

MI-DRA 3D: new model for reconstruction of US FDA drug-target network and theoretic-experimental studies of rasagiline derivatives inhibitors of AChE.

Francisco Prado-Prado^{1,2*}, Xerardo García-Mera¹, Nerea Alonso¹,
Francisco Aguirre-Crespo², Jorge Galicia-Vergara²,
Jorgelina Barrios De Tomasi² and Humberto González-Díaz³

¹ *Department of Organic Chemistry, Faculty of Pharmacy, University of Santiago de Compostela (USC), 15782, Spain.*

² *Biomedical Sciences Department, Health Science Division, University of Quintana Roo, 77039, Chetumal, Mexico*

³ *Department of Microbiology and Parasitology, Faculty of Pharmacy, USC, 15782, Spain*

Abstract. The Neurodegenerative diseases have been increasing in the last years. Many of the drug candidates to be used in the treatment of neurodegenerative disease present specific 3D structural features. One important protein in this sense is the acetylcholinesterase (AChE); which is the target of many Alzheimer's dementia drugs. Consequently, the prediction of Drug-Proteins Interactions (DPIs/nDPIs) between new drug candidates with specific 3D structure and targets it is about the major importance. For it, we can use Quantitative Structure-Activity Relationships (QSAR) models to carry out rational DPIs prediction. Unfortunately, many previous QSAR models developed to predict DPIs take into consideration only 2D structural information and codify the activity against only one target. To solve this problem we can develop one 3D multi-target QSAR (3D Mt-QSAR) models. In this communication, we introduce the technique MI-DRA 3D a new predictor for DPIs based two different well-known software. We use the software MARCH-INSIDE (MI) and DRAGON to calculate 3D structural parameters for drugs and targets respectively. Both classes of 3D parameters were used as input to train Artificial Neuronal Network (ANN) algorithms using as benchmark datasets the complex network (CN) formed by all DPIs between US FDA approved drugs and their targets. The entire dataset was downloaded from Drug Bank. The best 3D Mt-QSAR predictor found is one ANN of type Multilayer Perceptron (MLP) with profile MLP 37:37-24-1:1. This MLP classifies correctly 274 out of 321 DPIs (Sensitivity = 85.35%) and 1041 out of 1190 nDPIs (Specificity = 87.48%), corresponding to training Accuracy = 87.03%. We validated the model with external predicting series with Sensitivity = 84.16% (542/644 DPIs; Specificity = 87.51% (2039/2330 nDPIs) and Accuracy = 86.78%. The new CNs of DPIs reconstructed from US FDA can be used to explore large DPIs databases in order to discover both new drugs and/or targets. We carried out theoretic-experimental studies to illustrate the practical use of MI-DRA 3D. First, we reported the prediction and pharmacological assay of 22 different rasagiline derivatives with possible AChE inhibitory activity.

Keywords: Drug-Protein interaction complex networks; Protein Structure Networks; multi-target QSAR; Markov Model; AChE inhibitors

Corresponding authors: PRADO PRADO, F. (fenol1@hotmail.com), *Biomedical Sciences Department, Health Science Division, University of Quintana Roo, 77039, Chetumal, Mexico.*

1. Introduction

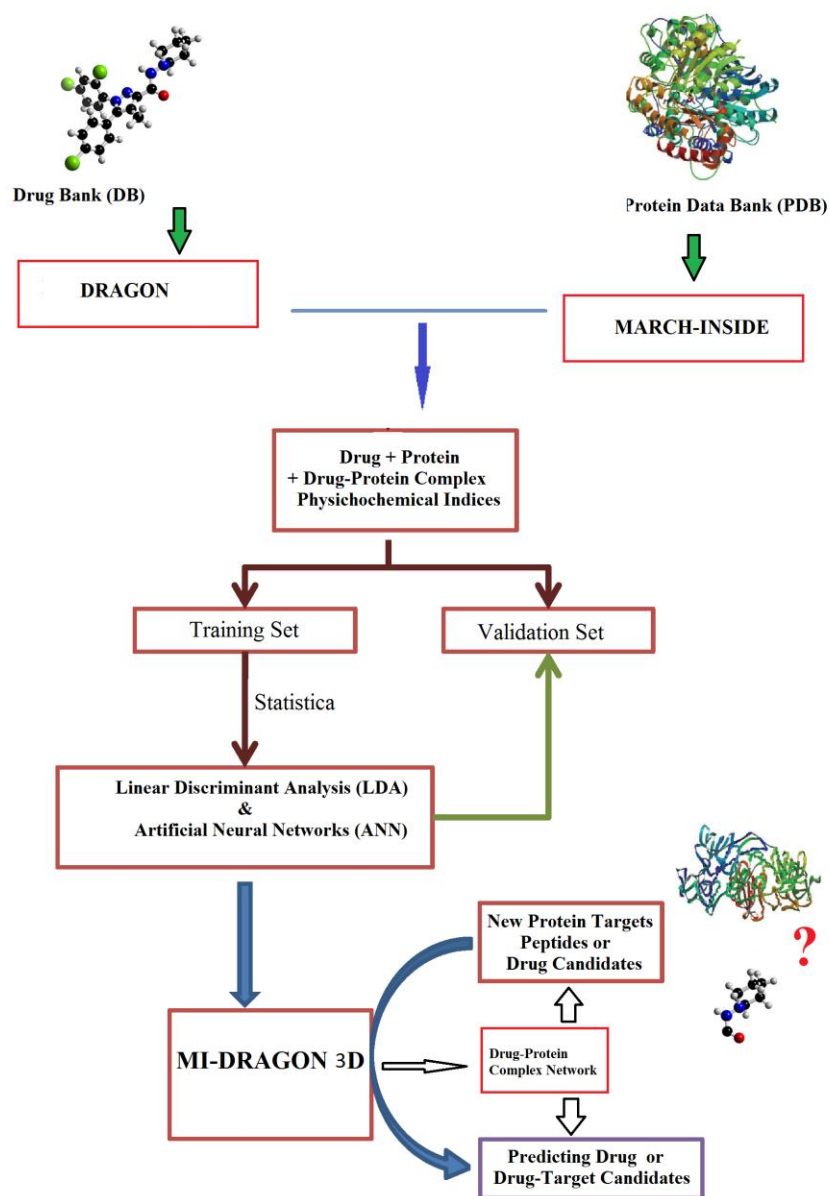
Yildirim, *et al.* have built a complex network (CN) of Drug-Protein Pairs (DPIs) with the form of a bipartite graph composed of all DPIs for all US Food and Drug Administration (US FDA) approved drugs and proteins linked by drug-target binary associations. The resulting CN connects most drugs into a highly interlinked giant component, with strong local clustering of drugs of similar types according to Anatomical Therapeutic Chemical classification. It was motivated due to the strong incentive to develop new methods able of predicting potential drug-target interactions complex networks (CNs) formed by DPIs [2]. For it, we can use Quantitative Structure-Activity Relationships (QSAR) models [5] to carry DPIs prediction. To solve this problem we can develop a 3D multi-target QSAR (3D mt-QSAR) models to predict DPIs [6]. One way to develop this class mt-QSAR is incorporated into the QSAR equation parameters of the structure of the target (protein, DNA, RNA, etc.) in addition to the structural parameters of the drug present in classic QSAR. Some of the more known software we can use to reach this goal are: DRAGON, CODESSA [7], MODES-LAB [8], TOMO-COMD [9], and MARCH-INSIDE (MI) [10]. The software DRAGON is one of the more complete calculating more than 1600 descriptors for drug structure including as zero- (0D) one- (1D), two- (2D), three-dimensional (3D) parameters.

Unfortunately several QSAR models are able to predict the activity of drugs against only one target and/or are unable to codify important 3D structural features. Speck-Planche, *et al.* [3, 4] have developed mt-QSAR for the design of multi-target inhibitors against chemokine receptors. This approach was focused on the construction of a mt-QSAR model for the classification and prediction of inhibitor chemokine receptors. For instance, very recently we have developed in a previous work a QSAR model base on the MARCH-INSIDE method to predict a large network of DTPs. This model was based on 2D structural parameters for drugs and 1D structural parameters for protein. After that we developed MIND-BEST [12] and NL MIND-BEST [13]. Both predictors are based on 3D structural parameters of proteins calculated with software MI but they used only 2D structural parameters of drugs (calculated also with MI). The accuracy of the MIND-BEST model found was 86.32% and NL MIND-BEST was Accuracy = 90.41%. However both models only use 2D parameters using MI software. After that, to improve and obtain better results we use the software MARCH-INSIDE (MI) to calculate 3D structural parameters for targets and the software DRAGON was used to calculate 2D molecular describe all drugs [14]. We introduce the technique 2D MI-DRAGON a new predictor for DPIs based on two different well-known software.

As was mentioned in the previous paragraph we can seek a QSAR predictor for DPIs using molecular descriptors of both drug and target. In this work, we introduce for first time MI-DRA 3D a new predictor for DPIs based on two different well-known software. We use the software MARCH-INSIDE (MI) to calculate 3D structural parameters for targets and the software DRAGON for 3D parameters of all DPIs present in the Drug Bank (US FDA benchmark datasets) [15-18]. Both classes of parameters were used as input of the different Artificial Neuronal Network (ANN) algorithms to seek an accurate non-linear mt-QSAR predictor. MI-DRA 3D offers a good opportunity for fast-track calculation of all possible

DPIs of one drug enabling us to re-construct large drug-target or DPIs Complex Networks (CNs). In this study, we reported the prediction and pharmacological assay of 22 different rasagiline derivatives with AChE inhibitory activity. The present work reports the attempts to calculate within unified DPIs. All this can help to design new inhibitors of AChE. A very good MI-DRA 3D QSAR model was obtained, and the subsequent combined QSAR & CN analysis may become of major importance for the prediction of the activity of new compounds against different targets or the discovery of new targets. In this sense we reported an illustrative study that combines both experiment and theory to show how to use this model in practical situations. We reported the prediction and pharmacological assay of rasagiline derivatives with AChE inhibitory activity. In **Figure 1** we depict a flowchart with the main steps given in this work to train and validate the ANN classifier.

Figure 1. Flowchart of all steps given in this work to develop the new model.



2. Materials and Methods

2.1. Computational methods

2.1.1 MOPAC AM1 Optimization geometry method using CS CHEM 3D.

Molecular structures of all FDA drugs were generated with CHEM 3D Ultra (version 2005). The energy of each intermediate was then minimized using the semi-empirical MOPAC method with a minimum RMS gradient of 0.100, which specifies the convergence criteria for the gradient of the potential energy surface. The geometry of the molecules was optimized and the values of the quantum chemical descriptors of each compound were calculated using AM1. AM1 theory was used with a closed shell function. The MOPAC AM1 method was selected because it was a semi-empirical quantum chemical method and the computational time was much shorter than that needed by ab initio method.

2.1.2. MI-DRAGON technique

3D Parameters for drugs. The DRAGON software 4.0 [19] was utilized here to calculate the 3D parameters of drugs. It depends on whether they are computed from the chemical formula, substructure list representation, molecular graph or geometrical representation of the molecule, respectively [20, 21]. In this work, we calculated only GETAWAY 3D descriptors. We use these descriptors after optimized for use with 3D descriptors.

3D Parameters of proteins. In previous works we have predicted protein function based on different protein structural parameters derived from a Markov matrix that account for electrostatic interactions between amino acid pairs in the 3D structure of the protein. One of the classes of parameters used was called the Shannon Entropy ${}^T\theta_k(R)$ of the Markov matrix. These values are used here as inputs to describe information about the structure of the drug target proteins (T) in order to construct the mt-QSAR models for DTPs. The detailed explanation has been published before [22-30] and reviewed in detail more recently [31]. As follows we give the formula for ${}^T\theta_k(R)$ values and some general explanations:

$${}^T\theta_k(R) = -\sum_{j \in R}^k p_j(R) \cdot \log[p_j(R)] \quad (1)$$

Where, ${}^k p_i(R)$ values are the absolute probabilities with which the effect of the electrostatic interaction propagates from the amino acid i^{th} to other amino acids j^{th} next to it and returns to i^{th} after k -steps. These probabilities refer to: aminoacids considered isolated in the space ($k = 0$), interaction between aminoacids in direct contact ($k = 1$) or spatial ($k > 1$) indirect interactions between amino acids placed at a distance equal to k -times the cut-off distance ($r_{ij} = k \cdot r_{\text{cut-off}}$) in the residue network. Euclidean 3D space $r_3 = (x, y, z)$ coordinates of the C_α atoms of amino acids listed in protein PDB files. For calculation, all water molecules and metal ions were removed [32]. All calculations were carried out with our in-house software MARCH-INSIDE 2.0 [32]. For the calculation, the MARCH-INSIDE software always uses the full matrix, never a sub-matrix, but the last summation term may run either for all amino acids or only for some specific protein regions (R) denoted as: *c* for core, *I* for inner, *m* for middle, and *is* for surface regions, respectively). Consequently, we can calculate different ${}^T\theta_k(R)$ for the amino acids contained in the regions (*c*, *i*, *m*, *s*, or *t*) and placed at a topological distance k each other within this orbit (k is the order) [22, 23, 33-35]. In this work, we have calculated altogether 5 (types of regions) x 6 (orders considered) = 30 ${}^T\theta_k(R)$ indices for each protein.

2.1.3 Statistical analysis. Let be ${}^D\theta_k(G)$ entropy descriptors molecular that codify information about drug structure and ${}^T\theta_k(R)$ entropy descriptors that codify information

about drug target proteins; we attempt to develop a simple mt-QSAR model in the form of a linear classifier with the general formula:

$$S(DTP)_{pred} = \sum_{k=0}^5 a_{G,k} \cdot^D \theta_k(G) + \sum_{k=0}^5 b_{R,k} \cdot^T \theta_k(R) + c_0 \quad (2)$$

We used Linear Discriminating Analysis (LDA) 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 lower values. In equation (6), b represents the coefficients of the classification function, determined by the LDA module of the STATISTICA 6.0 software package [36]. We used Forward Stepwise algorithm for a variable selection. The statistical significance of the LDA model was determined calculating the p-level (p) of error with the Chi - square test. We also inspected the Specificity, Sensitivity, and total Accuracy to determine the quality-of-fit to data in training. Cases for training set were selected at random out of the cases in full data set. The remnant cases were used to validate the model. The validation of the model was corroborated with these external prediction series; these cases were never used to train the model. The ration between training/validation set was 2/1 approximately. This procedure to select training and validation sets is largely known and used to train QSAR models [37-43].

2.1.4 ANN analysis. The non-linear mt-QSAR model was constructed using ANN analysis. All models trained were carried out in STATISTICA 6.0 [36]. In so doing, we used a very simple type of ANN called Three Layers Perceptron (MLP-3) 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 lower values. In equation (2), b represents the coefficients of the LNN classification function, determined by the ANN module of the STATISTICA 6.0 software package [36]. We used Forward Stepwise algorithm for a variable selection.

In addition, we can explore more complicated non-linear ANNs in order to improve the accuracy of the classifier. We processed our data with different ANNs looking for a better model. Four types of ANNs were used, namely, Probabilistic Neural Network (PNN), Radial Basic Function (RBF), Linear Neural Network (LNN), and Four Layer Perceptron (MLP-4) [44, 45]. The quality of all the ANNs (linear or nonlinear) was determined calculating values of Specificity, Sensitivity, and total Accuracy to determine the quality-of-fit to data in training. The validation of the model was corroborated with external prediction series. We also reported ROC-curve analysis (ROC curve can be used to select an optimum decision) for both training and validation series [44, 46].

2.1.5 Data set. The data set was formed by a set of marketing DPIs with a known affinity of drugs by targets. This dataset is the same benchmark data used in previous works [1, 5, 12, 13] in this area and contains all drugs approved by the US FDA. We download this dataset from the public resource called Drug Bank [12, 13, 16-18]. The data set was formed

for more than 519 drugs with their respectively 336 targets. Subsequently, we were able to collect above 4485 cases (drug-protein interactions) instead of 519 x 336 cases. In addition the data set was used to develop ANN models to perform the model.

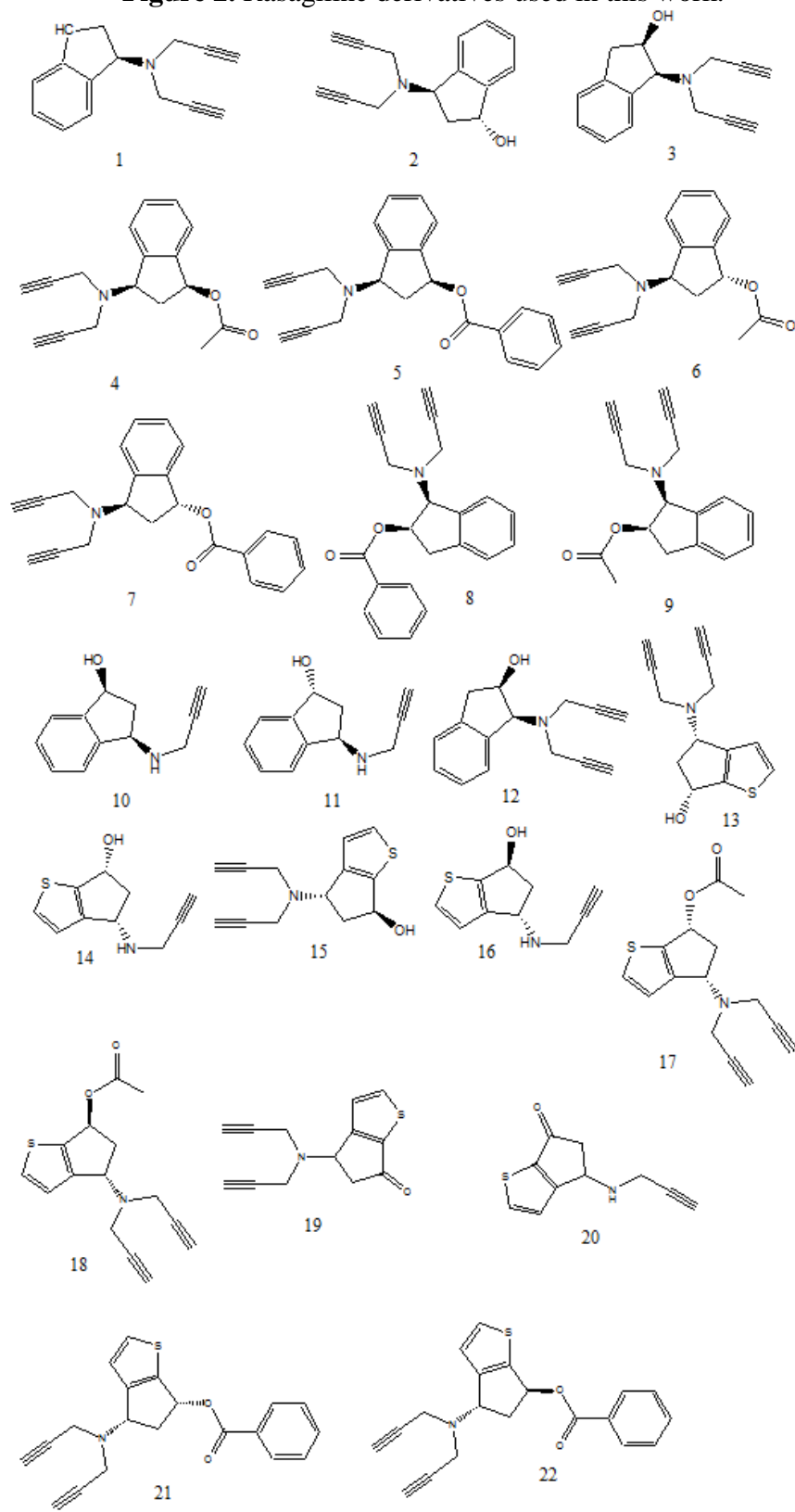
2.1.6 Complex network construction. We construct a DPIs network in order to achieve the drug and protein affinity with a network approach. Generally in this network, one node may represent a drug or a target. On the other hand, the edges represent the DPIs; express relationships between pairs of drugs with their targets [1]. Anyhow, the nodes representing targets may be of at least two types. In almost all cases reported up to date each target is represented only once in the network. In this class of “static” DPIs network the target is depicted by the node corresponding to the X-ray structure of itself. In this work, we build in total two complex networks. First, we constructed the DPIs networks for the observed data and second, DPIs network predicted by the model. The common steps to construct these networks are: First, using the Excel software in a column we introduce all the proteins, the drugs used quotation marks in our database. Then in another column lists all the cases. At the beginning of this column puts the total number of vertices, there are currently two columns of the name of drugs and protein and their corresponding number of vertices. After, at the end of the columns are placed bows in the first column put the number of vertices for the drug and in another column corresponding to the protein. Then, the file was saved as a .txt format file. After we had renamed the .txt file as a .net file we read it with the CentiBin software [47, 48]. Finally, using CentiBin we can not only represent the network but also highlight all drugs and targets (nodes) connected by a specific edge or link (DPI). Using this software we can calculate vertex centralities to analyze the relationships between drug targets.

2.2. Illustrative experiments

2.2.1. Synthesis of Rasagiline derivatives.

Synthesis. Synthesis of compounds **1-22** has been previously reported by us [5, 12], see **Figure 2**.

Figure 2. Rasagiline derivatives used in this work.



2.2.2 Determinations of cholinesterases activities

The cholinesterase assay method of Ellman was used to determine the in vitro cholinesterase activity [49]. The activity was measured by increase in absorbance at 412 NM due to the yellow color produced from the reaction of acetylthiocholine iodide with the dithiobisnitrobenzoate (DTNB) ion. Acetylcholinesterase from human erythrocytes, acetylcholinesterase recombinant expressed in HEK 293 cells and butyrylcholinesterase from human serum was obtained from Sigma.

2.2.3 Experimental conditions and kinetics. Enzyme activity was measured using a FLUOstar Optima microplate reader. The assay medium contained phosphate buffer, pH 8.0, 20 mM DTNB, 0.01 U/ml of enzyme and 0.75 μ M substrate (acetylthiocholine iodide or butyrylthiocholine iodide). The activity was determined by measuring the increase in absorbance at 412 nm at 1 min intervals for 10 min at 37 °C. In a dose-dependent inhibition studies, the substrate was added to the assay medium containing enzyme, buffer, and DTNB with inhibitor after 10 min of incubation time. All experiments were carried out in duplicate and expressed as mean \pm SEM. The relative activity is expressed as the percentage ratio of enzyme activity in the absence of inhibitor, see **Table 1**.

Table 1. Inhibitory activity of different rasagiline derivatives .

Compounds	hAChE (IC ₅₀ μ M)		hAChE (IC ₅₀ μ M)
1	>100 μ M	12	>100 μ M
2	>100 μ M	13	No tested
3	>100 μ M	14	No tested
4	No tested	15	No tested
5	**	16	No tested
6	No tested	17	**
7	**	18	**
8	No tested	19	**
9	**	20	**
10	>100 μ M	21	**
11	>100 μ M	22	**
Galantamine	1.43 \pm 0.03 ^a		
Eserine	151.40 \pm 5.63 nM		
Tacrine	130,90 \pm 6,83 nM		

Each IC₅₀ value is the mean \pm S.E.M. from five experiments.

3. Results

3.1. DPIs QSAR predictive models

3.1.1 LDA model. Common physicochemical properties like entropy have been demonstrated to be useful on protein QSAR [50, 51]. We used these properties as input of our model in addition to drug molecular descriptors. The present is the first mt-QSAR model combining DRAGON and MI to predict the probability with which occurs DPIs between a drug and a protein. This type of models lie within the frontiers between classic QSAR for drugs and protein QSAR [33]. Some applications of the present model are the prediction of new drugs, new protein receptors or drug targets, and drug binding sites. Based on the algorithms described in materials and methods the best linear model found was the following:

$$S(DTP)_{pred} = 11.01 \cdot d_1 - 36.37 \cdot d_2 - 12.77 \cdot d_3 - 9.34 \cdot d_4 - 52.17 \cdot d_5 + 1.62 \cdot d_6 \\ - 1.65 \cdot d_7 - 0.11 \cdot d_8 - 0.10 \cdot d_9 + 0.25d_{10} + 2.48 \quad (3) \\ N = 4485 \quad \chi^2 = 919.2988 \quad p\text{-level} < 0.001$$

Table 2. Detailed list of the symbols and description for all parameters present in the model.

Original Descriptor	Descriptor name	Code ID
H7v	H autocorrelation of lag 7 / weighted by atomic van der Waals volumes	d1
HATS5v	leverage-weighted autocorrelation of lag 5 / weighted by atomic van der Waals volumes	d2
HATS4e	leverage-weighted autocorrelation of lag 4 / weighted by atomic Sanderson electronegativities	d3
HATS6e	leverage-weighted autocorrelation of lag 6 / weighted by atomic Sanderson electronegativities	d4
R5e+	R maximal autocorrelation of lag 5 / weighted by atomic Sanderson electronegativities	d5
${}^T\theta_4(\text{core})$	Entropy of all aminoacids placed in the core region and all the neighbors at distance $k \leq 4$	d6
${}^T\theta_5(\text{core})$	Entropy of all aminoacids placed in the core region and all the neighbors at distance $k \leq 5$	d7
${}^T\theta_5(\text{inner})$	Entropy of all aminoacids placed in the inner region and all the neighbors at distance $k \leq 5$	d8
${}^T\theta_2(\text{middle})$	Entropy of all aminoacids placed in the middle region and all the neighbors at distance $k \leq 2$	p1
${}^T\theta_0(\text{surface})$	Entropy of all aminoacids placed in the surface region and all the neighbors at distance $k \leq 0$	d9

The nomenclature used in the descriptors of the equation is found in **Table 2**. In this equation, N is the number of cases, χ^2 is the Chi-square and p is the level of error. This model, with 10 variables, classifies correctly 256 out of 321 DPIs (Sensitivity of 79.75%) and 1014 out of 1190 nDPIs (Specificity of 85.21%). Overall Training Accuracy was 84.05%. The validation of the model was carried out by means of external predicting series. The model classifies correctly 498 out of 644 DPIs (77.33%) and 2000 out of 2330 nDPIs

(85.84%) in validation series. Accuracy of validation series (predictability) was 83.99%. These results (**Table 3**) indicate that we developed an accurate model according to previous reports on the use of LDA in QSAR [52, 53].

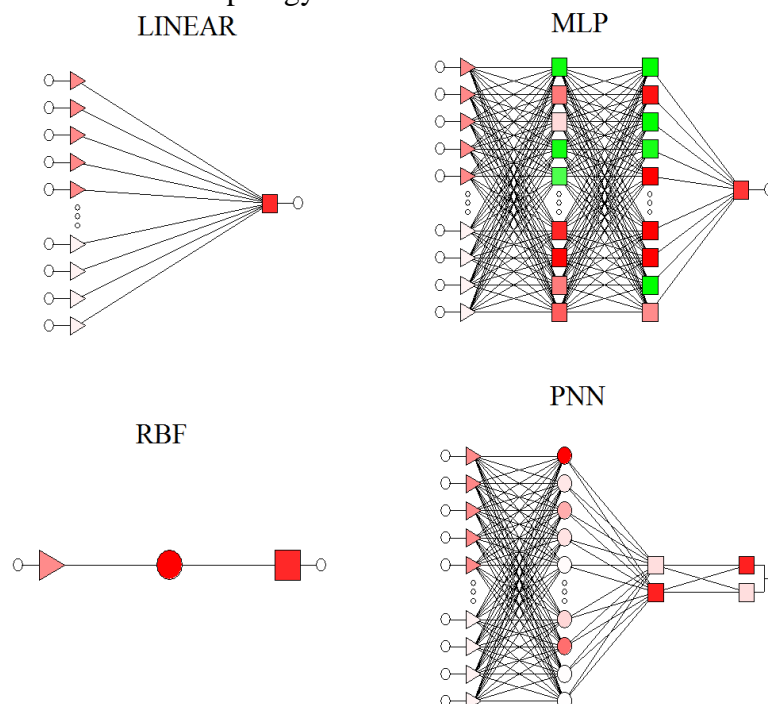
Table 3. Comparison of LDA and different ANNs classification models.

Model profile	Class	Train			Stat.	Validation		
		%	DPIs	nDPIs	Par.	%	DPIs	nDPIs
MI DRAGON 3D MLP 37:37-24-1:1	DPIs	85.36	274	47	Sn	84.16	542	102
	nDPIs	87.48	149	1041	Sp	87.51	291	2039
	Total	87.03			Ac	86.79		
LDA ^a 10:10-1:1	DPIs	79.75	256	65	Sn	77.33	498	146
	nDPIs	85.21	176	1014	Sp	85.84	330	2000
	Total	84.05			Ac	83.99		
PNN 227:227-14797-2-2:1	DPIs	0	0	644	Sn	0	0	321
	nDPIs	100	0	2346	Sp	100	0	1174
	Total	78.46			Ac	78.53		
RBF 1:1-1-1:1	DPIs	47.05	303	341	Sn	52.65	169	152
	nDPIs	56.01	1032	1314	Sp	54.86	530	644
	Total	54.08			Ac	54.38		
LNN 227:227-1:1	DPIs	53.73	346	298	Sn	45.79	147	174
	nDPIs	32.05	1594	752	Sp	31.52	804	370
	Total	36.72			Ac	34.58		

DPIs: Drug-Target Pairs for compounds with high affinity; nDPIs: Drug-Target Pair for compounds with non-affinity; Stat. is statistics, Par. is parameter

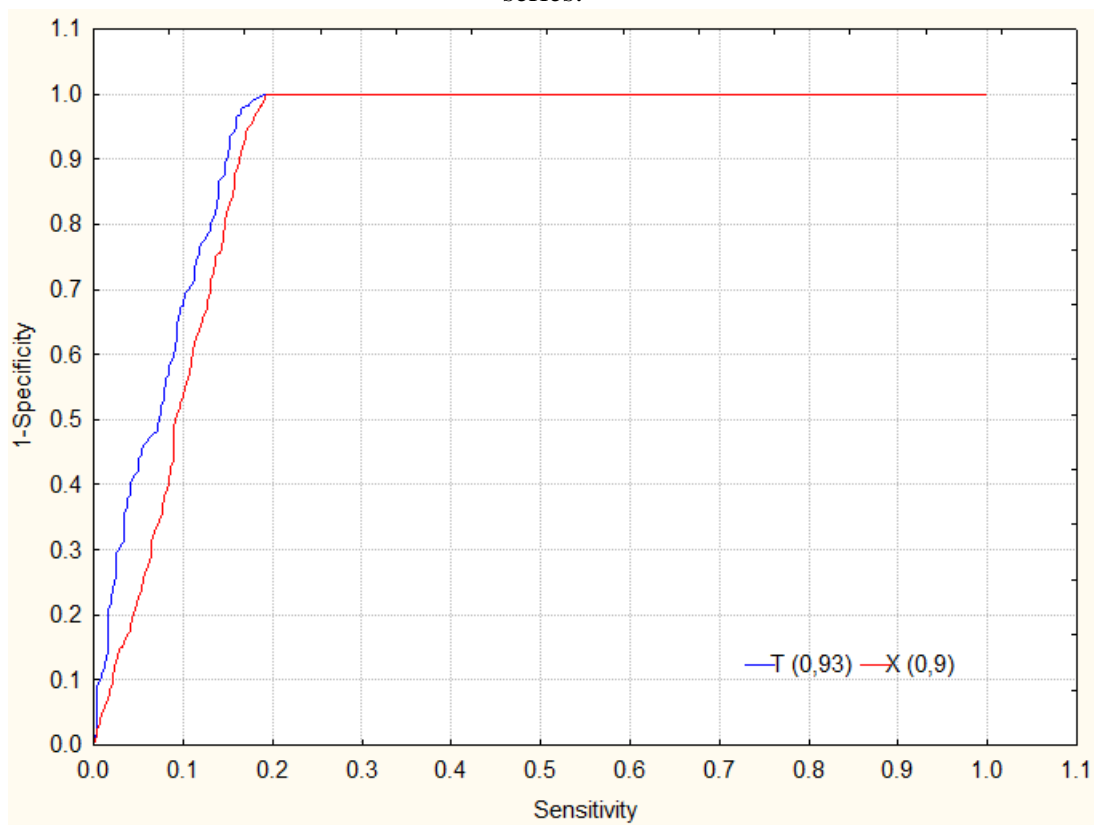
3.1.2 MI-DRA 3D ANN model. The previous model shows good results with a relatively small number of parameters (10 parameters) and a linear equation. However, as a result of the previous section we decided to carry out an ANN analysis to seek a better model using a non-linear method. Four types of ANNs were used, namely, Probabilistic Neural Network (PNN), Radial Basic Function (RBF), Three Layers Perceptron (MLP-3), and Four Layer Perceptron (MLP-4). See, previous works about the use of these ANNs in protein QSAR [5, 13]. The **Figure 3** depicts the network topology for some of the ANN models tested. In general, at least one ANN of every type tested was statically significant. However, one must note that the profiles of each network indicate that many of these are highly non-linear and complicated models.

Figure 3. Generic Topology of ANN models trained in this work.



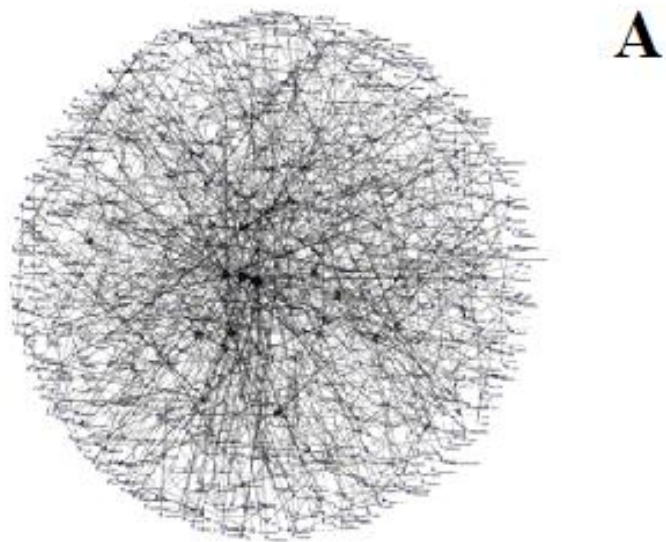
Models using ANN-QSAR has been demonstrated before; see, for instance, the works of Fernández and Caballero [54, 55]. We compare different types of networks to obtain a better model. In **Table 3** we show the classification matrix of the different networks. The profiles of networks tested were RBF 1:1-1-1:1 with only one variable; LNN 227:227-1:1, which present many variables, and PNN 227:227-14797-2-2:1, which has a very high number of hidden neurons, see **Table 3**. After that, the simpler but more accurate ANN model found was an MLP (MLP 37:37-24-1:1) with training Accuracy = 87.03 %. This was selected as the best network found because it presents both high accuracy and an adequate number of variables accounting for features relevant for DPIs. This ANN presents 37 input variables ($24 d_k + 13 \Theta_m$). This leads to 37 neurons in first or input layer (I), 24 neurons in the second layer or first hidden layer (H1) and only one neuron (DPI prediction) in the output layer (O). We depict the ROC-curve for MLP 37:37-24-1:1 to show how reliable was the network model developed, see **Figure 4**. Notably, the model presented had a ROC curve higher than 0.5. The model presented an area greater than 0.92. From now on we call the ANN MLP 37:37-24-1:1 as the 3D MI DRAGON predictor.

Figure 4. ROC Curve for MI-DRA 3DGON predictor (red = train series, blue = validation series).



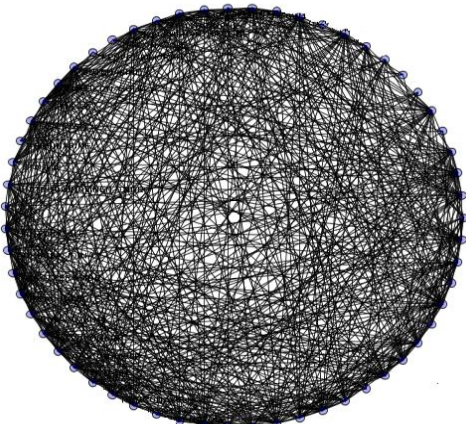
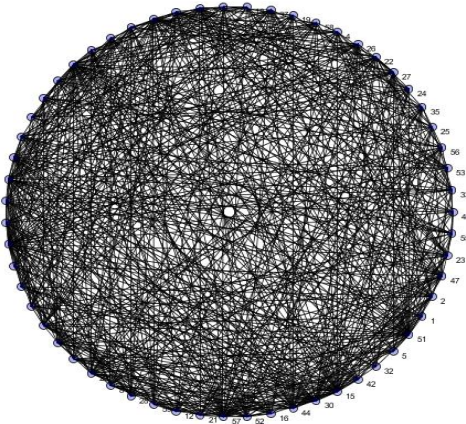
3.1.2.1 MI-DRA 3D assembly of CNs for DPIs. The construction of multi-protein CNs that incorporates protein affinity profile for drugs or the same CNs for DPIs is relevant to drug and target screening. And is one application of this model. In order to recall the capacity of MI-DRA 3D to predict new CNs of DPIs we selected the same benchmark database used in previous works [5, 13, 14]; which includes US FDA approved drugs with their targets. With these goals in mind, we constructed again and manually curated the above-mentioned CN obtaining a graph with 855 vertices or nodes (drugs and proteins) and $m = 1016$ DPIs (edges). This CN of DPIs have $D = 6.7$; average topological distances D_{ij} between all pairs of nodes. The same as before, we constructed a new CN of DPIs but connecting only pairs of nodes with DPIs predicted by MI-DRA 3D. In so doing, we obtained a value of $D = 7.2$ and $m = 1256$ DPIs. In **Figure 5** we illustrated visually both CNs (observed and predicted).

Figure 5. Observed vs. Predicted drug-target complex networks.



In the first instance, we compare this predicted network (MI-DRA 3D) with 2D MI-DRAGON predicted network [14]. We compare to observing the similar or dissimilar topology (connectivity pattern structure) between them. Measured in terms of TIs such as: number of nodes (n), number of edges (m), Wiener index (W), diameter (D), the Randic connectivity index (Xr), topological distance (Dist), network average values for radiality (R), node degree (δ), eccentricity (E). In **Table 4**, we observe all the TIs are similar excepting n, m and w. That means both CNs has a high similarity between them. These results are very interesting, because our MI-DRA 3D model present similar results to the 2D MI-DRAGON model, which results have been published successfully before.

Table 4 Comparison MI-DRA 3D versus 2D MI-DRAGON.

2D MI-DRAGON	Value	TIs	Value	MI-DRA 3D
	706	n	59	
	907	m	631	
	1826812	W	2057954	
	18	D	19	
	255.09	Xr	266.39	
	2.49	δ	2.44	
	6.7	Dist	7.2	
	0.078	E	0.083	
	12.43	R	11.39	

^aThe TIs used are: number of nodes (n), number of edges (m), Wiener index (W), diameter (D), the Randic connectivity index (Xr), topological distance (Dist), network average values for radiality (R), node degree (δ), eccentricity (E).

To see how reliable and valid is our model. Not only compared to TIs to observe the similarity between both predicted networks, but we study the centrality analysis of given networks too. This type of drug screening and drug target discovery is the calculation of those nodes (drugs or proteins) which are more relevant or important (central) in the graph. In it we can use numerical parameters that quantify the importance of a node in a graph which are called node centralities C_t of type t [56]. These nodes identifications using node centralities may help us to identify the most relevant drugs or proteins in analogy to similar procedures developed for PINs; networks of Protein-Protein Interactions (PPIs) [57]. In **Table 5** we show the predicted results of both node degree centrality (C_δ) and closeness centrality (C_{clo}) for proteins and drugs present in the database and compare with the predicted results of the 2D-MI-DRAGON model. The parameter C_δ measures the local importance of a node by counting the number of nodes directly attached to him [57]. Conversely, C_{clo} measures the global importance of a node in a CN by taking in consideration the inverse of the sum of D_{ij} ($C_{clo} = 1/\sum D_{ij}$) [58]. Consequently, the higher C_δ the higher is the local importance of the node but the higher C_{clo} the lower is the global importance of the node. For instance, the protein IHA2 is one important protein both locally and globally in this CNs with lower $C_{clo} > 4$ and a $C_\delta = 26$. It means that this protein is both locally and globally important because it is the target of many drugs (high C_δ). This result is similar to obtained from 2D MI-DRAGON model. Another interesting result was

simvastatin. Simvastatin is a hypolipidemic drug used to control elevated cholesterol, or hypercholesterolemia. It is a member of the statin class of pharmaceuticals. The primary use of simvastatin is for the treatment of dyslipidemia and the prevention of cardiovascular disease [59, 60]. Depending on our aims the more important nodes in pharmacological terms don't necessarily have to be the more central in the graph (those with higher C_δ and lower C_{clo}), see **Table 5**. We show in this example, our model predicts efficiently. We found that the MI-DRA 3D model shows very similar results to the previous model, which has been published with excellent results.

Table 5. Results of node degree (C_δ) and closeness centrality (C_{clo}) for 20 proteins and drugs.

Drug/PDB	C_δ <small>2D-MI-DRAGON</small>	C_δ <small>MI-DRA 3D</small>	Drug/PDB	C_{clo} <small>2D-MI-DRAGON</small>	C_{clo} <small>MI-DRA 3D</small>
1HA2	44	26	1HA2	4.80	3.54
1BNA	36	40	Simvastatin	4.22	4.11
NADH	35	33	Gliclazide	4.17	3.48
1R5K	27	29	Saquinavir	4.16	3.46
Simvastatin	18	17	1BNA	4.15	3.46
1EMI	16	21	Cefalotin	4.13	2.98
1CZM	14	17	Atorvastatin	4.09	3.44
1NHZ	14	14	1A8M	4.09	3.75
1MO8	14	14	1XF0	4.07	3.09
1UZF	13	13	Estrone	4.06	2.61
1SQN	13	13	Ketoprofen	4.02	2.86
1T9N	13	12	Testosterone	4.02	2.42
1BYW	11	15	1TZI	4.02	3.77
1VRU	11	10	1KED	4.00	3.16
1E3G	11	10	Captopril	3.99	3.49
Atorvastatin	11	10	Liothyronine	3.97	3.21
1ZNC	10	11	Diflunisal	3.96	3.20
1ODW	10	10	Halothane	3.95	3.11
Pyridoxal Phosphate	9	9	Digitoxin	3.94	3.45
1HWL	9	9	Pyridoxine	3.94	3.08

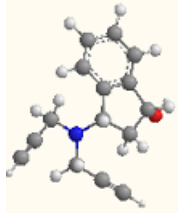
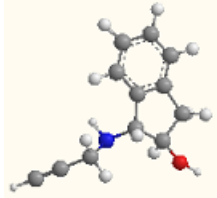
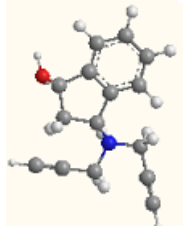
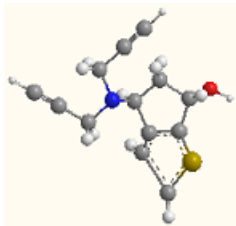

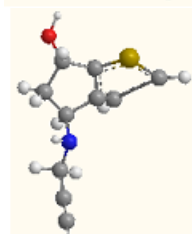
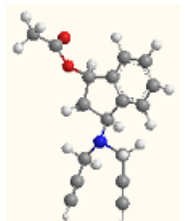
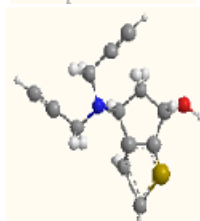
3.2. Theoretic-Experimental Study is using MI-DRA 3D predictor

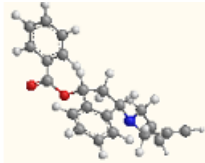

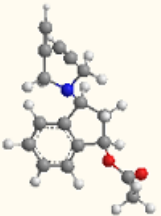
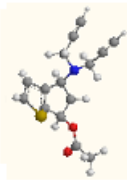
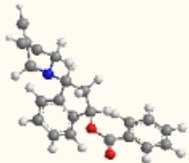
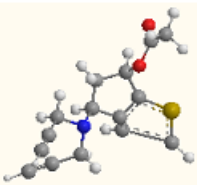
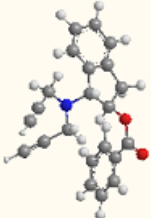
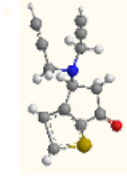
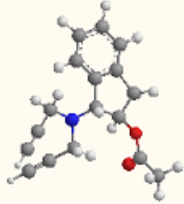
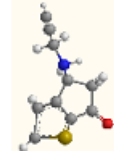
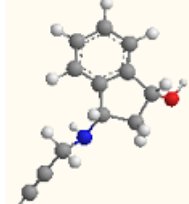
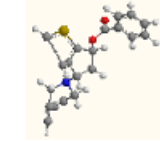
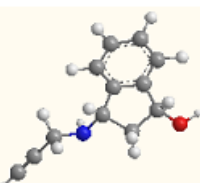
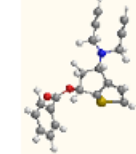
Finally, we illustrated in one theoretic-experimental study the practical use of MI-DRA 3D. We reported the prediction, synthesis, and pharmacological assay of 20 different rasagiline derivatives with AChE inhibitory activity.

3.2.1 MI-DRA 3D prediction of rasagiline derivatives vs. AChE. In this *in silico* experiment we used MI-DRA 3D to predict the interaction of the rasagiline derivatives with respect to AChE. For it, we downloaded the 3D structure of AChE protein with PDB ID 1EEA and calculated their structural parameters with MI. We also generated the SMILE codes for these compounds and we use MOPAC AM1 Optimization geometry method for these compounds for calculating their 3D structural parameters with DRAGON. After that, we predicted their propensity to undergo DPIs with AChE using as inputs for the MI-DRA 3D predictor the structural parameters of both the drugs and the protein. In **Table 6** we

confront the results obtained using this model and the outcomes of the pharmacological assay. No compounds are selective inhibitors of AChE, which is why we used as control galantamine for AChE was. We consider the observed class of active compounds OC = 1 if compound $IC_{50} < 10 \mu\text{M}$ this cutoff is in the similar range than other used in previous works [61, 62]. As we can see in this table all the compound rasagiline derivatives present some activity, But none of these compounds have inhibitory activity in the pharmacological assays. All of our compounds in the pharmacological assay (OC = 0) were inactive. MI-DRA 3D predicted as inactive all compounds, excepting 3. The model classified correctly 19 of 22 compounds tested (86.36%). In this test, our model was compared with pharmacological testing of 22 compounds synthesized by us. And we can observe the effectiveness of our model with experimental data. Also, we note that the model predicts all compounds tested as inactive, this is important because the model allows to discriminate between active and inactive compounds. However, some compounds were not tested by pharmacological assay, that compounds were predicted as inactive using MI-DRA 3D model. We discarded pharmaceutical assays of these compounds; because we consider our model reliable. This kind of model can be used to save efforts and money to perform the pharmacological tests. This is a good example of how reliable is the MI DRAGON 3D model.

Table 6. Prediction of rasagiline derivatives with MI-DRA 3D predictor

DRUG	OC	PC	Score	Structure	DRUG	OC	PC	Score	Structure
1	0	0	0.95		12	0	0	0.63	
2	0	0	0.95		13	0	0	1.00	
3	0	0	0.86		14	0	0	0.87	
4	0	0	0.95		15	0	0	1.00	

5	0	1	0.52		16	0	0	0.87	
6	0	0	0.95		17	0	0	1.00	
7	0	1	0.52		18	0	0	1.00	
8	0	0	0.88		19	0	0	1.00	
9	0	0	0.89		20	0	0	1.00	
10	0	0	0.75		21	0	0	0.97	
11	0	1	0.27		22	0	0	0.97	

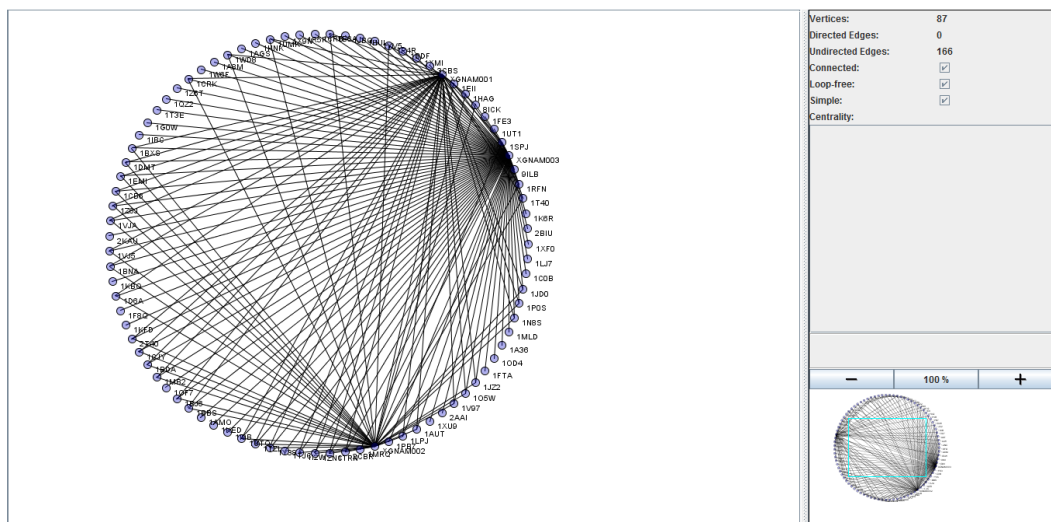
OC = Observed class; PC = Predicted class

3.2.2 MI-DRA 3D complex network of rasagiline derivatives vs. US FDA proteins. An additional use of MI-DRA 3D was to carry out the “*in silico*” or virtual screening of the new compounds with respect to all other targets previously approved by US FDA [14, 63]. It may help to find new targets for these drugs or discard possible toxicological effects depending on the other targets predicted and/or discarded for these compounds. This type of experiment is about the major importance due to the cost in terms of animal sacrifice, time, materials and human resources of the experimental assay of all compounds against all these targets, see recent reviews by Duardo-Sánchez *et al.* [64-67]. In fact, over a decade, the US FDA has been engaged in the applied research, development, and evaluation of computational toxicology methods used to support the safety evaluation of a diverse set of regulated products. The basis for evaluating computational toxicology methods is multi-factorial, including the potential for increased efficiency, reduction in the numbers of animals used, lower costs, and the need to explore emerging technologies that support the goals of the US FDA's Critical Path Initiative (e.g. To make decision support information available early in the drug review process) [68].

In this experiment, we downloaded the 3D structure of all proteins that are targets of US FDA approved drugs. Next, we calculated the structural parameters of all these proteins with MI. We also generated the SMILE codes for these compounds and we use MOPAC AM1 Optimization geometry method for these compounds for calculating their 3D structural parameters with DRAGON. After that, we predicted their propensity to undergo DPIs with all US FDA proteins using as inputs for the MI-DRA 3D predictor the structural parameters of both the drugs and proteins. We predicted all proteins in FDA dataset vs. the 22 rasagiline derivatives. We found that most of 22 derivatives were predicted as non-active (low DPIs scores) against most proteins in the FDA database. Consequently, MI-DRA 3D predicts a high selectivity of rasagiline derivatives as AChE inhibitors. We can reach this goal because the model predicts these compounds as non-active with respect to most proteins that are targets of FDA drugs.

Using these results, we constructed a DP-CN for rasagiline derivatives and the FDA dataset (see **Figure 6**). As a result we obtained a CN with 87 nodes (FDA drugs, proteins, or rasagiline derivatives) and 166 DP (edges, DTPs). As In this network we can see that protein 1EEA (AChE) is predicted to interact with compound 3, this protein is an AChE target [69]. These results are good because they agree with the experimental results presented in this paper where the compound 3 show low AChE activity. The use of such complex networks can help us find and predict new drugs-protein interactions, and therefore find new drugs with improved biological activity and fewer side effects, especially in neural disease.

Figure 6. DP-CN for rasagiline derivatives and the FDA dataset.



4. Conclusions

The MI-DRA 3D predictor based on structural parameters of drugs calculated with DRAGON and parameters of proteins calculated with MI. It is possible to seek excellent predictors for DPIs using as input structural parameters of drugs and proteins calculated with different programs and combined with ANN models. Combining **MARCH-INSIDE** and DRAGON approach and ANN is possible to seek one mt-QSAR classifier to predict with Accuracy > 85% the probability of drugs to bind more than 500 different drug target proteins approved by FDA of USA. MI-DRA 3D predictor is also useful to assemble CNs of DPIs. These CNs computationally assemble offer an alternative to discover new drugs or targets, and explore the selectivity of drugs. In this work, we exemplified these conclusions through the experimental-theoretical study of the AChE activity of new rasagiline derivatives.

5. Acknowledgments

The authors thank sponsorships for a research position at the University of Santiago de Compostela from *Angeles Alvariño*, Xunta de Galicia.

References

- [1] M.A. Yildirim, K.I. Goh, M.E. Cusick, A.L. Barabasi, M. Vidal, Drug-target network, *Nature biotechnology*, 25 (2007) 1119-1126.
- [2] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, M. Kanehisa, Prediction of drug-target interaction networks from the integration of chemical and genomic spaces, *Bioinformatics*, 24 (2008) i232-240.
- [3] A. Speck-Planche, V.V. Kleandrova, F. Luan, M.N. Cordeiro, Chemoinformatics in Multi-Target Drug Discovery for Anti-Cancer Therapy: In Silico Design Of Potent And Versatile Anti-Brain Tumor Agents, *Anticancer Agents Med. Chem.*, (2011).
- [4] A. Speck-Planche, V.V. Kleandrova, In silico design of multi-target inhibitors for C-C chemokine receptors using substructural descriptors, *Mol. Divers.*, (2011).
- [5] F. Prado-Prado, X. García-Mera, P. Abeijón, N. Alonso, O. Caamaño, M. Yañez, T. Garate, M. Mezo, M. González-Warleta, L. Muino, F.M. Ubeira, H. González-Díaz, Using entropy of drug and protein graphs to predict FDA drug-target network: theoretic-experimental study of MAO inhibitors and hemoglobin peptides from *Fasciola hepatica*, *Eur. J. Med. Chem.*, 46 (2011) 1074-1094.
- [6] F.J. Prado-Prado, F. Borges, L.G. Pérez-Montoto, H. González-Díaz, Multi-target spectral moment: QSAR for antifungal drugs vs. different fungi species, *Eur. J. Med. Chem.*, 44 (2009) 4051-4056.
- [7] A.M. Helguera, R.D. Combes, M.P. González, M.N.D.S. Cordeiro, Applications of 2D descriptors in drug design: a DRAGON tale, *Curr Top Med Chem*, 8 (2008) 1628-1655.
- [8] E. Estrada, E. Molina, D. Nodarse, E. Uriarte, Structural contributions of substrates to their binding to P-Glycoprotein. A TOPS-MODE approach, *Curr. Pharm. Des.*, 16 (2010) 2676-2709.
- [9] Y. Marrero-Ponce, G.M. Casanola-Martín, M.T. Khan, F. Torrens, A. Rescigno, C. Abad, Ligand-based computer-aided discovery of tyrosinase inhibitors. Applications of the TOMOCOMD-CARDD method to the elucidation of new compounds, *Curr. Pharm. Des.*, 16 (2010) 2601-2624.
- [10] H. González-Díaz, A. Duardo-Sánchez, F.M. Ubeira, F. Prado-Prado, L.G. Pérez-Montoto, R. Concu, G. Podda, B. Shen, Review of MARCH-INSIDE & complex networks prediction of drugs: ADMET, anti-parasite activity, metabolizing enzymes and cardiotoxicity proteome biomarkers, *Curr. Drug. Metab.*, 11 (2010) 379-406.
- [11] D. Viña, E. Uriarte, F. Orallo, H. González-Díaz, Alignment-free prediction of a drug-target complex network based on parameters of drug connectivity and protein sequence of receptors, *Mol. Pharm.*, 6 (2009) 825-835.
- [12] H. González-Díaz, F. Prado-Prado, X. García-Mera, N. Alonso, P. Abeijón, O. Caamaño, M. Yañez, C.R. Munteanu, A. Pazos, M.A. Dea-Ayuela, M.T. Gómez-Muñoz, M.M. Garijo, J. Sansano, F.M. Ubeira, MIND-BEST: Web server for drugs and target discovery; design, synthesis, and assay of MAO-B inhibitors and theoretical-experimental study of G3PDH protein from *Trichomonas gallinae*, *J. Prot. Res.*, 10 (2011) 1698-1718.
- [13] H. González-Díaz, F. Prado-Prado, E. Sobarzo-Sánchez, M. Haddad, S. Maurel Chevalley, A. Valentin, J. Quetin-Leclercq, M.A. Dea-Ayuela, M. Teresa Gómez-Muñoz, C.R. Munteanu, J.J. Torres-Labandeira, X. García-Mera, R.A. Tapia, F.M. Ubeira, 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*, *J. Theor. Biol.*, 276 (2011) 229-249.

- [14] F. Prado-Prado, X. García-Mera, M. Escobar, E. Sobarzo-Sánchez, M. Yañez, P. Riera-Fernández, H. González-Díaz, 2D MI-DRAGON: a new predictor for protein-ligands interactions and theoretic-experimental studies of US FDA drug-target network, oxoisoaporphine inhibitors for MAO-A and human parasite proteins, *Eur. J. Med. Chem.*, 46 (2011) 5838-5851.
- [15] J. Kirchmair, P. Markt, S. Distinto, D. Schuster, G.M. Spitzer, K.R. Liedl, T. Langer, G. Wolber, The Protein Data Bank (PDB), its related services and software tools as key components for in silico guided drug discovery, *J. Med. Chem.*, 51 (2008) 7021-7040.
- [16] C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A.C. Guo, D.S. Wishart, DrugBank 3.0: a comprehensive resource for 'omics' research on drugs, *Nucleic Acids Res.*, 39 (2011) D1035-1041.
- [17] D.S. Wishart, C. Knox, A.C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali, DrugBank: a knowledgebase for drugs, drug actions and drug targets, *Nucleic Acids Res.*, 36 (2008) D901-906.
- [18] D.S. Wishart, C. Knox, A.C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, *Nucleic Acids Res.*, 34 (2006) D668-672.
- [19] Talete srl, DRAGON for Windows (Software for Molecular Descriptor Calculations), 2005.
- [20] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, 2000.
- [21] E. Papa, F. Villa, P. Gramatica, Statistically validated QSARs, based on theoretical descriptors, for modeling aquatic toxicity of organic chemicals in Pimephales promelas (fathead minnow), *J. Chem. Inf. Model.*, 45 (2005) 1256-1266.
- [22] H. González-Díaz, Y. Pérez-Castillo, G. Podda, E. Uriarte, Computational Chemistry Comparison of Stable/Nonstable Protein Mutants Classification Models Based on 3D and Topological Indices, *J. Comput. Chem.*, 28 (2007) 1990-1995.
- [23] H. González-Díaz, L. Saíz-Urra, R. Molina, Y. González-Díaz, A. Sánchez-González, Computational chemistry approach to protein kinase recognition using 3D stochastic van der Waals spectral moments, *J. Comput. Chem.*, 28 (2007) 1042-1048.
- [24] G. Agüero-Chapín, A. Antunes, F.M. Ubeira, K.C. Chou, H. González-Díaz, Comparative Study of Topological Indices of Macro/Supramolecular RNA Complex Networks, *J. Chem. Inf. Mod.*, 48 (2008) 2265-2277.
- [25] M. Cruz-Montenegro, C.R. Munteanu, F. Borges, M.N.D.S. Cordeiro, E. Uriarte, K.-C. Chou, H. González-Díaz, Stochastic molecular descriptors for polymers. 4. Study of complex mixtures with topological indices of mass spectra spiral and star networks: The blood proteome case, *Polymer*, 49 (2008) 5575-5587.
- [26] M.A. Dea-Ayuela, Y. Pérez-Castillo, A. Meneses-Marcel, F.M. Ubeira, F. Bolas-Fernández, K.C. Chou, H. González-Díaz, HP-Lattice QSAR for dynein proteins: experimental proteomics (2D-electrophoresis, mass spectrometry) and theoretic study of a *Leishmania infantum* sequence, *Bioorg. Med. Chem.*, 16 (2008) 7770-7776.
- [27] G. Agüero-Chapín, H. González-Díaz, G. de la Riva, E. Rodríguez, A. Sánchez-Rodríguez, G. Podda, R.I. Vázquez-Padrón, MMM-QSAR recognition of ribonucleases without alignment: comparison with an HMM model and isolation from *Schizosaccharomyces pombe*, prediction, and experimental assay of a new sequence, *J. Chem. Inf. Mod.*, 48 (2008) 434-448.
- [28] G. Ferino, H. González-Díaz, G. Delogu, G. Podda, E. Uriarte, Using spectral moments of spiral networks based on PSA/mass spectra outcomes to derive quantitative

proteome-disease relationships (QPDRs) and predicting prostate cancer, *Biochem. Biophys. Res. Commun.*, 372 (2008) 320-325.

[29] H. González-Díaz, M.A. Dea-Ayuela, L.G. Pérez-Montoto, F.J. Prado-Prado, G. Agüero-Chapín, F. Bolas-Fernández, R.I. Vázquez-Padrón, F.M. Ubeira, QSAR for RNases and theoretic-experimental study of molecular diversity on peptide mass fingerprints of a new *Leishmania infantum* protein, *Mol. Divers.*, (2009).

[30] G. Agüero-Chapín, J. Varona-Santos, G.A. de la Riva, A. Antunes, T. González-Villa, E. Uriarte, H. González-Díaz, Alignment-Free Prediction of Polygalacturonases with Pseudofolding Topological Indices: Experimental Isolation from *Coffea arabica* and Prediction of a New Sequence, *J. Prot. Res.*, 8 (2009) 2122-2128.

[31] H. González-Díaz, F. Prado-Prado, F.M. Ubeira, Predicting antimicrobial drugs and targets with the MARCH-INSIDE approach, *Curr. Top. Med. Chem.*, 8 (2008) 1676-1690.

[32] H. González-Díaz, Y. González-Díaz, L. Santana, F.M. Ubeira, E. Uriarte, Proteomics, networks and connectivity indices, *Proteomics*, 8 (2008) 750-778.

[33] H. González-Díaz, L. Saiz-Urra, R. Molina, L. Santana, E. Uriarte, A model for the recognition of protein kinases based on the entropy of 3D van der Waals interactions, *J. Prot. Res.*, 6 (2007) 904-908.

[34] H. González-Díaz, R. Molina, E. Uriarte, Recognition of stable protein mutants with 3D stochastic average electrostatic potentials, *FEBS Lett.*, 579 (2005) 4297-4301.

[35] R. Concu, G. Podda, E. Uriarte, H. González-Díaz, Computational chemistry study of 3D-structure-function relationships for enzymes based on Markov models for protein electrostatic, HINT, and van der Waals potentials, *J. Comput. Chem.*, 30 (2009) 1510-1520.

[36] StatSoft.Inc., STATISTICA (data analysis software system), version 6.0, www.statsoft.com.Statsoft, Inc., in, 2002.

[37] G.M. Casanola-Martín, Y. Marrero-Ponce, M.T. Khan, S.B. Khan, F. Torrens, F. Pérez-Jimenez, A. Rescigno, C. Abad, Bond-based 2D quadratic fingerprints in QSAR studies: virtual and in vitro tyrosinase inhibitory activity elucidation, *Chem. Biol. Drug Des.*, 76 (2010) 538-545.

[38] J.A. Castillo-Garit, M.C. Vega, M. Rolon, Y. Marrero-Ponce, V.V. Kouznetsov, D.F. Torres, A. Gómez-Barrio, A.A. Bello, A. Montero, F. Torrens, F. Pérez-Gimenez, Computational discovery of novel trypanosomicidal drug-like chemicals by using bond-based non-stochastic and stochastic quadratic maps and linear discriminant analysis, *Eur. J. Pharm. Sci.*, 39 (2010) 30-36.

[39] R. Gozalbes, F. Barbosa, E. Nicolai, D. Horvath, N. Froloff, Development and validation of a pharmacophore-based QSAR model for the prediction of CNS activity, *ChemMedChem*, 4 (2009) 204-209.

[40] Y. Marrero-Ponce, A. Meneses-Marcel, O.M. Rivera-Borroto, R. García-Domenech, J.V. De Julian-Ortiz, A. Montero, J.A. Escario, A.G. Barrio, D.M. Pereira, J.J. Nogal, R. Grau, F. Torrens, C. Vogel, V.J. Aran, Bond-based linear indices in QSAR: computational discovery of novel anti-trichomonal compounds, *J. Comput. Aided Mol. Des.*, 22 (2008) 523-540.

[41] S.J. Patankar, P.C. Jurs, Classification of inhibitors of protein tyrosine phosphatase 1B using molecular structure based descriptors, *J. Chem. Inf. Comput. Sci.*, 43 (2003) 885-899.

[42] M. Murcia-Soler, F. Pérez-Gimenez, F.J. García-March, M.T. Salabert-Salvador, W. Díaz-Villanueva, P. Medina-Casamayor, Discrimination and selection of new potential antibacterial compounds using simple topological descriptors, *J. Mol. Graph. Model.*, 21 (2003) 375-390.

- [43] R.A. Cercos-del-Pozo, F. Pérez-Gimenez, M.T. Salabert-Salvador, F.J. García-March, Discrimination and molecular design of new theoretical hypolipemic agents using the molecular connectivity functions, *J. Chem. Inf. Comput. Sci.*, 40 (2000) 178-184.
- [44] F.J. Prado-Prado, X. García-Mera, H. González-Díaz, Multi-target spectral moment QSAR versus ANN for antiparasitic drugs against different parasite species, *Bioorg. Med. Chem.*, 18 (2010) 2225-2231.
- [45] Y. Rodríguez-Soca, C.R. Munteanu, J. Dorado, A. Pazos, F.J. Prado-Prado, H. González-Díaz, Trypano-PPI: a web server for prediction of unique targets in trypanosome proteome by using electrostatic parameters of protein-protein interactions, *J. Prot. Res.*, 9 (2010) 1182-1190.
- [46] F.J. Prado-Prado, H. González-Díaz, L. Santana, E. Uriarte, Unified QSAR approach to antimicrobials. Part 2: predicting activity against more than 90 different species in order to halt antibacterial resistance, *Bioorg. Med. Chem.*, 15 (2007) 897-902.
- [47] B.H. Junker, D. Koschützki, F. Schreiber, Exploration of biological network centralities with CentiBiN, *BMC Bioinformatics*, 7 (2006) 219.
- [48] D. Koschützki, CentiBiN Version 1.4.2, in, 2006, pp. CentiBiN Version 1.4.2, Centralities in Biological Networks © 2004-2006 Dirk Koschützki Research Group Network Analysis, IPK Gatersleben, Germany.
- [49] E. GL, C. KD, A.V. Jr, F.-S. RM, A new and rapid colorimetric determination of acetylcholinesterase activity, *Biochem. Pharmacol.*, 7 (1961) 88-95.
- [50] O. Ivanciuc, N. Oezguen, V.S. Mathura, C.H. Schein, Y. Xu, W. Braun, Using property based sequence motifs and 3D modeling to determine structure and functional regions of proteins, *Current Med. Chem.*, 11 (2004) 583-593.
- [51] C.H. Schein, O. Ivanciuc, W. Braun, Common physical-chemical properties correlate with similar structure of the IgE epitopes of peanut allergens, *J. Agric. Food Chem.*, 53 (2005) 8752-8759.
- [52] Y.M. Álvarez-Ginarte, Y. Marrero-Ponce, J.A. Ruíz-García, L.A. Montero-Cabrera, J.M. Vega, P. Noheda Marín, R. Crespo-Otero, F.T. Zaragoza, R. García-Domenech, Applying pattern recognition methods plus quantum and physico-chemical molecular descriptors to analyze the anabolic activity of structurally diverse steroids, *J. Comp. Chem.*, (2007).
- [53] A.H. Morales, J.E. Rodríguez-Borges, X. García-Mera, F. Fernández, M.N.D.S. Cordeiro, Probing the Anticancer Activity of Nucleoside Analogues: A QSAR Model Approach Using an Internally Consistent Training Set, *J. Med. Chem.*, 50 (2007) 1537-1545.
- [54] M. Fernández, J. Caballero, A. Tundidor-Camba, Linear and nonlinear QSAR study of N-hydroxy-2-[(phenylsulfonyl)amino]acetamide derivatives as matrix metalloproteinase inhibitors, *Bioorg. Med. Chem.*, 14 (2006) 4137-4150.
- [55] J. Caballero, M. Fernández, Linear and nonlinear modeling of antifungal activity of some heterocyclic ring derivatives using multiple linear regression and Bayesian-regularized neural networks, *J. Mol. Model.*, 12 (2006) 168-181.
- [56] H. González-Díaz, Y. González-Díaz, L. Santana, F.M. Ubeira, E. Uriarte, Proteomics, networks and connectivity indices, *Proteomics*, 8 (2008) 750-778.
- [57] H. Jeong, S.P. Mason, A.L. Barabasi, Z.N. Oltvai, Lethality and centrality in protein networks, *Nature*, 411 (2001) 41-42.
- [58] E. Estrada, Virtual identification of essential proteins within the protein interaction network of yeast, *Proteomics*, 6 (2006) 35-40.

- [59] P.A. Todd, K.L. Goa, Simvastatin. A review of its pharmacological properties and therapeutic potential in hypercholesterolaemia, *Drugs*, 40 (1990) 583-607.
- [60] O. Hernández-Perera, D. Pérez-Sala, J. Navarro-Antolín, R. Sánchez-Pascuala, G. Hernández, C. Díaz, S. Lamas, Effects of the 3-hydroxy-3-methylglutaryl-CoA reductase inhibitors, atorvastatin and simvastatin, on the expression of endothelin-1 and endothelial nitric oxide synthase in vascular endothelial cells, *J. Clin. Invest.*, 101 (1998) 2711-2719.
- [61] L. Santana, E. Uriarte, H. González-Díaz, G. Zagotto, R. Soto-Otero, E. Méndez-Álvarez, A QSAR model for in silico screening of MAO-A inhibitors. Prediction, synthesis, and biological assay of novel coumarins, *J. Med. Chem.*, 49 (2006) 1149-1156.
- [62] L. Santana, H. González-Díaz, E. Quezada, E. Uriarte, M. Yañez, D. Viña, F. Orallo, Quantitative structure-activity relationship and complex network approach to monoamine oxidase a and B inhibitors, *J. Med. Chem.*, 51 (2008) 6740-6751.
- [63] N.T. Nguyen, D.M. Cook, L.A. Bero, The decision-making process of US Food and Drug Administration advisory committees on switches from prescription to over-the-counter status: a comparative case study, *Clin. Therap.*, 28 (2006) 1231-1243.
- [64] A. Duardo-Sánchez, G. Patlewicz, A. López-Díaz, Current topics on software use in medicinal chemistry: intellectual property, taxes, and regulatory issues, *Curr. Top. Med. Chem.*, 8 (2008) 1666-1675.
- [65] H. González-Díaz, F. Prado-Prado, L.G. Pérez-Montoto, A. Duardo-Sánchez, A. López-Díaz, QSAR Models for Proteins of Parasitic Organisms, Plants and Human Guests: Theory, Applications, Legal Protection, Taxes, and Regulatory Issues, *Curr. Proteomics*, 6 (2009) 214-227.
- [66] H. González-Díaz, A. Duardo-Sánchez, F.M. Ubeira, F. Prado-Prado, L.G. Pérez-Montoto, R. Concu, G. Podda, B. Shen, Review of MARCH-INSIDE & Complex Networks prediction of Drugs: ADMET, Anti-parasite Activity, Metabolizing Enzymes and Cardiotoxicity Proteome Biomarkers Current Drug Metabolism, 11 (2010) 379-406.
- [67] H. González-Díaz, F. Romaris, A. Duardo-Sánchez, L.G. Pérez-Montoto, F. Prado-Prado, G. Patlewicz, F.M. Ubeira, Predicting drugs and proteins in parasite infections with topological indices of complex networks: theoretical backgrounds, applications, and legal issues, *Curr. Pharm. Des.*, 16 (2010) 2737-2764.
- [68] C. Yang, L.G. Valerio, Jr., K.B. Arvidson, Computational toxicology approaches at the US Food and Drug Administration, *Altern. Lab. Anim.*, 37 (2009) 523-531.
- [69] M.L. Raves, M. Harel, Y.P. Pang, I. Silman, A.P. Kozikowski, J.L. Sussman, Structure of acetylcholinesterase complexed with the nootropic alkaloid, (-)-huperzine A, *Nat. Struct. Biol.*, 4 (1997) 57-63.