



Published in final edited form as:

*Open Bioinforma J.* 2020 ; 13: 1–14.

## Machine Learning Neuroprotective Strategy Reveals a Unique Set of Parkinson Therapeutic Nicotine Analogs

Felipe Rojas-Rodríguez<sup>1,\*</sup>, Carlos Morantes<sup>2</sup>, Andrés Pinzón<sup>3</sup>, George E. Barreto<sup>4</sup>, Ricardo Cabezas<sup>1</sup>, Leonardo Mariño-Ramírez<sup>5</sup>, Janneth González<sup>1</sup>

<sup>1</sup>Departamento de Nutrición y Bioquímica, Pontificia Universidad Javeriana. Bogotá D.C, Republic of Colombia

<sup>2</sup>Departamento de Biología, Universidad Nacional de Colombia. Bogotá, Republic of Colombia

<sup>3</sup>Instituto de Genética, Universidad Nacional de Colombia, Bogotá, Republic of Colombia

<sup>4</sup>Department of Biological Sciences, University of Limerick, Limerick, Ireland

<sup>5</sup>National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 8600 Rockville Pike, Bethesda, MD 20894, USA

### Abstract

**Aims:** Present a novel machine learning computational strategy to predict the neuroprotection potential of nicotine analogs acting over the behavior of unpaired signaling pathways in Parkinson's disease.

**Background:** Dopaminergic replacement has been used for Parkinson's Disease (PD) treatment with positive effects on motor symptomatology but low progression and prevention effects.

---

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: (<https://creativecommons.org/licenses/by/4.0/legalcode>). This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

\*Address correspondence to the author at the Departamento de Nutrición y Bioquímica, Pontificia Universidad Javeriana. Bogotá D.C, Republic of Colombia; [arojas.r@javeriana.edu.co](mailto:arojas.r@javeriana.edu.co).

#### AUTHORS' CONTRIBUTIONS

FR-R developed the structural analysis for the ligands with the receptor, the manual reconstruction of the signaling pathway and the dataset for the agonist and antagonist ligands related to the receptor including the physicochemical featurization. CM developed the MCMC decomposition of the PI3K/AKT signaling pathway. CM and FR-R developed the iterative process for the architecture for the ANN and the corresponding training datasets. RC contribute to the writing and biological analysis of the results. AP and JG were major contributors to the biological significance of the results and improved the methods and applications along with the development of the project. GEB and LM contributed to the biological impact of the computational framework and applications. All authors contributed to analyzing the ANN data, manuscript writing and approved the final manuscript.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

#### ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

#### HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

#### CONSENT FOR PUBLICATION

Not applicable.

#### AVAILABILITY OF DATA AND MATERIAL

The datasets used and/or analyzed during the current study are available from the corresponding author [F.R.] on request.

#### SUPPLEMENTARY MATERIAL

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

Epidemiological studies have shown that nicotine consumption decreases PD prevalence through neuroprotective mechanisms activation associated with the overstimulation of signaling pathways (SP) such as PI3K/AKT through nicotinic acetylcholine receptors (*e.g.*  $\alpha 7$  nAChRs) and over-expression of anti-apoptotic genes such as Bcl-2. Nicotine analogs with similar neuroprotective activity but decreased secondary effects remain as a promissory field.

**Objective:** The objective of this study is to develop an interdisciplinary computational strategy predicting the neuroprotective activity of a series of 8 novel nicotine analogs over Parkinson's disease.

**Methods:** We present a computational strategy integrating structural bioinformatics, SP manual reconstruction, and deep learning to predict the potential neuroprotective activity of 8 novel nicotine analogs over the behavior of PI3K/AKT. We performed a protein-ligand analysis between nicotine analogs and  $\alpha 7$  nAChRs receptor using geometrical conformers, physicochemical characterization of the analogs and developed manually curated neuroprotective datasets to analyze their potential activity. Additionally, we developed a predictive machine-learning model for neuroprotection in PD through the integration of Markov Chain Monte-Carlo transition matrix for the 2 SP with synthetic training datasets of the physicochemical properties and structural dataset.

**Results:** Our model was able to predict the potential neuroprotective activity of seven new nicotine analogs based on the binomial Bcl-2 response regulated by the activation of PI3K/AKT.

**Conclusion:** Hereby, we present a robust novel strategy to assess the neuroprotective potential of biomolecules based on SP architecture. Our theoretical strategy can be further applied to the study of new treatments related to SP deregulation and may ultimately offer new opportunities for therapeutic interventions in neurodegenerative diseases.

## Keywords

Parkinson's disease; Nicotine analogs; PI3K/AKT; Machine learning; Artificial Neural Networks; Substantia Nigra

## 1. INTRODUCTION

Parkinson's Disease (PD) is the second most common neurodegenerative disorder, mainly affecting the population over 60 years of age [1]. As a complex multifactorial disease, PD is caused by the accumulation of unfolded proteins, genetic predisposition, mitochondrial dystrophies, epigenetic imbalance and environmental factors that increase the degeneration of dopaminergic neurons in the *substantia nigra* of the brain [2 – 7]. Traditional approaches of dopaminergic replacement, such as levodopa, carbidopa, and dopamine agonists (apomorphine, amantadine, cabergoline, pergolide, etc) restore direct or indirect dopaminergic supply, modulate dopaminergic neuron activity and stimulate postsynaptic receptors [2, 8]. However, classical therapeutics are directed towards treating PD symptoms (dyskinesia, motor fluctuations and systemic complications) in exchange for reducing disease progression coupled with a lack of long-term efficacy [1, 2, 9].

Epidemiological studies demonstrated an inverse association between cigarette smoking and PD risk [10 – 12]. Pre-clinical studies using *in vivo* and *in vitro* models showed that nAChR agonists protect nigrostriatal and other neuronal cell populations against cytotoxic damage, suggesting a neuroprotective activity [12 – 15]. nAChRs are ligand-gated ion channels capable to respond to several ligands such as acetylcholine and nicotine [16]. Classified between neuronal and muscle subtype receptors, only one pentameric subtype of nAChR ( $2\alpha1, \beta1, \gamma/\epsilon, \delta$ ) has been reported for muscles compared with the 12 subtypes in the neural system ( $\alpha2-10, \beta2-4$ ). Neuroprotective activity of nicotine and nicotine agonists has been linked to the activation of  $\alpha7$ -nAChRs [14, 17]. Highly expressed in dopaminergic neurons,  $\alpha7$ -nAChRs activation increases calcium flow and activates downstream SP such as ERK/MAPK, JAK2/STAT3, calmodulin and PI3K/AKT. SP activation has been related with neuronal survival, synaptic plasticity and decreased apoptosis due to the overexpression of cell survival proteins such as Bcl-x, CREB and Bcl-2 [17 – 19]. In parallel, tumor suppressor p53 directly regulates cell survival by modulating Bax:Bcl-2 expression, influencing the apoptotic fate of a cell in response to stress [20]. Bcl-2 has been suggested as a candidate for the nicotine-mediated neuroprotection activity in neurons. For example, nicotine induces the phosphorylation of Bcl-2 through either protein kinase C (PKC) and the MAPKs ERK1 and ERK2 or through PI3K-induced phosphorylation of AKT causing the inhibition of apoptosis in neuronal cultures [14, 17, 18, 21 – 25]. Additionally, overexpression of Bcl-2 Prevents neuronal death in *Mus musculus* models and Bcl-2 inhibition results in viability loss in dopaminergic cells through the induction of caspase-3 [26, 27].

## 2. METHODS

In this sense, nAChRs agonists can reduce secondary effects of classical therapies such as levodopa-induced dyskinesia, increased dopamine release and improve cytochrome P450 activity [28, 29]. Nicotine usage as a pharmacological agent is limited considering toxicity and addiction [30]. Nicotine analogs retaining neuroprotective activity, but reducing toxicity and addiction offer an excellent pharmacological alternative in PD [31 – 35]. We have previously the antioxidant potential of two nicotine analogs against rotenone-induced ROS generation in a PD *in vitro* model founding that 10  $\mu$ M of (E)-Nicotinaldehyde O-Cinnamyloxime was able to reduce superoxide anion and hydrogen peroxide production in neuronal SH-SY5Y cells treated with rotenone for 24h [10]. Consistently, other studies showed that nicotine analogs decrease superoxide anion generation and oxidative stress in rodents *Mus musculus* and *Rattus rattus*, as well as primates *Macaca fascicularis* and *Saimiri sciureus* through the activation of  $\alpha7$  nAChRs and PI3K/AKT [17, 35 – 38]. Computational methods have been used to infer the activity of candidate molecules but, due to the intrinsic PD complexity, is important to improve drug scanning efficiency by integrating structural and systemic data [39 – 40]. The development of holistic computational methodologies for SP and their modulatory mechanisms can lead to the discovery of new pharmacological agents [40, 41]. In this aspect, the use of Artificial neural networks (ANN) have been previously used in a simplified model of the Toll-like receptor signaling pathway, and the PI3K/AKT pathway in the context of cancer [42]. Nevertheless, the absence of integrative modeling in PD has led to a decrease in drug discovery efficiency. In the present study, we predicted the effect of 8 novel nicotine analogs as possible neuroprotective agents acting on

PI3K/AKT by integrating structural bioinformatic methods, SP manual reconstruction and ANN (Fig. 1).

### 3. RESULTS

#### 3.1. Geometry and Structure Analysis

Structures of (3R,5S)-1, methyl-5-(pyridine-3-yl) pyrrolidine-3-ol (A1), 3-(1,3-dimethyl-4,5-dihydro-1h-pyrazole-5-yl) pyridine (A2), 3-(3-methyl-4,5-dihydro-1h-pyrazole-5-yl) pyridine (A3), 3-(((2S,4R)-1,4-dimethylpyrrolidine-2-yl)) (A4), 3-(((2S,4R)-4-(fluoromethyl)-1-methylpyrrolidine-2-yl))pyridine (A5), 3-(((2S,4R)-4-methoxy-1-methylpyrrolidine-2-yl) pyridine (A6), 3-(((2S,3S)-1,3-dimethylpyrrolidine-2-yl) pyridine (A7) and 5-methyl-3-(pyridine-3-yl)-4,5-dihydroisoxazole (A8) were used for subsequent methods (Table 1). The neutral molecules optimized at the B3LYP/631G level and conformationally analyzed at the PM6 level showed rotations around C1-C1' bonds and bonds between radical atoms and rings. The minimum energy geometry found was used as input for subsequent analysis and the corresponding energy values (kJ/mol) for the local minimum geometrical conformers were reported in Table 1. Geometrical rotations around the bonds of the rings and the radicals were categorized as true minima of the potential energy of the surface based on the absence of imaginary vibrational frequencies.

#### 3.2. Nicotine Analogs Docking and $\alpha 7$ -nAChR Binding Interaction

$\alpha 7$  nAChR has 5 active pockets located at each subunit interactions of the protein complex. Considering that the homopentamer has 5 identical active pockets, we modeled the interaction within one of the pockets considering all the 5 pockets in the homopentamer subunits interaction region are identical. AChBP ligand-binding structure possesses conserved domain residues forming a narrow hydrophobic pocket including A (Tyr91), B (Trp145), C (Tyr184 and Tyr191), D (Trp53), E (Leu106, Gln114 and Leu116). Whereby A, B, and C are present in a different AChBP ligand-binding subunit than D and E (Fig. 2). Following the structural analysis, the nicotine analogs were set to interact with the active pocket within the interaction of the  $\alpha 7$  nAChR homo-subunits. Nicotine structure resulted in a geometrical docked conformation in which the pyrrole ring is oriented towards the C loop favoring Van der Waals interactions with amino acids TYR91, TRP145, and TYR184 of the chain A, in addition to LEU106 and TRP53 of the chain B (Table 2).

Docking models were performed with the top ligand conformations and the energy interaction values are presented in Table 3. Evaluated nicotine analogs, except for A1 and A3, have minor binding energy compared with nicotine (Table 3). Energetically minimized protein structures interacting with nicotine and analogs were used to calculate Root-Mean-Square Deviation of atomic positions (RMSD) in a comparative manner against the unbonded receptor and  $\alpha 7$  nAChR docked with nicotine (Table 3). Such an unbonded model was set as the minimized structure of  $\alpha 7$  nAChR without ligand on the active side.

#### 3.3. Clustering of Similarity

Two manually curated groups of compounds with known  $\alpha 7$  nAChR agonistic and antagonistic activities were used as input for the cluster. The first group consisted of

$\alpha 7$  nAChR agonists SAK3 [43], Nicotine, Acetylcholine, TC-1698 [44], PNU-282987 [45], DMXB [46], and ABT-107 [47]. The second group was composed of receptor  $\alpha 7$  antagonists and included methyllycaconitine [48], mecamylamine [49], neostigmine [50], anisodamine [51] and bupropion [52]. To the best of our knowledge, these are all the compounds reported with agonistic neuroprotective activity and agonistic activity over  $\alpha 7$  nAChR. In this sense, we set a positive response of PI3K/AKT for the first group and we hypothesized a null activation of PI3K/AKT SP or repression by other molecular mechanisms, and therefore no neuroprotective activity [53]. The same procedure previously mentioned was applied to obtain the optimized conformers for each group and all were categorized considering general descriptors of physicochemical properties [54, 55]. Surprisingly, the dendrogram of similarity was found to be related to the variables and the neuroprotective activity but with inconclusive scores for similarity (Supplementary Fig S1). Nevertheless, cluster values of similarity for PaDEL-Descriptors structural, physical and chemical values were used to enrich the known experimental data of the ligands.

### 3.4. Signaling Pathway Reconstruction and Predictive Model of Interaction-Response

The reconstructed model of PI3K/AKT SP presented both activation or inhibition of Bcl-2, as it occurs within the cell [20]. The network is composed of 28 nodes (proteins) that represent relevant components of the pathway ( $\alpha 7$  nAChR, PI3K, AKT, CREB, etc), 2 final states (activation/inhibition) and 43 interactions between the nodes (Fig. 3). The network was constructed using nodes with biological relevance associated with PI3K/AKT and P53 proliferative regulatory activity.

With previously mentioned clustering and PaDEL-Descriptors data, 1848 variables were obtained characterizing physico-chemical and pharmacological properties for each one of the molecules, both agonist/antagonist and nicotine analogs. Reducing the number of variables for subsequent analysis, principal component analysis (PCA, Supplementary Fig S2) and K-mean decomposition (Supplementary Fig S3) were used maintaining data variance explanation. First PCA three dimensions explained 98.9% of variance across the original dataset (Supplementary Fig S2) while the K-mean showed a variance explanation of 57.1% for two-dimensions and the corresponding coordinate values (Supplementary Fig S3).

Based on the PI3K/AKT architecture the resulting MCMC topology representing probabilities of transition within the network (Fig. 4). In this case, the resulting Markov Chain Monte Carlo (MCMC) model was able to quantify the transitions between the directed network resulting in a model capable to determine the ANN architecture. The most effective number of hidden layers of the ANN was determined by iterative topology generation and convergence was reached at 100.000 iterations. Such an approach resulted in a multi-perceptron ANN with 4 hidden layers with 5, 1, 4 and 5 nodes respectively (Fig. 5). In general, the resulting topology lack a canonical hidden layer architectures considering these models tend to reduce the number of nodes per layer across the model [56].

Training datasets from the PCA and K-mean were evaluated for each of the ANN learning with 1000 iterations with randomized series and reported values of misclassification (Table 4). Backpropagation with weight tracking coupled with PCA model was capable of generating consistent results. In this sense, the consistency of resilient backpropagation

with weight backtracking and PCA was 100% but resilient backpropagation without weight backtracking and K-mean showed 98%. These results suggest that both methods were able to predict the binomial output for the training and testing datasets. Nevertheless, throughout the iterations not all the predictions were identical in resilient backpropagation without weight backtracking and K-mean.

The resulting best ANN model implementing resilient backpropagation with weight backtracking is based on the evolving adaptation rule with a learning process based on error function [57].

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ * \Delta_{ij}^{(t-1)}, & \text{if } \frac{\partial E}{\partial \omega_{ij}}^{(t-1)} * \frac{\partial E}{\partial \omega_{ij}}^{(t)} > 0 \\ \eta^- * \Delta_{ij}^{(t-1)}, & \text{if } \frac{\partial E}{\partial \omega_{ij}}^{(t-1)} * \frac{\partial E}{\partial \omega_{ij}}^{(t)} < 0 \\ \Delta_{ij}^{(t-1)}, & \text{else} \end{cases} \quad (\text{A})$$

where  $0 < \eta^- < 1 < \eta^+$

$$\Delta \omega_{ij}^{(t)} = \begin{cases} -\Delta_{ij}^{(t)}, & \text{if } \frac{\partial E}{\partial \omega_{ij}}^{(t)} > 0 \\ +\Delta_{ij}^{(t)}, & \text{if } \frac{\partial E}{\partial \omega_{ij}}^{(t)} < 0 \\ 0, & \text{else} \end{cases} \quad (\text{B})$$

$$\omega_{ij}^{(t+1)} = \omega_{ij}^{(t)} + \Delta \omega_{ij}^{(t)}$$

and

$$\Delta \omega_{ij}^{(t)} = -\Delta \omega_{ij}^{(t-1)}, \text{ if } \frac{\partial E}{\partial \omega_{ij}}^{(t-1)} * \frac{\partial E}{\partial \omega_{ij}}^{(t)} < 0 \quad (\text{C})$$

Briefly, for the adaptation rule each update-value  $\Delta_{ij}^{(t)}$  changes depending on the sign for each partial derivative of  $\dot{u}_{ij}$  at point  $t$  (A). If the values are too big then the algorithm skips a local minimum and the function decreases by  $\bar{\eta}$ . Once  $\Delta_{ij}^{(t)}$  adapts,  $\Delta \omega_{ij}^{(t)}$  changes according to  $\frac{\partial E}{\partial \omega_{ij}}$  results (B). Positive values of the derivative mean an error increase; therefore,

the function is decreased by its  $\Delta_{ij}^{(t)}$ . Finally, the exception presented deals with changes in partial derivative sign (C) [57]. In that case, the step was too large and missing the minimum, so the previous  $\Delta \omega_{ij}^{(t)}$  is reverted. Finally, with the best trained structure method with PCA and Resilient backpropagation with weight backtracking and the MCMC derived ANN multi-perceptron architecture it was possible to predict the neuroprotective potential of the nicotine analogs. This model predicted A1, A2, A3, A4, A6, A7 and A8 as potential neuroprotective agents. In this aspect, our model predicted that 7 out of 8 analogs had putative neuroprotective activity due to the association of the binomial output of the ANN with the activation of Bcl-2.



## 4. DISCUSSION

### 4.1. Conformer Geometry and Ligand-Receptor Interaction

Even though research efforts have been done regarding the pharmaceutical potential of nicotine analogs, little assessment of some critical aspects of their conformational variability has been studied [15, 58]. Recently, the crystallographic structure of the extracellular domain of the nicotinic receptor  $\alpha 7$  in complex with epibatidine was reported [59]. Although interactions described for the complex agree with our results, it was found that THR146, CYS187, and CYS186 could also be involved in the protein complex due to molecular proximity. Docking with analog A5 had a high-affinity score ( $-4.7$ ) with respect to nicotine ( $-5.5$ ) (Table 3). A5 affinity can be attributed to the geometrical conformation establishing hydrogen bonds and radicals orientation towards TRP145, similar to epibatidine. The high affinity of epibatidine has been related to large inter-nitrogen distance previously reported as 4.6 Å [34].

Considering H-bond acceptors and donors number, analogs and nicotine had similar values. Specifically, A2, A3, A4 and A7 have two H-bond donors in comparison with three for A1, A5, A6 and A8. In terms of rotatable bonds, A5 and A6 differ from the single bond in nicotine considering the presence of radicals across the pyrrole ring (Table 1). Fluorine radical of A5 and the methyl group attached to the oxygen radical of A6 are rotation-capable bonds relative to the ring plane. Distinct similarity values in the clustering analysis for analogs A3 and A8 (Supplementary Fig S1) are not reflected in the values of energy affinity or interacting residues. Interestingly, A3 and A8 are the only analogs with a single methyl group addition to the pyrrolidine ring. In this sense, the group generated between A3 and A8 can be associated with the presence of the methyl group and the double bond with nitrogen in the pyrrolidine ring. Nevertheless, clustering was capable of generating patterns based on atom pair descriptors and molecular fingerprints, the model is inefficient to replicate the docking data. All analogs clustered with nicotine and TC-1698 with a node value of 0.4 in the dendrogram, but A3 and A8 showed significant similarity with TC-1698, suggesting agonistic and neuroprotective activity. Interestingly, A3 and A8 also have a null net charge of interaction with the protein (Table 3). Nevertheless, further *in vitro* and *in vivo* experiments are necessary to assess the neuroprotective capacity of the tested molecules and *in silico* modeling must be performed to establish a correlation between structural signatures and biological activity.

Besides affinity energy, the type and amount of interactions are crucial to determine the role of the docking in the protein function [13]. In this case, nicotine interacts with TRP53, Tyr91, Leu116, Trp145, Tyr184, Cys186, Cys187 and Tyr191 through hydrophobic interactions and without hydrogen bonds (Table 2). Compared with nicotinic receptor  $\alpha 7$  in complex with epibatidine interaction, only residues Leu106 and Gln114 were absent in the interaction [60]. Compared with the analogs, hydrophobic interactions with Leu106 were present in the analogs A1, A5, A6, and A8, but no hydrophobic interactions or hydrogen bonds were found with Gln114 (Table 2). The rest of the residues interacting with nicotine were found interacting at least once with any analog in similar geometrical orientations. Interestingly, none of the analogs presented the same interactions as nicotine, but the energy

changes were not higher than 0.4 kJ/mol, except for A5 with 0.8 kJ/mol. Such values of A5 could be attributed to the highly electronegative fluorine radical interacting with Thr146. An increase in Affinity energy could be associated with the presence of hydroxyl radical in the polar sidechain. Moreover, fluorine has been associated with secondary effects in humans and, if consumed in higher doses, has been reported to be toxic, thus suggesting that A5 is not a suitable neuroprotective agent [61, 62]. However, further studies must be performed in order to identify the role of specific residues in the activity of the complex and specific analog modifications effect over protein interactions.

Additionally, taking into account the comparison of RMSD values, after deuration of 25% of the misplaced atoms, unbonded protein and interacting complex with nicotine were lower than 0,1. Nicotine and analogs with the unbonded protein, RMSD validated our docking result (Table 3). RMSD for A2, A3 and A8 are higher than the rest of analogs when compared with the bounded complex with nicotine meaning that the geometrical modifications of the protein are higher. In this sense, A3 and A8 have the potential of interacting with the receptor similar to other neuroprotective compounds but changing the receptor response when compared with nicotine. The similar biological activity could be deduced also by considering the residues interacting, A3 and A8 as well as A6 and A7 present hydrogen bond interactions similar to nicotine (Table 2). With the structural and geometrical analysis is possible to conclude that all analogs except for A5 have neuroprotective potential and that A3 and A8 have an interesting behavior that must be addressed with *in vitro* and *in vivo* experiments. For A5, we present evidence of difference in energy affinity, interacting residues and fluoride radicals lead to the potential toxicity of A5.

Nevertheless, inconclusive relations between ligand structure, complex bonding and neuroprotective response occurred within the clustering. Even though consistent aggrupation occurred for all the neuroprotective compounds, antagonists generate a disperse distribution in the graph. The clustering FOF method was able to discriminate between biological effects but only strong aggrupation occurred only for neuroprotection positive molecules (Supplementary Figure S1). Acetylcholine was discarded considering divergence and ubiquitous neurotransmitter distribution coupled with no associated neuroprotective function [63]. Worth noticing, the potential toxicity of the other analogs should be addressed by experimental methods considering no information is available in cell culture or animal models neither *in silico* testing

## 4.2. Neuroprotective Prediction Modeling

Only a few studies have employed ANN in the study of signaling pathways [42, 64]. For instance, a 3-layered multi-perceptron with backpropagation learning and sigmoid activation function was implemented across 96 genes to identify biomarkers in children sarcomas [64]. Moreover, a simplified neural network comprising two microenvironmental input nodes (growth factor and death signal), and two phenotype output nodes (pro-growth and pro-death) was used to integrate environmental and molecular characteristics of cancer progression [42]. However, to the best of our knowledge, this is the first study that combines structural data, docking simulations, MCMC and ANN evaluating the modulation



of PI3K/AKT allowing the prediction of neuroprotective new compounds. By including a higher set of variables, analogs cluster near to nicotine for both PCA and K-mean decomposition analysis (Supplementary Figs 2 & 3). Nevertheless, the absence of strong groups in the PCA and K-mean analysis showed that synthetic data was not conclusive to predict the biological activity by themselves. With this in mind, the ANN model was essential to discriminate between the groups and underline the fundamental properties of the analogs based on the SP structure, using both PCA and K-mean synthetic values

SP are dynamic systems in which crosstalk with neighbor paths is essential to modulate signal intensity [65]. P53 negative feedback increased the robustness of the model by modifying Bcl-2 binomial output and information flow across the model supporting network regulation (Fig. 3). Additionally, ionotropic channels, such as  $\alpha 7$ -nAChR acting over PI3K/AKT lead to an activation without explicitly acknowledging the presence of JAK [14]. MCMC PI3K/AKT-derived model considered a matrix of transition representing network structure and probabilities of transition across PI3K/AKT leading to robust computational inferences [66]. By applying MCMC to the network topology it was possible to transfer the biological structure of the network to a quantitative model capable to modulate ANN architecture. The resulting non-canonical multi perceptron structure resulted as biological-based discrimination of ANN architectures through iterations using the MCMC matrix. Similar approaches such as the *drop-out* technique allow to drop units from the neural network during training hence preventing an excessive coadaptation and overfitting of the network [12]. Iterative processes converged at the minimal topology with a significant predictive capacity of the ANN model reducing layer and nodes number [67]. The hidden layers were locked according to the MCMC product (Fig. 4), ensuring that the subsequent topology represented the biological information contained in the SP. In this aspect, the synthetic iterative randomized data sets integrated with dimensional minimization and ANN, allowed the determination of the best predictive method reducing error and ensuring consistent predictions during training.

PCA and K-mean dimensional reduction showed to be an effective way to effectively decrease sample variability to a low number of synthetic values [68]. By reducing the number of variables from 1848 variables to a small non-autocorrelated dataset, PCA and K-mean reproduced clustering results. This approach was implemented in order to optimize the performance of the strategy and reduce overfitting of the trained ANN model. Through variable reduction, we increased the sensibility and reliability of the model to predict the binomial neuroprotective signal output. This data was used as training/testing for the ANN model evaluating some of the developed learning algorithms [69]. Identical split differences in misclassification for smallest absolute derivative, backpropagation and smallest learning rate with PCA and K-mean, and resilient backpropagation with weight backtracking with K-mean were identified (Table 4). Improvement was found for 25% misclassification for both resilient backpropagation with weight backtracking with PCA and resilient backpropagation without weight backtracking with K-mean. With this last method inconsistency in the predicting output was found meaning the output change through 2% of the iterations. Resilient backpropagation with weight backtracking with PCA resulted as the best combination for neuroprotection prediction, considering 100% of the predictions were consistent along with the iterations. The resulting model was based on the implementation

of an adaptation rule with a learning based on error values [57]. By predicting the binomial neuroprotective output of the ANN with synthetic PCA data from structural conformational, docking analysis and physicochemical featuring, made possible to determine that A1, A2, A3, A4, A6, A7 and A8 have potential neuroprotective agents. In general, the results are consistent with the experimental data available for some of the compounds [10, 47 – 52].

Even though this model considers the main aspects of SP response to external stimuli, some key considerations are necessary. Structurally, the number of reference compounds used was low, making it necessary to improve  $\alpha 7$ -nAChR agonist/antagonist data through identification and screening. Data restriction can lead to predicting bias of the model, overfitting or fixed prediction. In this matter, we recommend implementing both agonist and antagonist *in silico* and experimental screening to increase the training dataset. Although, the inclusion of synthetic molecules, such as SAK3, PNU-282987 and allosteric modulators allowed to improve predictions considering previously reported induction of neuroprotection through the modulation of PI3K/AKT/Bcl-2 [14]. Furthermore, the model can be used to address the secondary effects of nicotine analogs by increasing the dataset with addiction/reward molecules like nicotine and methyllycaconitine [48].

Correcting inner functions for noisy data needs to be coupled with further testing, essential to determine the true potential of nicotine analogs and antagonists for PD therapy. In this regard, we propose a mathematical *in silico* model that has to be used integrated with *in vitro* and *in vivo* data, either on dopaminergic neuronal cell lines, cerebral organoids or animal models. Database cluster enrichment with experimental data allows to increase the reference compounds and therefore reducing possible prediction bias and overfitting. Even so, the model strategy was in accordance with previously published experimental data [10, 47 – 52] proving to be robust enough as a testing method for further research. Worth noticing, SP are dynamic systems, therefore we suggest that continuous parameters need to be addressed to consider essential intermediate states [65]. Finally, this strategy opens the door to improve neuroprotection prediction in the future by integrating protein dynamics, partial differential equations and other machine learning approaches. These type of approaches are necessary to improve drug scanning efficiency in PD and any other disease founded upon SP.

### 4.3. Conformational Analysis and Protein Structure Preparation

8 novel nicotine analogs (Table 1) were used in the present study: (3R,5S)-1, methyl-5-(pyridine-3-yl) pyrrolidine-3-ol (A1), 3-(1,3-dimethyl-4,5-dihydro-1H-pyrazole-5-yl) pyridine (A2), 3-(3-methyl-4,5-dihydro-1H-pyrazole-5-yl) pyridine (A3), 3-(((2S,4R)-1,4-dimethylpyrrolidine-2-yl)) (A4), 3-((2S,4R)-4-(fluoromethyl)-1-methylpyrrolidine-2-yl)pyridine (A5), 3-((2S,4R)-4-methoxy-1-methylpyrrolidine-2-yl) pyridine (A6), 3-((2S,3S)-1,3-dimethylpyrrolidine-2-yl) pyridine (A7) and 5-methyl-3-(pyridine-3-yl)-4,5-dihydroisoxazole (A8). Molecular structures were sketched on Avogadro using MMFF94s forcefield, correcting atom type and chirality [70]. Calculations were carried out with Gaussian 16 [71], using B3LYP level of theory and cc-PVDZ basis set. In order to find the minimum energetic conformation, the structures were optimized at the DFT B3LYP/6–31G level. The conformational analysis was carried out on minimized rotating

bonds between pyrrole, derivate rings, and pyridine rings, including the bonds of all the radicals of the pyrrole ring for each ligand. The maximum number of conformers for each molecule was set to 30, with the 10 lowest energy conformations used for the docking simulations.

The crystal structure of  $\alpha 7$ -nAChR in complex with lobeline was obtained from the RCSB Protein Data Bank (PDB ID: 5AFJ) [60]. AutoDockTools [72] was used to assign polar hydrogens and add Gasteiger charges. The geometry of the receptor was optimized using the MM2 molecular mechanics force field. The neutral and ionized states of aliphatic amine and carboxylic acid groups of compounds to be docked were protonated and deprotonated separately.

#### 4.4. Molecular Docking

To determine the interaction between the studied molecules and the receptor, docking simulations were performed with AutoDock4 (version 4.2) [72]. The active site of the pentameric structure of  $\alpha 7$ -nAChR was defined as the interfaces between subunits that were within 12 Å from the geometric centroid of the ligand [60]. Within the 3 domains of the protein (extracellular, intracellular and transmembrane), the pocket is located in the extracellular side of the receptor with residues from loops A-C of the principal subunit and loops D-E of the complementary subunit. Default settings for small molecule-protein docking were used throughout the simulations.

Docked conformations were clustered according to the interacting energy combined with geometrical matching quality. The complexes with the best score were taken as the lead conformation for each compound. The correlation between key interactions obtained from the computational simulations and the interactions reported by the crystallographic structure of  $\alpha 7$ -nAChR in complex with lobeline was further explored. To further analyze the differences between the interaction of the analogs and the receptor, we calculated the RMSD values between analogs, nicotine and the receptor. RMSD was calculated using 85% of the atoms resulted after 3 outlier rejection cycles. The energetically minimized docked structures of the  $\alpha 7$ -nAChR receptor were made with Chimera [73], using 10000 steepest descent steps, 0.001 Å steepest descent steps size, 10 conjugate gradient steps, 0.001 Å conjugate gradient steps size, update per 10 intervals, gasteiger method for charge and Amber ff14sb algorithms for residues.

#### 4.5. Physicochemical Clustering and Dimensional Decomposition

To determine the features for each molecule and their structural similarity, physicochemical property predictions, classification and clustering of structures were performed on ChemmineR, R package [54]. Structural similarities between analogs were compared to a manually curated dataset of 7 agonists and 5 antagonists of  $\alpha 7$ -nAChR with reported neuroprotective activity [43 – 52]. This method allowed us to increase the size of the available data, robustness and reproducibility. In the dataset, agonists of  $\alpha 7$ -nAChR were associated with the induction of a neuroprotective pathway, either expressed through the induction of cell proliferation or apoptotic prevention. Antagonists were set to block either

the proliferative activity by PI3K/AKT of the signaling pathway or the associated  $\text{Ca}^{2+}$  mobilization of the receptor [51, 52, 74, 75].

Clustering analysis was based on molecular descriptors, such as molecular formula, molecular weight, atom frequency and functional groups. To determine the optimal cluster, a matrix of molecular descriptors for all molecules was used, categorizing the ligands according to the structural, physical and chemical variables. To increase the robustness of the clustering model, additional physicochemical data from *de-novo* featurization was included [55]. To ensure a proper physicochemical featurization and increase the number of variables available for further analysis, data from *de-novo* characterization using PaDEL-Descriptors [55] was also included. Dimensionality reduction of the dataset was done by principal component analysis (PCA) and k-mean decomposition. The mentioned approaches were used to identify the multiparametric data capable to describe the agonistic and antagonistic function and the relation between the activity and structure of the ligands in the dataset.

#### 4.6. Signal Reconstruction

To establish a reliable topology of the SP related with the neuroprotective capacity of nicotine, a manually reconstructed network for the PI3K/AKT SP was developed using information from KEGG [76 – 78] and PantherDB [79]. To ensure the biological coherence of the network, our model was enriched with additional protein-protein interactions that are present in the PI3K/AKT/mTOR interactive pathway [80]. Finally, this PI3K/AKT SP was integrated with a manually curated model of p53 to generate a negative feedback over Bcl-2. In this aspect, the model included 6 specific proteins (Myc, Bim, P53, Bax, Noxa, Puma) associated with the partial inhibition of the Bcl-2, which emulates the SP biological behavior [66].

#### 4.7. Activity Prediction and Interaction-Response Model

In order to elucidate the quantitative relationship between  $\alpha 7$ -nAChR receptor and the activation of Bcl-2, a Markov Chain Monte Carlo (MCMC) model for the whole network was implemented. This model was based on the reconstructed PI3K/AKT SP network and was generated using the R package *MCMCpack* [81]. A matrix of transitional stages predicting the most suitable path across the nodes was associated with the end stages of the MCMC that represented Bcl-2 expression as a binomial logical argument. The construction of the Artificial Neural Network model (ANN) was performed with the *NeuralNet* package in R [82]. To determine the canonical architecture for the ANN we used an optimization algorithm using 100.000 iterations using the MCMC transition matrix as an input. A stable multi-perceptron structure was set to have a minimal number of hidden layers with the capacity of representing the MCMC transition matrix, as shown below:

Briefly, a matrix  $x$  is used as input in the first layer of the multiperceptron. Each node of the hidden layer computes  $f_{(x)}$  to finally generate a binomial  $y$  output.

The dimensional reduction methods were used to train the ANN using the coordinates in the PCA and K-mean. One thousand randomized training datasets were selected and associated with the binomial activity of the ligand dataset. To optimize the predictive

capability of the model using the training datasets, four ANN algorithms were compared: Backpropagation with a learning rate of 0.001, Resilient backpropagation with weight backtracking, Resilient backpropagation without weight backtracking, smallest absolute derivative model and smallest learning rate. Additionally, for each possible combination of dimensional minimizations and ANN methods, the random data subsets used as training were iteratively generated, after excluding the testing values. To ensure the robustness of the prediction, values for misclassification error were obtained for each training and testing combination. In this aspect, the best model acquired was set to minimize the error to reduce the amount of false-positive predictions.

## CONCLUSION

We present a novel computational strategy as a novel quantitative approach used to predict the potential neuroprotective activity of 7 nicotine analogs in the context of PD therapeutics (A1, A2, A3, A4, A6, A7 and A8). To the best of our knowledge, this is the first machine learning neuroprotective computational strategy for ligand activity prediction over  $\alpha 7$ -nAChR activating PI3K/AKT/Bcl-2. By conformational analysis, molecular docking, the use of PaDEL variables and geometrical optimization, we obtained robust evidence suggesting nicotine analogs activity compared with nicotine.

The model also showed the potential of ANN integrations with a dimensional reduction to avoid overfitting by autocorrelation. Resilient backpropagation with weight backtracking with PCA showed to be the best combination for ANN training to validate the neuroprotective activity. Nevertheless, additional data is needed to ensure model robustness, decrease the error of misclassification and avoid fixing predictions. Nicotine analogs have promissory future as candidates for PD treatment if future improvements to these strategies are performed. Nevertheless, further studies must be performed to identify the potential of nicotine analogs as neuroprotective compounds and the application of these methods in neurodegenerative research.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

We would like to thank Alix Loaiza for providing the nicotine analogs structure and the DIT support team at Pontificia Universidad Javeriana for troubleshooting the server related issues.

## FUNDING

This research was funded in part by the Intramural Research Program of the National Institutes of Health (USA), National Library of Medicine (USA), and National Center for Biotechnology Information (NCBI) ZIA LM082713-06.

## LIST OF ABBREVIATIONS

<b>PD</b>	Parkinson's Disease
<b>DN</b>	Dopaminergic Neurons

<b>PI3K/AKT</b>	Phosphatidylinositol-4,5-bisphosphate 3- kinase/Protein kinase B
<b><math>\alpha 7</math>-nAChRs</b>	alpha7 nicotinic Acetylcholine Receptor
<b>MCMC</b>	Markov Chain Monte Carlo
<b>PCA</b>	Principal Component Analysis
<b>ANN</b>	Artificial Neural Networks
<b>SP</b>	Signaling Pathway

## REFERENCES

- [1]. Burbulla LF, Song P, Mazzulli JR, et al. Dopamine oxidation mediates mitochondrial and lysosomal dysfunction in Parkinson's disease. *Science* 2017; 357(6357): 1255–61. [10.1126/science.aam9080] [PubMed: 28882997]
- [2]. Cacabelos R Pharmacogenomics of Alzheimer's and Parkinson's diseases. *Neurosci Lett* 2018.133807 [10.1016/j.neulet.2018.09.018] [PubMed: 30236877]
- [3]. Goldman SM. Environmental toxins and Parkinson's disease. *Annu Rev Pharmacol Toxicol* 2014; 54: 141–64. [10.1146/annurev-pharmtox-011613-135937] [PubMed: 24050700]
- [4]. Lill CM. Genetics of Parkinson's disease. *Mol Cell Probes* 2016; 30(6): 386–96. [10.1016/j.mcp.2016.11.001] [PubMed: 27818248]
- [5]. Logroscino G The role of early life environmental risk factors in Parkinson disease: What is the evidence? *Environ Health Perspect* 2005; 113(9): 1234–8. [10.1289/ehp.7573] [PubMed: 16140634]
- [6]. Shulman JM, De Jager PL, Feany MB. Parkinson's disease: Genetics and pathogenesis. *Annu Rev Pathol* 2011; 6: 193–222. [10.1146/annurev-pathol-011110-130242] [PubMed: 21034221]
- [7]. Spatola M, Wider C. Genetics of Parkinson's disease: The yield. *Parkinsonism Relat Disord* 2014; 20(Suppl. 1): S35–8. [10.1016/S1353-8020(13)70011-7] [PubMed: 24262184]
- [8]. Ellis JM, Fell MJ. Current approaches to the treatment of Parkinson's Disease. *Bioorg Med Chem Lett* 2017; 27(18): 4247–55. [10.1016/j.bmcl.2017.07.075] [PubMed: 28869077]
- [9]. Berson A, Nativio R, Berger SL, Bonini NM. Epigenetic regulation in neurodegenerative diseases. *Trends Neurosci* 2018; 41(9): 587–98. [10.1016/j.tins.2018.05.005] [PubMed: 29885742]
- [10]. Jurado-Coronel JC, Loaiza AE, Díaz JE, et al. (E)-Nicotinaldehyde o-cinnamyloxime, a nicotine analog, attenuates neuronal cells death against rotenone-induced neurotoxicity. *Mol Neurobiol* 2018; 41: 587–98. [10.1007/s12035-018-1163-0]
- [11]. Ma C, Liu Y, Neumann S, Gao X. Nicotine from cigarette smoking and diet and Parkinson disease: A review. *Transl Neurodegener* 2017; 6: 18. [10.1186/s40035-017-0090-8] [PubMed: 28680589]
- [12]. Srinivasan R, Henley BM, Henderson BJ, et al. Smoking-relevant nicotine concentration attenuates the unfolded protein response in dopaminergic neurons. *J Neurosci* 2016; 36(1): 65–79. [10.1523/JNEUROSCI.2126-15.2016] [PubMed: 26740650]
- [13]. Huang S-Y, Zou X. Advances and challenges in protein-ligand docking. *Int J Mol Sci* 2010; 11(8): 3016–34. [10.3390/ijms11083016] [PubMed: 21152288]
- [14]. Quik M, Zhang D, McGregor M, Bordia T. Alpha7 nicotinic receptors as therapeutic targets for Parkinson's disease. *Biochem Pharmacol* 2015; 97(4): 399–407. [10.1016/j.bcp.2015.06.014] [PubMed: 26093062]
- [15]. Jurado-Coronel JC, Janneth G, Echeverria V, Barreto G. Nicotine analogues as potential therapeutic agents in Parkinson's disease by targeting nicotinic acetylcholine receptors (nAChRs) in astrocytes. 2015; 9.
- [16]. Yang T, Xiao T, Sun Q, Wang K. The current agonists and positive allosteric modulators of  $\alpha 7$  nAChR for CNS indications in clinical trials. *Acta Pharm Sin B* 2017; 7(6): 611–22. [10.1016/j.apsb.2017.09.001] [PubMed: 29159020]

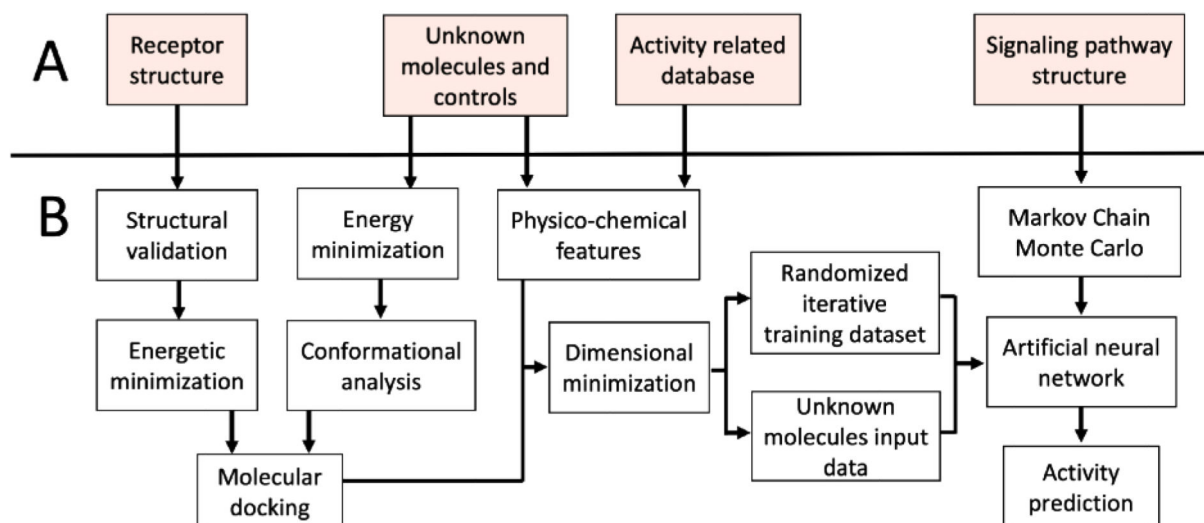


- [17]. Quik M, Perez XA, Bordia T. Nicotine as a potential neuroprotective agent for Parkinson's disease. *Mov Disord* 2012; 27(8): 947–57. [10.1002/mds.25028] [PubMed: 22693036]
- [18]. Dauer W, Przedborski S. Parkinson's disease: Mechanisms and models. *Neuron* 2003; 39(6): 889–909. [10.1016/S0896-6273(03)00568-3] [PubMed: 12971891]
- [19]. Zhu Z, Wang Y-W, Ge D-H, et al. Downregulation of DEC1 contributes to the neurotoxicity induced by MPP+ by suppressing PI3K/Akt/GSK3 $\beta$  pathway. *CNS Neurosci Ther* 2017; 23(9): 736–47. [10.1111/cns.12717] [PubMed: 28734031]
- [20]. Hemann MT, Lowe SW. The p53-Bcl-2 connection. *Cell Death Differ* 2006; 13(8): 1256–9. [10.1038/sj.cdd.4401962] [PubMed: 16710363]
- [21]. Buckingham SD, Jones AK, Brown LA, Sattelle DB. Nicotinic acetylcholine receptor signalling: Roles in Alzheimer's disease and amyloid neuroprotection. *Pharmacol Rev* 2009; 61(1): 39–61. [10.1124/pr.108.000562] [PubMed: 19293145]
- [22]. Cai Y, Zhang X, Zhou X, et al. Nicotine suppresses the neurotoxicity by MPP+/MPTP through activating  $\alpha$ 7nAChR/PI3K/Trx-1 and suppressing ER stress. *Neurotoxicology* 2017; 59: 49–55. [10.1016/j.neuro.2017.01.002] [PubMed: 28082123]
- [23]. Kawamata J, Suzuki S, Shimohama S. Enhancement of nicotinic receptors alleviates cytotoxicity in neurological disease models. *Ther Adv Chronic Dis* 2011; 2(3): 197–208. [10.1177/2040622310397691] [PubMed: 23251750]
- [24]. Shimohama S. Nicotinic receptor-mediated neuroprotection in neurodegenerative disease models. *Biol Pharm Bull* 2009; 32(3): 332–6. [10.1248/bpb.32.332] [PubMed: 19252273]
- [25]. Nakaso K, Ito S, Nakashima K. Caffeine activates the PI3K/Akt pathway and prevents apoptotic cell death in a Parkinson's disease model of SH-SY5Y cells. *Neurosci Lett* 2008; 432(2): 146–50. [10.1016/j.neulet.2007.12.034] [PubMed: 18201823]
- [26]. Robinson EJ, Aguiar SP, Kouwenhoven WM, et al. Survival of midbrain dopamine neurons depends on the Bcl2 factor Mcl1. *Cell Death Discov* 2018; 4: 107. [10.1038/s41420-018-0125-7] [PubMed: 30479840]
- [27]. Mostafapour SP, del Puerto NM, Rubel EW. Overexpression eliminates deprivation-induced cell death of brainstem auditory neurons. *J Neurosci* 2002; 22: 4670–4. [PubMed: 12040073]
- [28]. Quik M, O'Leary K, Tanner CM. Nicotine and parkinson's disease: Implications for therapy. *Mov Disord* 2008; 23(12): 1641–52. [10.1002/mds.21900] [PubMed: 18683238]
- [29]. Pandey S, Srivanthapoom P. Levodopa-induced dyskinesia: Clinical features, pathophysiology, and medical management. *Ann Indian Acad Neurol* 2017; 20(3): 190–8. [10.4103/aian.AIAN\_239\_17] [PubMed: 28904447]
- [30]. Yildiz D. Nicotine, its metabolism and an overview of its biological effects. *Toxicol* 2004; 43(6): 619–32. [10.1016/j.toxicol.2004.01.017] [PubMed: 15109883]
- [31]. Zouridakis M, Giastas P, Zarkadas E, Chroni-Tzartou D, Bregestovski P, Tzartos SJ. Crystal structures of free and antagonist-bound states of human  $\alpha$ 9 nicotinic receptor extracellular domain. *Nat Struct Mol Biol* 2014; 21(11): 976–80. [10.1038/nsmb.2900] [PubMed: 25282151]
- [32]. Wang S, Zhang C, Zhang M, et al. Activation of AMP-activated protein kinase  $\alpha$ 2 by nicotine instigates formation of abdominal aortic aneurysms in mice in vivo. *Nat Med* 2012; 18(6): 902–10. [10.1038/nm.2711] [PubMed: 22561688]
- [33]. Polindara-García LA, Vazquez A. Combinatorial synthesis of nicotine analogs using an Ugi 4-CR/cyclization-reduction strategy. *Org Biomol Chem* 2014; 12(36): 7068–82. [10.1039/C4OB00767K] [PubMed: 25087489]
- [34]. Abreo MA, Lin NH, Garvey DS, et al. Novel 3-Pyridyl ethers with subnanomolar affinity for central neuronal nicotinic acetylcholine receptors. *J Med Chem* 1996; 39(4): 817–25. [10.1021/jm9506884] [PubMed: 8632405]
- [35]. Pogocki D, Ruman T, Danilczuk M, Danilczuk M, Celuch M, Wałajtyś-Rode E. Application of nicotine enantiomers, derivatives and analogues in therapy of neurodegenerative disorders. *Eur J Pharmacol* 2007; 563(1–3): 18–39. [10.1016/j.ejphar.2007.02.038] [PubMed: 17376429]
- [36]. Akaike A, Takada-Takatori Y, Kume T, Izumi Y. Mechanisms of neuroprotective effects of nicotine and acetylcholinesterase inhibitors: role of  $\alpha$ 4 and  $\alpha$ 7 receptors in neuroprotection. *J Mol Neurosci* 2010; 40(1–2): 211–6. [10.1007/s12031-009-9236-1] [PubMed: 19714494]

- [37]. Baumann RJ, Mckean HE, Jameson HD, Turbek JA. Nicotine exposure and parkinson disease. *Am J Epidemiol* 1981; 114: 191–200. [10.1093/oxfordjournals.aje.a113182] [PubMed: 7304554]
- [38]. Lin N-H, Carrera GM Jr, Anderson DJ. Synthesis and evaluation of nicotine analogs as neuronal nicotinic acetylcholine receptor ligands. *J Med Chem* 1994; 37(21): 3542–53. [10.1021/jm00047a012] [PubMed: 7932582]
- [39]. Sliwoski G, Kothiwale S, Meiler J, Lowe EW Jr. Computational methods in drug discovery. *Pharmacol Rev* 2013; 66(1): 334–95. [10.1124/pr.112.007336] [PubMed: 24381236]
- [40]. Cho K-H, Wolkenhauer O. Analysis and modelling of signal transduction pathways in systems biology. *Biochem Soc Trans* 2003; 1: 1503–9. [10.1042/bst0311503]
- [41]. Berg JM, Rogers ME, Lyster PM. Systems biology and pharmacology. *Clin Pharmacol Ther* 2010; 88(1): 17–9. [10.1038/clpt.2010.69] [PubMed: 20562889]
- [42]. Kim E, Gerlee P, Anderson ARA. Using neural networks to bridge scales in cancer: Mapping signaling pathways to phenotypes. *bioRxiv* 2018; 324038 [10.1101/324038]
- [43]. Fukunaga K, Yabuki Y. SAK3-Induced neuroprotection is mediated by nicotinic acetylcholine receptors bt - nicotinic acetylcholine receptor signaling in neuroprotection. In: Akaike A, Shimohama S, Misu Y, Eds. Singapore: Springer Singapore 2018; pp. 159–71.
- [44]. Marrero MB, Papke RL, Bhatti BS, Shaw S, Bencherif M. The neuroprotective effect of 2-(3-pyridyl)-1-azabicyclo[3.2.2]nonane (tc-1698), a novel  $\alpha 7$  ligand, is prevented through angiotensin ii activation of a tyrosine phosphatase. *J Pharmacol Exp Ther* 2004; 309: 16–27. [PubMed: 14722323]
- [45]. Mata D, Linn DM, Linn CL. Retinal ganglion cell neuroprotection induced by activation of alpha7 nicotinic acetylcholine receptors. *Neuropharmacology* 2015; 99: 337–46. [10.1016/j.neuropharm.2015.07.036] [PubMed: 26239818]
- [46]. Meyer EM, King MA, Meyers C. Neuroprotective effects of 2,4-dimethoxybenzylidene anabaseine (DMXB) and tetrahydroaminoacridine (THA) in neocortices of nucleus basalis lesioned rats. *Brain Res* 1998; 786(1–2): 252–4. [10.1016/S0006-8993(97)00300-4] [PubMed: 9555043]
- [47]. Bordia T, McGregor M, Papke RL, Decker MW, McIntosh JM, Quik M. The  $\alpha 7$  nicotinic receptor agonist ABT-107 protects against nigrostriatal damage in rats with unilateral 6-hydroxydopamine lesions. *Exp Neurol* 2015; 263: 277–84. [10.1016/j.expneurol.2014.09.015] [PubMed: 25261754]
- [48]. Panagis G, Kastellakis A, Spyraiki C, Nomikos G. Effects of methyllycaconitine (MLA), an O07 nicotinic receptor antagonist, on nicotine- and cocaine-induced potentiation of brain stimulation reward. 2000; 149.
- [49]. Bacher I, Wu B, Shytle DR, George TP. Mecamylamine - a nicotinic acetylcholine receptor antagonist with potential for the treatment of neuropsychiatric disorders. *Expert Opin Pharmacother* 2009; 10(16): 2709–21. [10.1517/14656560903329102] [PubMed: 19874251]
- [50]. Zheng J-Q, He X-P, Yang A-Z, Liu C-G. Neostigmine competitively inhibited nicotinic acetylcholine receptors in sympathetic neurons. *Life Sci* 1998; 62(13): 1171–8. [10.1016/S0024-3205(98)00043-5] [PubMed: 9519798]
- [51]. Liu C, Shen F-M. Antishock effect of anisodamine involves a novel pathway for activating  $\alpha 7$  nicotinic acetylcholine receptor\*. *Crit Care Med* 2009; 1: 3(2): 634–41. [10.1097/CCM.0b013e31819598f5]
- [52]. Slemmer JE, Martin BR, Damaj MI. Bupropion is a nicotinic antagonist. *J Pharmacol Exp Ther* 2000; 295: 321–7. [PubMed: 10991997]
- [53]. Kihara T, Shimohama S, Sawada H, et al. Alpha7 nicotinic receptor transduces signals to PI3 kinase to block A $\beta$ -induced neurotoxicity. *J Biol Chem* 2001. [10.1074/jbc.M008035200]
- [54]. Cao Y, Charisi A, Cheng L-C, Jiang T, Girke T. ChemmineR: A compound mining framework for R. *Bioinformatics* 2008; 24(15): 1733–4. [10.1093/bioinformatics/btn307] [PubMed: 18596077]
- [55]. Yap CW. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* 2011; 32(7): 1466–74. [10.1002/jcc.21707] [PubMed: 21425294]
- [56]. Castro W, Oblitas J, Santa-Cruz R, Avila-George H. Multilayer perceptron architecture optimization using parallel computing techniques. *PLoS One* 2017; 12(12) e0189369. [10.1371/journal.pone.0189369] [PubMed: 29236744]

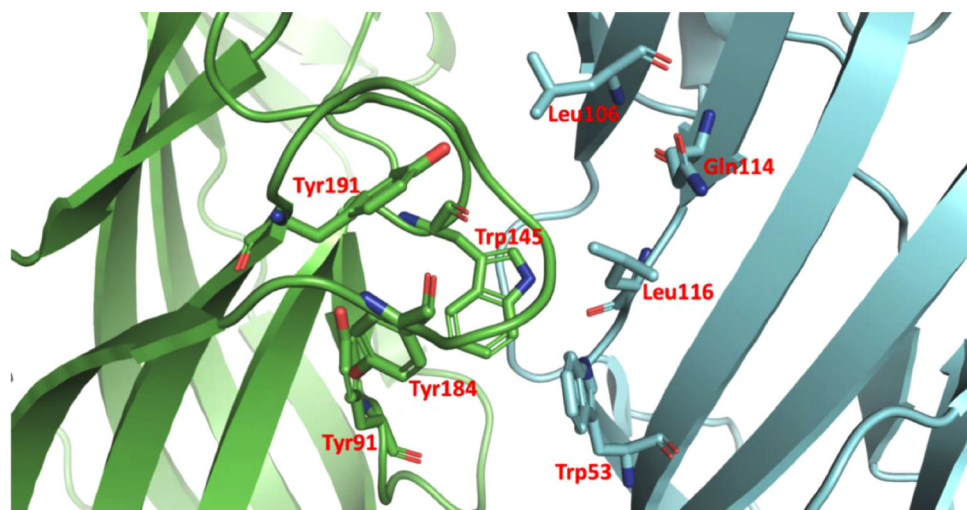
- [57]. Riedmiller M, Braun H. A direct adaptive method for faster backpropagation learning. RPROP Algorithm 1993; 1 [10.1109/ICNN.1993.298623]
- [58]. Gao J, Adam B-L, Terry AV Jr. Evaluation of nicotine and cotinine analogs as potential neuroprotective agents for Alzheimer's disease. *Bioorg Med Chem Lett* 2014; 24(6): 1472–8. [10.1016/j.bmcl.2014.02.008] [PubMed: 24581918]
- [59]. Thompson AJ, Metzger S, Lochner M, Ruepp MD. The binding orientation of epibatidine at  $\alpha 7$  nACh receptors. *Neuropharmacology* 2017; 116: 421–8. [10.1016/j.neuropharm.2017.01.008] [PubMed: 28089847]
- [60]. Li S-X, Huang S, Bren N, et al. Ligand-binding domain of an  $\alpha 7$ -nicotinic receptor chimera and its complex with agonist. *Nat Neurosci* 2011; 14(10): 1253–9. [10.1038/nn.2908] [PubMed: 21909087]
- [61]. Dhar V, Bhatnagar M. Physiology and toxicity of fluoride. *Indian J Dent Res* 2009; 20(3): 350–5. [10.4103/0970-9290.57379] [PubMed: 19884722]
- [62]. Ullah R, Zafar MS, Shahani N. Potential fluoride toxicity from oral medicaments: A review. *Iran J Basic Med Sci* 2017; 20(8): 841–8. [10.22038/IJBMS.2017.9104] [PubMed: 29085574]
- [63]. Picciotto MR, Higley MJ, Mineur YS. Acetylcholine as a neuromodulator: Cholinergic signaling shapes nervous system function and behavior. *Neuron* 2012; 76(1): 116–29. [10.1016/j.neuron.2012.08.036] [PubMed: 23040810]
- [64]. Tong DL, Boocock DJ, Dhondalay GK, Lemetre C, Ball GR. Artificial neural network inference (ANNI): A study on gene-gene interaction for biomarkers in childhood sarcomas. *PLoS One* 2014; 9(7): e102483. [10.1371/journal.pone.0102483] [PubMed: 25025207]
- [65]. Kholodenko BN. Cell-signalling dynamics in time and space. *Nat Rev Mol Cell Biol* 2006; 7(3): 165–76. [10.1038/nrm1838] [PubMed: 16482094]
- [66]. Montero J, Letai A. Why do BCL-2 inhibitors work and where should we use them in the clinic? *Cell Death Differ* 2017; 25: 56. [10.1038/cdd.2017.183] [PubMed: 29077093]
- [67]. Beer M Neural network based Monte Carlo simulation of random processes. 2005.
- [68]. Celebi ME, Kingravi HA, Vela PA. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst Appl* 2013; 40: 200–10. [10.1016/j.eswa.2012.07.021]
- [69]. Lancashire LJ, Lemetre C, Ball GR. An introduction to artificial neural networks in bioinformatics--application to complex microarray and mass spectrometry datasets in cancer studies. *Brief Bioinform* 2009; 10(3): 315–29. [10.1093/bib/bbp012] [PubMed: 19307287]
- [70]. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform* 2012; 4(1): 17. [10.1186/1758-2946-4-17] [PubMed: 22889332]
- [71]. Frisch MJ, Trucks GW, Schlegel HB, et al. Gaussian 16 Revision C01. 2016.
- [72]. Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and autoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 2009; 30(16): 2785–91. [10.1002/jcc.21256] [PubMed: 19399780]
- [73]. Pettersen EF, Goddard TD, Huang C, Couch G, Greenblatt DM, Meng E. UCSF Chimera—A visualization system for exploratory research and analysis. 2004; 25: pp. (13)1605–2.
- [74]. Cheng Q, Yakel JL. Activation of  $\alpha 7$  nicotinic acetylcholine receptors increases intracellular cAMP levels *via* activation of AC1 in hippocampal neurons. *Neuropharmacology* 2015; 95: 405–14. [10.1016/j.neuropharm.2015.04.016] [PubMed: 25937212]
- [75]. Arias HR, López JJ, Feuerbach D, Fierro A, Ortells MO, Pérez EG. Novel 2-(substituted benzyl)quinuclidines inhibit human  $\alpha 7$  and  $\alpha 4\beta 2$  nicotinic receptors by different mechanisms. *Int J Biochem Cell Biol* 2013; 45(11): 2420–30. [10.1016/j.biocel.2013.08.003] [PubMed: 23954208]
- [76]. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000; 28(1): 27–30. [10.1093/nar/28.1.27] [PubMed: 10592173]
- [77]. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017; 45(D1): D353–61. [10.1093/nar/gkw1092] [PubMed: 27899662]

- [78]. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 2016; 44(D1): D457–62. [10.1093/nar/gkv1070] [PubMed: 26476454]
- [79]. Thomas PD, Campbell MJ, Kejariwal A, et al. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res* 2003; 13(9): 2129–41. [10.1101/gr.772403] [PubMed: 12952881]
- [80]. Ersahin T, Tuncbag N, Cetin-Atalay R. The PI3K/AKT/mTOR interactive pathway. *Mol Biosyst* 2015; 11(7): 1946–54. [10.1039/C5MB00101C] [PubMed: 25924008]
- [81]. Park JH, Quinn K. MCMCpack: Markov chain Monte Carlo in R. 2011; 42.
- [82]. Günther F, Fritsch S. Neuralnet: Training of Neural Networks. *R Journal* 2010; 2(1): 30–8.



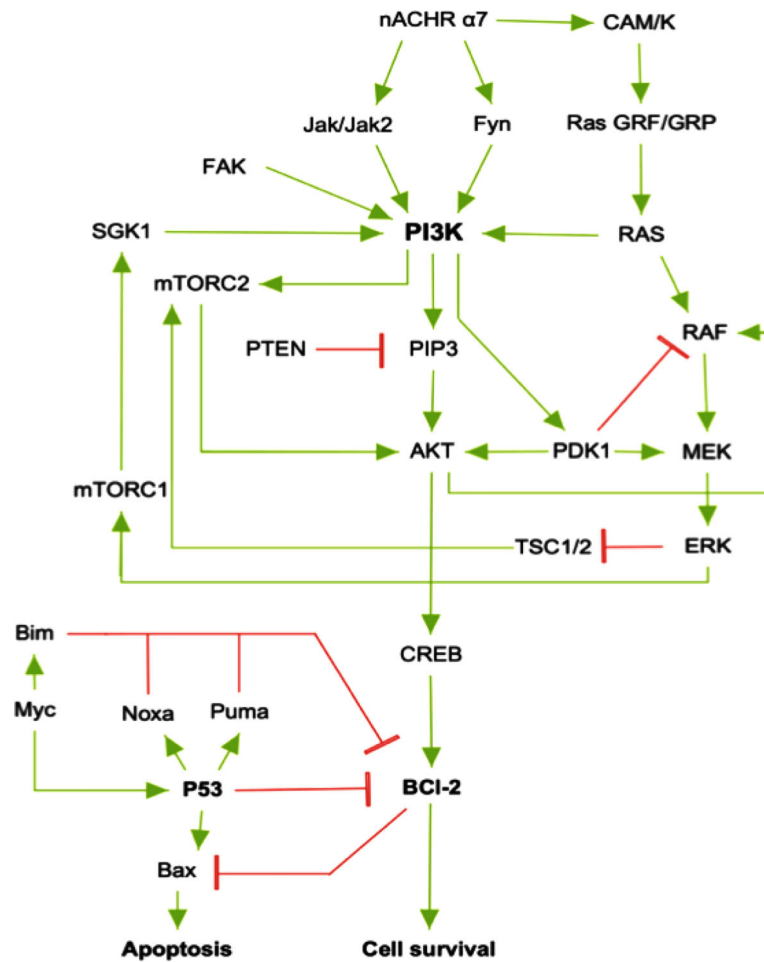
**Fig. (1).**

Diagrammatic workflow of the computational strategy for ligand prediction over SP architecture. The pipeline for the strategy implementation uses a receptor crystallographic structure, structure of the unknown ligands, known ligands data interacting with the receptor of interest and SP architecture. The SP must interact with the receptor by signaling cascade of secondary messengers and be triggered by the interactions of known and unknown ligands with the receptor. **A)** Input of biological information **B)** Computational methods to predict the activity of a series of ligands. Subgroups of computational methods aim to represent biological stages of the SP response to external ligands: 1) Structural validation, energy minimization, conformational search and molecular docking stand as protein ligand complex and 2) Markov Chain Monte Carlo and Artificial neural network as methods to generate a quantitative model with SP architecture embedded. Physico-chemical featuring, dimensional minimization and training/testing datasets were implemented to integrate the structural data with the systemic model of the SP and the machine learning model.

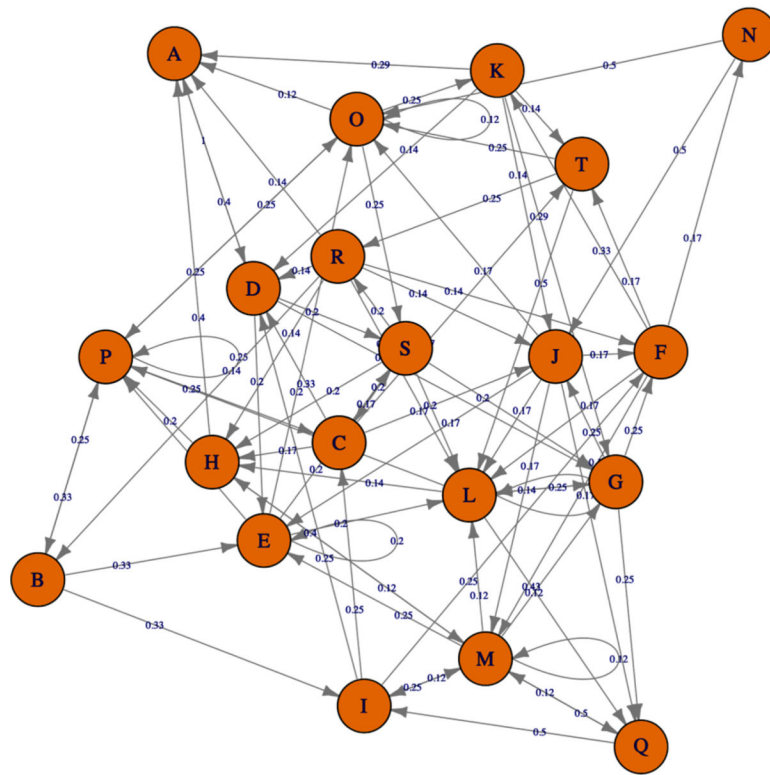


**Fig. (2).** Structural composition of the active site of  $\alpha 7$  nAChR. Green and blue represent the two  $\alpha 7$  nAChRs subunits composing the active pocket. All  $\alpha 7$  nAChRs have 5 identical active sites across the intersections of the  $\alpha 7$  homosubunits. Additionally, the localization of the residues composing the active site of the pentameric  $\alpha 7$  nAChR is presented highlighting the geometrical orientation.

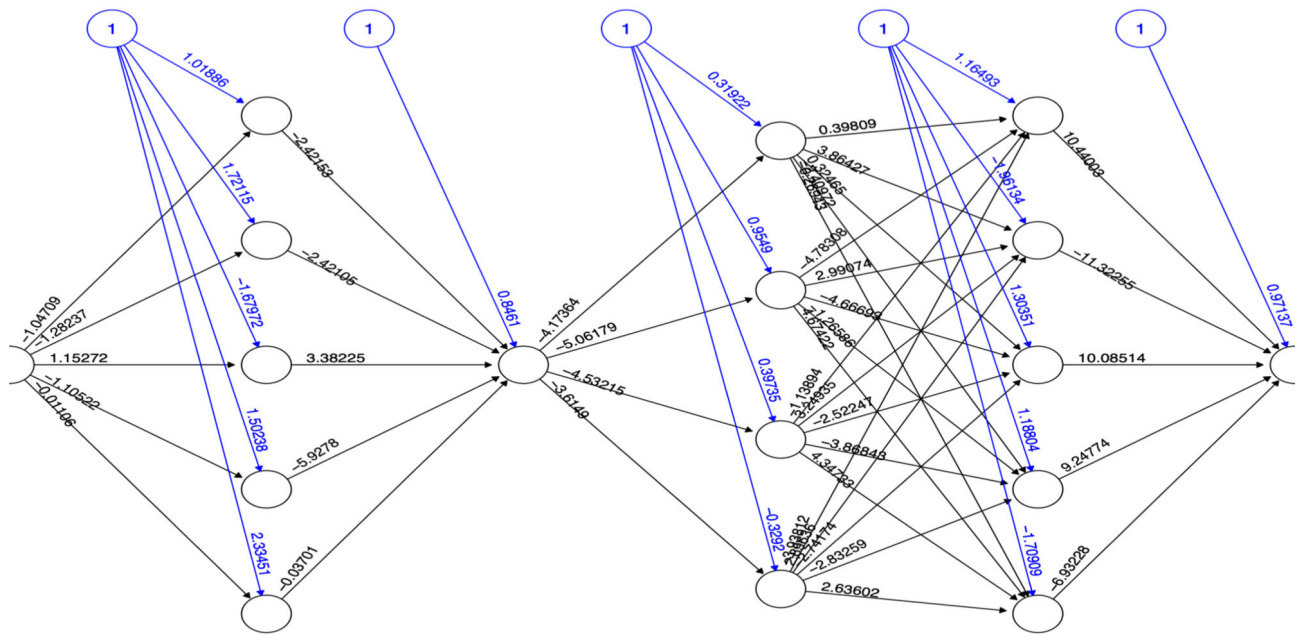




**Fig. (3).** PI3K/AKT Signaling pathway network reconstruction. Each node represents the proteins selected for the SP modulating the activity/expression of Bcl-2. Positive response to proliferation and cell survival was modulated by PI3K/ATK core SP and repression of Bcl-2 was added to the model by including P53. Such a network has a repressive activity over the binomial output. Worth noticing, the network is directional and with feature specific edges, meaning that green edges have a positive modulation of the ending node and red have a negative regulation of the nodes across the network.



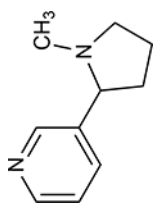
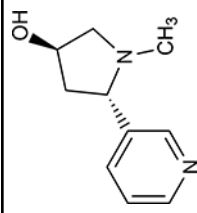
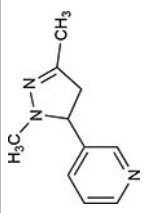
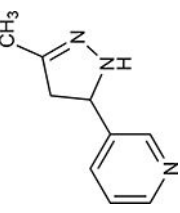
**Fig. (4).** Graphical representation of the probability transitions in the MCMC model of the PI3K/AKT directed network. Optimization was reached when the model stabilizes the probabilities between nodes. In general, each node represents a state involved in the SP and the edges in the graph establish a relationship between the states.



**Fig. (5).** ANN multi-perceptron architecture obtained with the MCMC SP topology. The layers and the number of neurons related to each one of the layers are correlated to the architecture of the MCMC optimal topology. The output of the ANN was set as an activation or repression binomial response.

Nicotine and nicotine analogs studied. 2D structure of each one of the ligands and energy values of each compound (kJ/mol) are presented. Structural similarity with nicotine was based upon energy values and the pyridine and pyrrolidine rings present in their structures. Nicotine analogs have modifications across the pyrrolidine ring of the nicotine backbone focusing in the addition of methylations, hydroxyl groups or highly electronegative elements. All the corresponding structures were provided by the laboratory of chemistry at the Pontificia Universidad Javeriana. Further analysis of all nicotine analogs will reference the ID provided ranging from A1 to A8.

Table 1.

Molecule	IUPAC Name	Structure	Minimum Energy Value (kJ/mol)
Nicotine	Nicotine		105.636
A1	(3R,5S)-1, methyl-5-(pyridine-3-yl) pyrrolidine-3-ol		199.330
A2	3-(1,3-dimethyl-4,5-dihydro-1H-pirazole-5-yl) pyridine		120.593
A3	3-(3-methyl-4,5-dihydro-1H-pirazole-5-yl) pyridine		82.944

Molecule	IUPAC Name	Structure	Minimum Energy Value (kJ/mol)
A4	3-((2S,4R)-1,4-dimethylpyrrolidine-2-yl)pyridine		130.095
A5	3-((2S,4R)-4-(fluoromethyl)-1-methylpyrrolidine-2-yl)pyridine		111.131
A6	3-((2S,4R)-4-methoxy-1-methylpyrrolidine-2-yl)pyridine		216.058
A7	3-((2S,3S)-1,3-dimethylpyrrolidine-2-yl)pyridine		127.645
A8	5-methyl-3-(pyridine-3-yl)-4,5-dihydroisoxazole		174.670

Diagrammatic representation of receptor interaction, residues involved in the receptor interaction with analogs and nicotine. Dark shaded cells mean the presence of the interaction between the ligand and the receptor. Discrimination between hydrogen bonds and hydrophobic interactions are shown.

**Table 2.**

Molecule	Hydrogen Bonds					Hydrophobic Interactions								
	Tyr91	Thr145	Tyr184	Tyr191	Trp53	Tyr91	Leu106	Leu116	Trp145	Thr146	Tyr184	Cys186	Cys187	Tyr191
Nicotine														
A1														
A2														
A3														
A4														
A5														
A6														
A7														
A8														



**Table 3.**

Energy and RMSD values for the molecular docking between analogs and nicotine. For RMSD there is a differentiation between  $\alpha 7$  nAChRs interacting with nicotine and  $\alpha 7$  nAChRs (unbonded) receptor without ligand. For the chimera optimization, the charges for the steepest descent minimization of the ligand were manually added. RMSD values represent the average distance between atoms, in this case, backbone atoms, between protein structures. The lower the RMSD value, the more geometrical similarity between the proteins.

Molecule	Charge of Interaction	Energy of Interaction (kJ/mol)	RMSD	
			Unbonded	Nicotine
Nicotine	+1	-5.5	0.049	-
A1	+1	-5.6	0.069	0.033
A2	0	-5.5	0.013	0.054
A3	0	-5.6	0.039	0.075
A4	+1	-5.1	0.053	0.014
A5	+1	-4.7	0.052	0.030
A6	+1	-5.5	0.047	0.016
A7	+1	-5.5	0.051	0.020
A8	0	-5.5	0.020	0.052

Values of misclassification error for all different ANN methods and dimensional minimization approaches. Normalized data is shown (0–1).

**Table 4.**

Learning algorithm	Dimensional reduction method	
	Principal Component Analysis	K-mean
Backpropagation	0.5	0.5
Resilient backpropagation with weight backtracking	0.25	0.5
Resilient backpropagation without weight backtracking	0.375	0.25
Smallest absolute derivative	0.5	0.5
Smallest learning rate	0.5	0.5