$\text{accuracy} = \frac{NTP + NTN}{NTP + FN + FP + TN} \quad (4)$$

Where NTN means number of true negatives, NFP is the number of false positives, NTP is the number of true positives, NFN is the number of false negatives, FN is false negatives, FP is false positives and TN is true negatives.

The data set used in this article was obtained from ChEMBL database (29-33). It has more than 11000 cases and more than 6500 different compounds. ChEMBL normalizes the bioactivities into a uniform set of endpoints and units where possible, and also tags the links between a molecular target and a published assay with a set of varying confidence levels. The data is abstracted and curated from primary scientific literature, and covers a significant fraction of the structure activity relationship (SAR) and discovery of modern drugs.

3. RESULTS AND DISCUSSION

3.1. ANN Multi-target model of drug-hypertension receptor interaction.

The ANN models are non-linear models used to predict the biological activity of a large dataset of molecules. This technique is an alternative to linear methods such as LDA (34, 35). However, one must note that the profiles of each network indicate that these are highly nonlinear and complicated models. Different types of networks were compared to obtain a better model; Table 1 shows the classification matrix of the ANN network. MLP 42: 42-27-1:1 was taken as the main network because it presented a wider range of variables, 42 inputs in the first layer and 27 neurons in the second layer, and two sets of cases (Training and Validation). The network found was MLP and it showed a training performance higher than 89.9%. The best model correctly classifies 667 out of 794 active compounds (84.00 %) and 2971 out of 3251 non-active compounds (91.39 %) in the training series. Overall training performance was 89.94%. Validation of the model was carried out by means of external predicting series, the validation correctly classifies 6083 out of 6733 non-active compounds (90.34%) and 1325 out of 1527 active compounds (86.77%). Overall Validation predictability performance was 89.7%. The values of accuracy higher than 75% are acceptable for ANN models; according to previous reports (36-45).

Table 1. Results of the MLP classification model

Drug-Enzyme interaction	Training series			Statistical Parameter
	Positive	Negative	%	
Negative	280	2971	91.39	Sensitivity
Positive	667	127	84.01	Specificity
			89.94	Accuracy
Drug-Enzyme interaction	External validation series			Statistical Parameter
	Positive	Negative	%	
Negative	650	6083	90.34	Sensitivity
Positive	1325	202	86.77	Specificity
			89.7	Accuracy

^a Sensitivity = Recall = True Positive / (True Positive + False Negative);

Specificity = True Negative / (True Negative + False Positive);

Accuracy = Ac = True Total / Total = (True Positive + True Negative) / Total

Notably, the model presented had a ROC curve higher than 0.5. The model presented an area greater than 0.9, see Figure 1. One compound may lead to 1 or more statistical cases because it may give different outcomes for alternative biological assays carried out in different organisms with different enzymes as targets (46). We used a big data from ChEMBL database, only using anti-hypertension drugs and the model have good results. It is the

first work on hypertension in the mt-QSAR is used within the model using different enzymes, assays, and types of proteins.

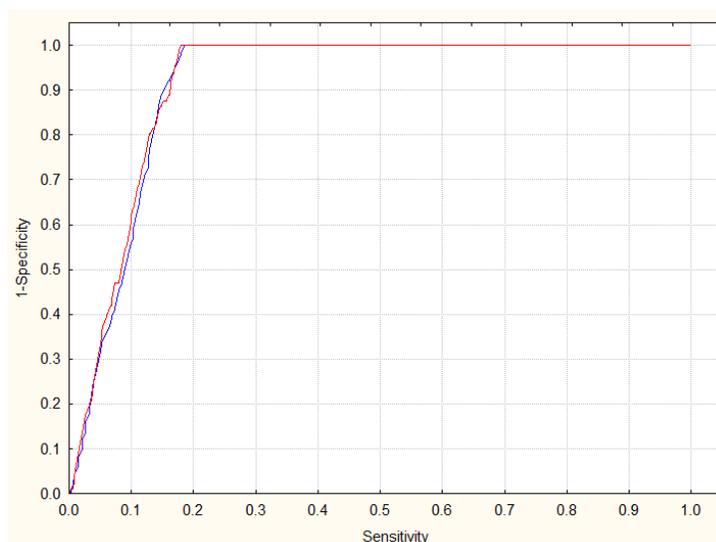


Figure 1. ROC Curve for MLP predictor (red = train series, blue = validation series)

4. CONCLUSION

Theoretic mt-QSAR models based on ANN and Topological Indices descriptors may become a useful tool in design and development of new drugs with anti-hypertension activity. Drug-Protein interaction is a goal of major importance to discover a cure of hypertension. In this work, we developed a new MLP model using the Dragon descriptors, with a large data base using about 6000 different drugs obtained from the ChEMBL database. The use of different assays, organisms, targets and assay type conditions for each target involved in the model is a big advantage for seek the treatment and possible cure of hypertension more effective and less toxic drugs.

5. ACKNOWLEDGMENTS

The authors thank the sponsorship DCS-UQROO PROFOCIE (**P/PROFOCIE-2014-23MSU0140Z-09 DCS**) and sponsorships for research project PRODEP-Fortalecimiento Cuerpos Académicos 2012.

REFERENCES

1. Whitworth JA. 2003 World Health Organization (WHO)/International Society of Hypertension (ISH) statement on management of hypertension. *Journal of hypertension*. [Guideline Practice Guideline]. 2003 Nov;21(11):1983-92.
2. Moser M. World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension-Do These Differ From the U.S. Recommendations? Which Guidelines Should the Practicing Physician Follow? *J Clin Hypertens (Greenwich)*. 1999 Jul;1(1):48-54.
3. Abifadel M, Pakradouni J, Collin M, Samson-Bouma ME, Varret M, Rabes JP, et al. Strategies for proprotein convertase subtilisin kexin 9 modulation: a perspective on recent patents. *Expert opinion on therapeutic patents*. [Research Support, Non-U.S. Gov't Review]. 2010 Nov;20(11):1547-71.
4. Riera-Fernandez I, Martin-Romalde R, Prado-Prado FJ, Escobar M, Munteanu CR, Concu R, et al. From QSAR models of Drugs to Complex Networks: State-of-Art Review and Introduction of New Markov-Spectral Moments Indices. *Curr Top Med Chem*. 2012 Feb 14.
5. Prado-Prado F, Garcia-Mera X, Escobar M, Sobarzo-Sanchez E, Yanez M, Riera-Fernandez P, et al. 2D MI-DRAGON: a new predictor for protein-ligands interactions and theoretic-experimental studies of US FDA drug-target network, oxoisoaporphine inhibitors for MAO-A and human parasite proteins. *Eur J Med Chem*. [Research Support, Non-U.S. Gov't]. 2011 Dec;46(12):5838-51.

6. Garcia I, Fall Y, Gomez G, Gonzalez-Diaz H. First computational chemistry multi-target model for anti-Alzheimer, anti-parasitic, anti-fungi, and anti-bacterial activity of GSK-3 inhibitors in vitro, in vivo, and in different cellular lines. *Molecular Diversity*. 2011 MAY 2011;15(2):561-7.
7. Gonzalez-Diaz H, Duardo-Sanchez A, Ubeira FM, Prado-Prado F, Perez-Montoto LG, Concu R, et al. Review of MARCH-INSIDE & complex networks prediction of drugs: ADMET, anti-parasite activity, metabolizing enzymes and cardiotoxicity proteome biomarkers. *Current drug metabolism*. [Research Support, Non-U.S. Gov't Review]. 2010 May;11(4):379-406.
8. Prado-Prado FJ, Garcia-Mera X, Gonzalez-Diaz H. Multi-target spectral moment QSAR versus ANN for antiparasitic drugs against different parasite species. *Bioorganic & medicinal chemistry*. [Research Support, Non-U.S. Gov't]. 2010 Mar 15;18(6):2225-31.
9. Prado-Prado FJ, Borges F, Uriarte E, Perez-Montoto LG, Gonzalez-Diaz H. Multi-target spectral moment: QSAR for antiviral drugs vs. different viral species. *Analytica chimica acta*. [Research Support, Non-U.S. Gov't]. 2009 Oct 5;651(2):159-64.
10. Prado-Prado FJ, Borges F, Perez-Montoto LG, Gonzalez-Diaz H. Multi-target spectral moment: QSAR for antifungal drugs vs. different fungi species. *European journal of medicinal chemistry*. [Research Support, Non-U.S. Gov't]. 2009 Oct;44(10):4051-6.
11. Prado-Prado FJ, Martinez de la Vega O, Uriarte E, Ubeira FM, Chou KC, Gonzalez-Diaz H. Unified QSAR approach to antimicrobials. 4. Multi-target QSAR modeling and comparative multi-distance study of the giant components of antiviral drug-drug complex networks. *Bioorganic & medicinal chemistry*. [Research Support, Non-U.S. Gov't]. 2009 Jan 15;17(2):569-75.
12. Prado-Prado FJ, Gonzalez-Diaz H, de la Vega OM, Ubeira FM, Chou KC. Unified QSAR approach to antimicrobials. Part 3: first multi-tasking QSAR model for input-coded prediction, structural back-projection, and complex networks clustering of antiprotozoal compounds. *Bioorganic & medicinal chemistry*. [Comparative Study Research Support, Non-U.S. Gov't]. 2008 Jun 1;16(11):5871-80.
13. Gonzalez-Diaz H, Prado-Prado FJ. Unified QSAR and network-based computational chemistry approach to antimicrobials, part 1: multispecies activity models for antifungals. *Journal of computational chemistry*. [Research Support, Non-U.S. Gov't]. 2008 Mar;29(4):656-67.
14. Prado-Prado FJ, Gonzalez-Diaz H, Santana L, Uriarte E. Unified QSAR approach to antimicrobials. Part 2: predicting activity against more than 90 different species in order to halt antibacterial resistance. *Bioorganic & medicinal chemistry*. [Research Support, Non-U.S. Gov't]. 2007 Jan 15;15(2):897-902.
15. Gonzalez-Diaz H, Prado-Prado FJ, Santana L, Uriarte E. Unify QSAR approach to antimicrobials. Part 1: predicting antifungal activity against different species. *Bioorganic & medicinal chemistry*. [Research Support, Non-U.S. Gov't]. 2006 Sep 1;14(17):5973-80.
16. Nunez MB, Maguna FP, Okulik NB, Castro EA. QSAR modeling of the MAO inhibitory activity of xanthenes derivatives. *Bioorg Med Chem Lett*. 2004 Nov 15;14(22):5611-7.
17. Todeschini R, Consonni V. *Handbook of Molecular Descriptors*. Wiley VCH. 2000.
18. . !!! INVALID CITATION !!!
19. Wen Y, Planel E, Herman M, Figueroa HY, Wang L, Liu L, et al. Interplay between cyclin-dependent kinase 5 and glycogen synthase kinase 3 beta mediated by neuregulin signaling leads to differential effects on tau phosphorylation and amyloid precursor protein processing. *J Neurosci*. 2008;28(10):2624-32.
20. Todeschini R, Consonni V. *Handbook of Molecular Descriptors*. 2000.
21. StatSoft.Inc. STATISTICA (data analysis software system), version 6.0, www.statsoft.com. Statsoft, Inc. 6.0 ed2002.
22. Gonzalez-Diaz H, Bonet I, Teran C, De Clercq E, Bello R, Garcia MM, et al. ANN-QSAR model for selection of anticancer leads from structurally heterogeneous series of compounds. *Eur J Med Chem*. 2007 May;42(5):580-5.
23. Vilar S, Santana L, Uriarte E. Probabilistic neural network model for the in silico evaluation of anti-HIV activity and mechanism of action. *J Med Chem*. 2006;49(3):1118-24.

24. González-Díaz H. QSAR and Complex Networks in Pharmaceutical Design, Microbiology, Parasitology, Toxicology, Cancer and Neurosciences *Curr Pharm Des.* 2010;16(24):2598-600.
25. Gonzalez-Diaz H. Network topological indices, drug metabolism, and distribution. *Curr Drug Metab.* 2010 May;11(4):283-4.
26. Concu R, Podda G, Ubeira FM, Gonzalez-Diaz H. Review of QSAR Models for Enzyme Classes of Drug Targets: Theoretical Background and Applications in Parasites, Hosts, and other Organisms. *Curr Pharm Des.* 2010;16(24):2710-23.
27. González-Díaz H, González-Díaz Y, Santana L, Ubeira FM, Uriarte E. Proteomics, networks and connectivity indices. *Proteomics.* 2008;8:750-78.
28. Vina D, Uriarte E, Orallo F, Gonzalez-Diaz H. Alignment-free prediction of a drug-target complex network based on parameters of drug connectivity and protein sequence of receptors. *Mol Pharm.* 2009 May-Jun;6(3):825-35.
29. Overington J. ChEMBL. An interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). Interview by Wendy A. Warr. *Journal of computer-aided molecular design.* [Interview]. 2009 Apr;23(4):195-8.
30. Heikamp K, Bajorath J. Large-scale similarity search profiling of ChEMBL compound data sets. *Journal of chemical information and modeling.* 2011 Aug 22;51(8):1831-9.
31. Mok NY, Brenk R. Mining the ChEMBL database: an efficient cheminformatics workflow for assembling an ion channel-focused screening library. *Journal of chemical information and modeling.* [Research Support, Non-U.S. Gov't]. 2011 Oct 24;51(10):2449-54.
32. Wassermann AM, Bajorath J. BindingDB and ChEMBL: online compound databases for drug discovery. *Expert opinion on drug discovery.* 2011 Jul;6(7):683-7.
33. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic acids research.* [Research Support, Non-U.S. Gov't]. 2012 Jan;40(Database issue):D1100-7.
34. Tenorio-Borroto E, Penuelas Rivas CG, Vasquez Chagoyan JC, Castanedo N, Prado-Prado FJ, Garcia-Mera X, et al. ANN multiplexing model of drugs effect on macrophages; theoretical and flow cytometry study on the cytotoxicity of the anti-microbial drug G1 in spleen. *Bioorganic & medicinal chemistry.* 2012 Oct 15;20(20):6181-94.
35. Gonzalez-Diaz H, Prado-Prado F, Sobarzo-Sanchez E, Haddad M, Maurel Chevalley S, Valentin A, et al. NL MIND-BEST: a web server for ligands and proteins discovery--theoretic-experimental study of proteins of *Giardia lamblia* and new compounds active against *Plasmodium falciparum*. *Journal of theoretical biology.* [Research Support, Non-U.S. Gov't]. 2011 May 7;276(1):229-49.
36. Patankar SJ, Jurs PC. Classification of inhibitors of protein tyrosine phosphatase 1B using molecular structure based descriptors. *J Chem Inf Comput Sci.* 2003 May-Jun;43(3):885-99.
37. Garcia-Garcia A, Galvez J, de Julian-Ortiz JV, Garcia-Domenech R, Munoz C, Guna R, et al. New agents active against *Mycobacterium avium* complex selected by molecular topology: a virtual screening method. *J Antimicrob Chemother.* 2004 Jan;53(1):65-73.
38. Marrero-Ponce Y, Castillo-Garit JA, Olazabal E, Serrano HS, Morales A, Castanedo N, et al. Atom, atom-type and total molecular linear indices as a promising approach for bioorganic and medicinal chemistry: theoretical and experimental assessment of a novel method for virtual screening and rational design of new lead anthelmintic. *Bioorg Med Chem.* 2005 Feb 15;13(4):1005-20.
39. Marrero-Ponce Y, Machado-Tugores Y, Pereira DM, Escario JA, Barrio AG, Nogal-Ruiz JJ, et al. A computer-based approach to the rational discovery of new trichomonacidal drugs by atom-type linear indices. *Curr Drug Discov Technol.* 2005 Dec;2(4):245-65.
40. Casanola-Martin GM, Marrero-Ponce Y, Khan MT, Ather A, Sultan S, Torrens F, et al. TOMOCOMD-CARDD descriptors-based virtual screening of tyrosinase inhibitors: evaluation of different classification model combinations using bond-based linear indices. *Bioorg Med Chem.* 2007 Feb 1;15(3):1483-503.
41. Casanola-Martin GM, Marrero-Ponce Y, Tareq Hassan Khan M, Torrens F, Perez-Gimenez F, Rescigno A. Atom- and bond-based 2D TOMOCOMD-CARDD approach and ligand-based virtual screening for the drug discovery of new tyrosinase inhibitors. *J Biomol Screen.* 2008 Dec;13(10):1014-24.

42. Casanola-Martin GM, Marrero-Ponce Y, Khan MT, Khan SB, Torrens F, Perez-Jimenez F, et al. Bond-based 2D quadratic fingerprints in QSAR studies: virtual and in vitro tyrosinase inhibitory activity elucidation. *Chemical biology & drug design*. [Research Support, Non-U.S. Gov't]. 2010 Dec;76(6):538-45.
43. Rodriguez-Soca Y, Munteanu CR, Dorado J, Pazos A, Prado-Prado FJ, Gonzalez-Diaz H. Trypano-PPI: A Web Server for Prediction of Unique Targets in Trypanosome Proteome by using Electrostatic Parameters of Protein-protein Interactions. *Journal of Proteome Research*. 2010 FEB 2010;9(2):1182-90.
44. Gonzalez-Diaz H, Muino L, Anadon AM, Romaris F, Prado-Prado FJ, Munteanu CR, et al. MISS-Prot: web server for self/non-self discrimination of protein residue networks in parasites; theory and experiments in *Fasciola* peptides and *Anisakis* allergens. *Molecular Biosystems*. 2011 2011;7(6):1938-55.
45. Riera-Fernandez P, Munteanu CR, Escobar M, Prado-Prado F, Martin-Romalde R, Pereira D, et al. New Markov-Shannon Entropy models to assess connectivity quality in complex networks: from molecular to cellular pathway, Parasite-Host, Neural, Industry, and Legal-Social networks. *Journal of theoretical biology*. [Research Support, Non-U.S. Gov't]. 2012 Jan 21;293:174-88.
46. Martinez-Romero M, Vazquez-Naya JM, Rabunal JR, Pita-Fernandez S, Macenlle R, Castro-Alvarino J, et al. Artificial intelligence techniques for colorectal cancer drug metabolism: ontology and complex network. *Curr Drug Metab*. 2010 May;11(4):347-68.