MP2 and DFT Calculations of Ligand Binding in Acetylcholine Binding Protein's Aromatic Box

Erin Elizabeth Carter

Department of Chemistry

Rhodes College

Memphis, Tennessee

2012

Submitted in partial fulfillment of the requirements for the

Bachelor of Science degree with Honors in Chemistry

This Honors paper by Erin Carter has been read and approved for Honors in Chemistry.

Dr. Mauricio Cafiero

Project Advisor

_____

Dr. Loretta Jackson-Hayes

Second Reader

_____

Dr. Jonathan Fitz Gerald

Extra-Departmental Reader

_____

Dr. Darlene Loprete

Department Chair

_____

CONTENTS

ILLUSTRATIONS AND TABLES

ABSTRACT

MP2 and DFT Calculations of Ligand Binding in Acetylcholine Binding Protein's Aromatic Box

by

Erin Elizabeth Carter

Acetylcholine Binding Protein's active site, called the Aromatic Box for its five aromatic amino acid components, is a receptor for nicotine that closely mimics $\langle_7$ Nicotinic Acetylcholine Receptors (nAChRs) in the brain. In this study, several experiments were conducted in order to design novel drug candidates that capitalize on the dispersion-dominated Aromatic Box. To better understand how charge interactions are involved in ligand binding, the migration of partial atomic charges were evaluated for several component structures of the active site using both correlated WFT and select DFT methods and variety of basis sets. Additional ligands (morphine, cocaine, and galantamine) were docked in the Aromatic Box and their interaction energies with the residues were calculated to further map the ligand binding behavior of the Aromatic Box. Electrostatic potential maps were generated to visually evaluate the known ligands' binding efficacy. Using information from these studies, novel drug candidate molecules have been designed. The interaction energies between these candidates and the Aromatic Box were evaluated. The results from this *in silico* experiment show that bicyclic molecules with heteroatoms bind more effectively in the Aromatic Box than the known

ligands studied. Further improvement can be gained through the addition of hydroxyl and cyano functional groups around the bicylcic portion of the drug candidates. Quantitative Structure Activity Relationship (QSAR) calculations were used to draw correlations between the molecular properties of the known ligands and their literature potency values. Equations were calibrated using known ligands' QSAR data to predict the potencies of the novel drug candidates.

I give permission for public access to my Honors paper and for any copying or digitization to be done at the discretion of the College Archivist and/or the College Librarian.

Signed_____

Erin Elizabeth Carter

Date _____

1. Introduction

Neurodegenerative diseases are among the most pharmacologically elusive of fatal ailments. Both Alzheimer's Disease (AD) and Parkinson's Disease (PD) have no known cure, and current treatment strategies leave much to be desired. Both diseases are characterized by progressive failure of the brain.[1] Symptoms include cognitive, motor, and behavioral impairments that intensify as the disease progresses, ultimately leading to an early death.[1,2] Most of the manifestations of PD are degenerations of the parts of the brain that control motor movements and AD's degeneration is mostly cognitive; it affects the patient's memory and ability to think, learn, communicate, and perform daily activities.[1,2] Neurodegeneration in both cases is gradual, and those afflicted eventually become severely disabled. In both AD and PD, slowing down or halting neuronal cell death is one of the primary focuses of treatment. However, this only provides symptomatic relief and is not a cure.

This discrepancy is due to cell death of midbrain dopamine-producing neurons in the substantia nigra (SN).[3] Specific activation of a variety of nicotinic receptors has been associated with the release of dopamine into the striatum, where PD's neuronal loss is concentrated. While the exact cause of the cell death is unknown, several symptomatic drugs have been created in hopes of slowing the progression of the disease. However, these treatments come with negative side effects, which may be ascribed to a lack of ligand specificity with the receptors in question.

Relatively new treatments attempt to take advantage of the nicotinic cholinergic system in order to prevent and slow neurodegeneration as well as maintain striatal dopamine levels, which are $\geq$ 70% lower than normal in PD patients at the time of

diagnosis.[3,4] Nicotinic acetylcholine receptors (nAChRs) describe a family of proteins that are found in the central nervous system, skeletal tissue, and peripheral nervous system.[4] They are all pentameric ligand-gated ion channels (LGICs) of the Cys-loop receptor family of membrane proteins.[5] Stimulation of several nAChR subtypes has shown benefits in the treatment of PD and AD.[4] Two of the most abundant subtypes in the central nervous system, which are of particular interest for PD and AD drug design, are the α_7 and $\alpha_4\beta_2$ receptors.[5] Acetylcholine Binding Protein (AChBP) is a CNS protein that was first isolated from neural glia cells in a freshwater snail, *Lymnaea stagnalis* (Ls-AChBP) (Figure 1).[5,6] This particular organism has been useful to neurobiology for over a decade, as the components of its CNS are relatively large scale and have many conserved functions.[6] Ls-AChBP performs an analogous function to human nAChRs, regulating dopamine and acetylcholine levels for use in synaptic transmission.[6] Additionally, Ls-AChBP is easily scaled in large quantities for use in crystallizing the structure in complex with a variety of ligands, which gives insight into the binding properties of nAChRs.[5] Overall, these qualities make AChBP a good model or "template" for the nAChR family of proteins.[7]

Figure 1: Ribbon structure of Ls-AChBP in complex with nicotine.[8]

Smit et. al. showed that the ligand binding region of AChBP is homologous to the ligand binding region of human nAChRs in general, but that it is virtually identical to the ligand binding region in the α_7 subtypes.[6] While there is a need for drugs that target specific nAChR subtypes, AChBP is an accepted model for the α_7 subtype. nAChRs are distinguished by five distinct subunits, each of which has an N-terminal extracellular domain, and a C-terminal domain, which contain the transmembrane components and LGIC pore.[6] AChBP has a moderate to high overall level of conservation with the nAChR superfamily.[6] The major difference is that AChBP lacks the C-terminal. However, these elements generally remain static upon ligand binding, so they do not significantly affect the ability of AChBP to act as a homologue to nAChRs.[6] The active sites of these proteins are located in the N-terminus and are characterized by a single aromatic residue-lined pocket at the junction of two subunits.[5,6] AChBP has been shown to contain all of the amino acid residues needed to mimic the active site, most closely resembling the α_7 nAChR subtype.[6] The active site is often referred to as the Aromatic

Box because its spatial organization loosely resembles a cube, with three tyrosine molecules on the bottom and two adjoining side faces of the cube and two tryptophan molecules on the remaining two adjoining side faces.

AChBP and α_7 nAChRs have high affinities for nicotine and acetylcholine, the endogenous ligand, though AChBP is thought to show higher affinity for nicotine.[5] Both short-term and chronic administration of nicotine has been shown to prevent SN neuronal loss to some degree in mouse and primate models, leading to the belief that nicotine has some neuroprotective properties.[3] Stimulation of α_7 nAChRs has some anti-inflammatory qualities and is also an established target for AD, PD, epilepsy, chronic pain, and addiction.[5] Akdemir et. al. notes that AChBPs have been successfully used as structural templates with regards to the α_7 and $\alpha_4\beta_2$ nAChR subtypes in order to describe their selectivity for a variety of known nAChR ligands. These include, but are not limited to, nicotine, morphine, galantamine, and cocaine. Galantamine is a symptomatic AD drug that is currently on the market under the brand names Reminyl® or Razadyne®. While galantamine helps relieve some of the symptoms of AD, it can present any of several unpleasant side effects. In researching this active site, the present study evaluated new drug candidates and compared them with several types of data from the known ligands in the hopes of improving galantamine and other symptomatic AD and PD drugs that bind to α_7 and $\alpha_4\beta_2$ nAChR receptors.

Many research groups use automated protein-ligand docking studies as a means for finding suitable molecules to propose as drug candidates. These studies are typically performed using a large database of small molecules with a variety of functional groups and geometries.[9] These molecules are quickly optimized in the active site in question

using a fast, inexpensive method.[9] Depending on the method used, the results of these studies may not necessarily yield very accurate binding energies. Instead of using the automated approach, the drug design in this study will focus on mapping the internal electronic structure of the active site and understanding how known substrates bind. This method is especially useful for AChBP's active site, since the interactions in the Aromatic Box are predominantly weak induction and dispersion forces, which the commonly used methods do not describe well.

In order to create novel drug candidates that effectively capitalize on these weak interactions, several different computational experiments were used. First, the charge migration in the Aromatic Box upon nicotine binding was explicitly evaluated, tracking the partial charge on each of the atoms in the ligand/active site complex. The interaction energies of known ligands were evaluated via interaction energies in ligand/residue pairs. Next, electrostatic potential maps were generated for each of the known ligands as well as the empty Aromatic Box in order to visually show the areas of interest for drug design. Several drug candidates were designed based upon this information, and their interaction energies were calculated. Finally, several molecular properties were analyzed for all of the known ligands and drug candidates; these properties were used to predict the relative potencies of the candidates through correlation with the literature potencies of the known ligands. Through these *in silico* studies, novel drug candidates were successfully created to have both high affinity and high specificity for AChBP's Aromatic Box.

2. Methods

2.1 *Brief Description of Computational Methods and Basis Sets*

Wave Functional Theory (WFT) methods, which explicitly use the wave functions (Ψ) of the electrons in a molecule to determine a variety of molecular properties, were developed in the early twentieth century. Hartree-Fock (HF) was one of the first WFT methods to be developed, but it does not include any information about electron correlation—which describes the tendency for electrons to avoid each other—leading to a poor representation of induction and dispersion forces.[10] MP2 contains algorithms that take into account approximately 85% of the correlation, but it is a more computationally expensive method. More accurate WFT methods, including the "gold standard" couple cluster singles-doubles-triples (CCSDT) take into account up to 95% of the correlation, but they are too expensive for most supercomputing resources and can take weeks to months to complete a single calculation; these methods are not within the reach of the present study.

Density Functional Theory (DFT) methods were developed in the 1950's, but came to maturity in the 1980-90's. They use different algorithms to calculate information about the system in question. There are two components to DFT methods, which differ between methods: the exchange functional and the correlation functional.[10] Exchange refers to the interactions between delocalized electrons in a system. This is especially important in dispersion-dominated complexes due to the high possibility of individual electrons being spread between multiple orbitals. Correlation takes into account the tendency of electrons to avoid each other by allowing the electrons to move into excited states to minimize repulsion and instability. Hundreds of DFT methods exist today,

compared with the handful of WFT methods; only a few were used in the present study. None of the DFT methods perfectly replicate the coupled cluster methods, but the DFT methods in this study were carefully selected in order to replicate the experimental values for aromatic-dominated systems.

SVWN[11,12], B3LYP[13], HCTH407[14,15], and BHANDHLYP[16] are the four main DFT methods used in this study. SVWN and B3LYP are two of the oldest DFT methods; B3LYP is still one of the most widely used and cited methods. Zhao and Truhlar showed that B3LYP was not accurate for describing complexes with hydrogen bonds, ones that are dispersion-dominated, or complexes that have both of these qualities.[10] On the other hand, this study showed that SVWN (sometimes referred to as LSDA) is one of the most accurate methods for describing dispersion-dominated complexes and mixed complexes, being only slightly less accurate than MP2 for mixed complexes.[10] The present study with AChBP is dispersion-dominated and has hydrogen bonds, so SVWN was selected for use and B3LYP was used as a point of comparison. HCTH407 and BHANDHLYP were both used to quickly optimize the geometry of structures.

2.2 *Isolation of the Active Site*

The Aromatic Box was isolated from a crystal structure of AChBP on the Protein Data Bank.[17] Crystal structures of AChBP in complex with nicotine, cocaine, and galantamine were also available via the Protein Data Bank. For these crystal structures, all five residues in the active site were manually isolated with the program ArgusLab[18] based upon their proximity to the bound ligand and each other. Peptide bonds connecting the surrounding residues were severed and the dangling nitrogen atoms were replaced with hydroxyl groups in order to mimic the electronegativity of the peptide bond. Missing

protons were added. The geometry of this structure was optimized using the HCTH407/6-31G method and basis set, keeping the heavy atoms of the residues static while allowing the ligand and hydrogens to move. All calculations were performed using Gaussian03.[19]

2.3 *Preliminary Studies, Evaluating Methods Used*

Preliminary studies calculated the Mulliken Atomic Charges of isolated benzene, sandwiched benzenes, and t-shaped benzenes using a variety of methods (HF, MP2, QCISD[20], SVWN, B3LYP, and HCTH407) and basis sets (3-21G, 3-21+G*, 6-31G, 6-31+G*, 6-311G, 6-311+G*, aug-cc-pvdz, and aug-cc-pvtz). Mulliken Atomic Charges quantify the partial positive or negative charge on every atom of the molecule or complex in question. From this study, MP2, B3LYP, and HCTH407 were selected along with the 6-31G, 6-31+G*, 6-311G, and 6-311+G* basis sets for use in all subsequent Mulliken Atomic Charge calculations. QCISD as well as the aug-cc-pvdz and aug-cc-pvtz basis sets were all either too expensive or unreliable to use with the available supercomputing resources. All reported charges in this document are from using MP2/6-311+G*.

2.4 *Evaluation of Mulliken Atomic Charges*

The Aromatic Box was built by gradually adding functional groups and other chemical moieties to benzene, the most common component of the residues in the active site. These moieties were added in a stepwise manner, forming a series of component structures that led up to the full amino acid structures and the full Aromatic Box. Mulliken Atomic Charges, which quantify the partial charge on each of atoms in the system, were calculated for each of these structures. The component structures used to create tyrosine included benzene, toluene, and *p*-cresol, and the component structures used to create tryptophan included benzene and indole. The Mulliken Atomic Charges of

each of these component structures was evaluated both isolated and in complex with benzene in a variety of conformations. The geometry of each complex was optimized, allowing all atoms to move. Redundant structures were eliminated and represented by a single structure. Mulliken Atomic Charges were calculated at each step, tracking the partial charge change on each atom. Mulliken Atomic Charges were also calculated for isolated nicotine, the Aromatic Box, and the Aromatic Box with bound nicotine.

2.5 *Docking of Ligands*

Ligands for which there was no pre-existing crystal structure, including morphine and all of the candidates, were manually docked using ArgusLab's GADock and ArgusDock docking engines. The active site was isolated from the crystal structure of AChBP in complex with nicotine and the ligand was manually constructed in GaussView. Both structures' geometries were individually pre-optimized using BHANDHLYP/3-21G. Subsequent calculations were performed in the same way as those ligands that did have an existing crystal structure.

2.6 *Evaluation of Interaction Energies*

Using the optimized structures of the ligands bound in the Aromatic Box, each ligand/amino acid pair was isolated in turn. After the ligand/residue complexes were isolated, the counterpoise-corrected interaction energies of these complexes were calculated using MP2, SVWN, and B3LYP and the 6-311+G* basis set. Interaction energy calculations for the candidate molecules were performed using only SVWN and B3LYP. B3LYP values were used throughout the interaction energy calculations to provide a reference, though SVWN has been shown to outperform B3LYP for π-stacking induction and dispersion forces.[22]

2.7 *Electrostatic Potential Maps*

Electrostatic potential maps were generated for the empty Aromatic Box as well as the Aromatic Box in complex with each of the known ligands using HF and the 3-21G basis set.

2.8 *Quantitative Structural Activity Relationships (QSAR) Calculations*

The structure of each of the isolated ligands, including both the known ligands and the candidates, were optimized using HCTH407 and the 6-31G basis set in the gas phase and using an implicit solvent of water and chloroform. Quantitative information about the solvation energies, ratio of solvation energies, HOMO, LUMO, and dipole moment was gathered. Aggregate pEC50 values found in the literature were correlated with each of these molecular properties and linear regressions were assigned to each of the properties. Equations to predict the candidates' pEC50 values were constructed using regressions in Microsoft Excel.

3. Results and Discussion

3.1 *Preliminary Charge Migration Studies Using Benzene and Intermediate Structures*

The charge distribution of the Aromatic Box was evaluated via Mulliken Atomic Charges in stages, starting with the most common moiety present in all of the residues: benzene. The Mulliken Atomic Charges of benzene were evaluated and moieties were gradually added until both tyrosine and tryptophan residues had been formed (Figure 2). Mulliken Atomic Charge calculations quantitatively show the partial atomic charge on each atom in the complex. The charge distribution in benzene showed a negative interior and positive exterior of the ring, correlating with the carbons and hydrogens,

respectively. In order to make tyrosine, two intermediate steps were used: toluene and *p*-cresol. For Tryptophan, the only intermediate step used was indole.

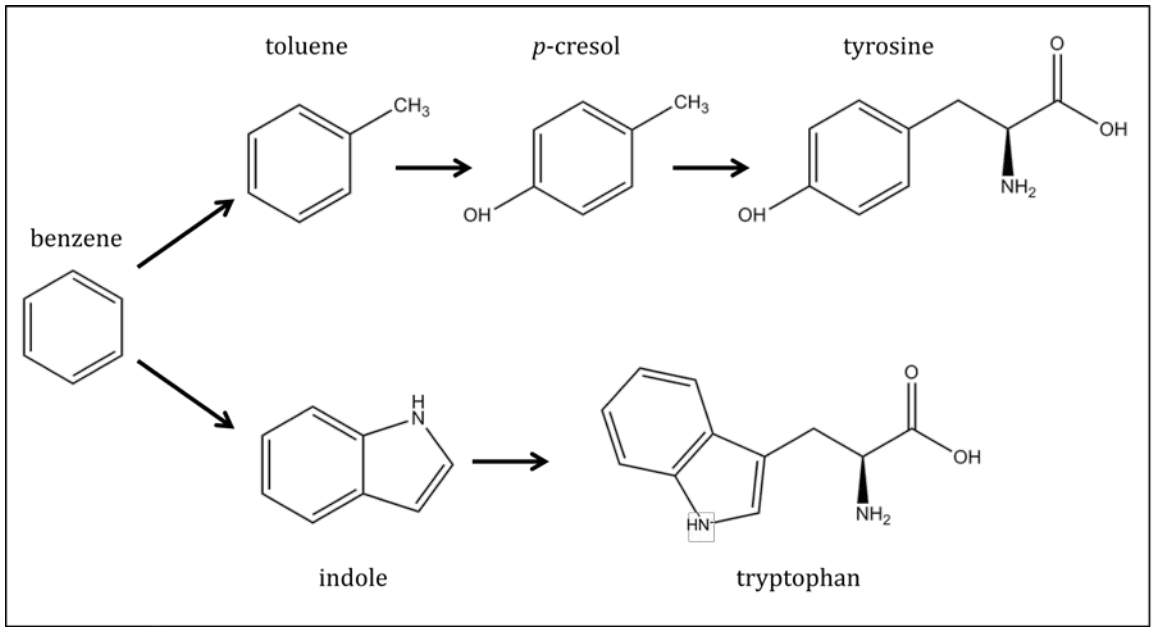


Figure 2: Shows the intermediate structures used to build the two types of aromatic residues found in the Aromatic Box.

Several of the component structure/benzene complexes optimized to similar geometries. All redundant structures were eliminated from calculations, letting one serve as the prototype for the rest. The aromatic ring interiors were all negative, while the exteriors showed partial positive charges. The majority of the large negative charges were concentrated on the nitrogens and oxygens. At each step the charge distribution of the isolated molecule was evaluated. The Mulliken Atomic Charges were then calculated using the charges for the empty Aromatic Box, the Aromatic Box in complex with nicotine, and isolated nicotine (Figure 3). The geometry of each component structure in complex with benzene was optimized using HCTH407/6-31G, allowing only the hydrogens to move. When nicotine binds to the Aromatic Box active site of AChBP, the

greatest change in charge distribution occurs within the nicotine molecule, while the active site itself exhibits comparatively less charge migration. Each atom in the amino acid residues' partial charge only changed between zero and 0.07 atomic charge units, while the atoms in nicotine changed between 0.04 and 0.15 atomic charge units.

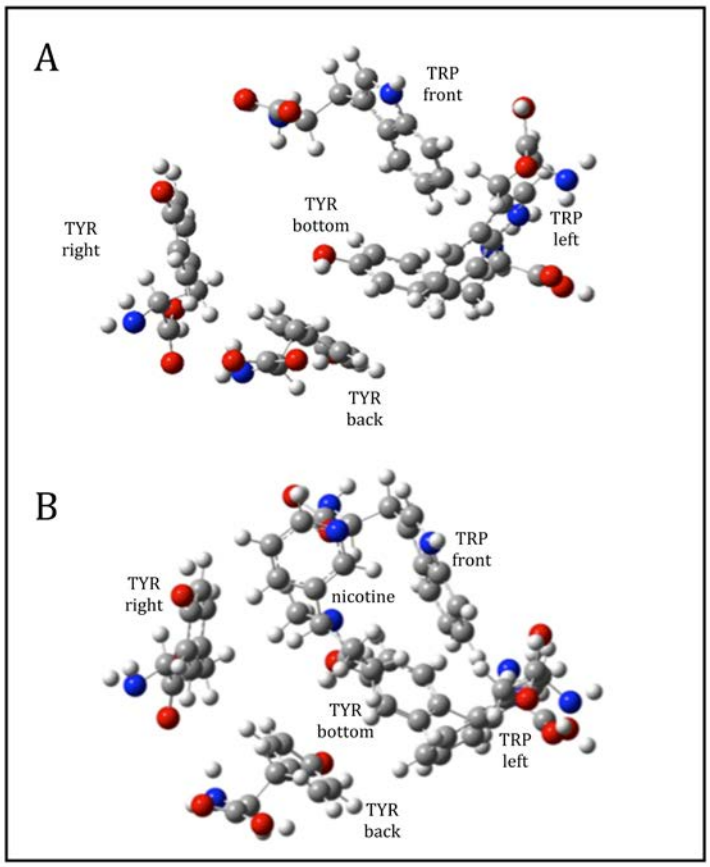


Figure 3: Empty Aromatic Box (A) and Nicotine in complex with the Aromatic Box (B). Both images show the top-down view of the active site. Residues are labeled with the nomenclature used for the remainder of this study.

3.2 *Interaction Energy Calculations*

The interaction energies between each residue of the Aromatic Box and each ligand was calculated using MP2, SVWN, and B3LYP with the 6-311++g** basis set. Calculated interaction energies are shown in Tables 1.1, 1.2, and 1.3 and Table 2 shows a summary of the total interaction energies calculated for each of the ligands.

Table 1.1: MP2 interaction energies (kcal/mol) for known ligands. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

MP2	Nicotine	Cocaine	Galantamine	Morphine (1)	Morphine (4)
TRP Left	-1.71	--	-2.91	-1.44	0.28
TRP Front		-7.59	--	21.88	-3.46
TYR Back	-2.19	-3.96	-0.47	10.80	47.57
TYR Right	-3.56	-6.30	1.25	1.05	-0.06
TYR Bottom	2.28	-2.76	-7.52	14.59	0.08
TOTAL:		-20.61	-9.65	46.88	44.41

Table 1.2: SVWN interaction energies (kcal/mol) for known ligands. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

SVWN	Nicotine	Cocaine	Galantamine	Morphine (1)	Morphine (4)
TRP Left	-1.56	--	-3.19	17.17	-1.82
TRP Front	-2.97	-9.24	--	-4.51	-6.60
TYR Back	-1.46	-4.45	-0.14	0.88	33.27
TYR Right	-4.05	-9.47	-4.50	-2.26	-2.87
TYR Bottom	-0.56	-3.02	-13.42	2.69	-2.12
TOTAL:	-10.60	-26.18	-21.25	13.97	19.86

Table 1.3: B3LYP interaction energies (kcal/mol) for known ligands. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

B3LYP	Nicotine	Cocaine	Galantamine	Morphine (1)	Morphine (4)
TRP Left	0.70	--	1.64	5.83	6.30
TRP Front	10.22	3.81	--	47.37	5.60
TYR Back	0.27	1.94	0.20	22.10	65.55
TYR Right	0.54	0.26	5.99	6.34	6.53
TYR Bottom	4.39	1.31	1.02	18.18	2.20
TOTAL:	16.12	7.32	8.85	99.81	86.18

Table 2: Summary table of the total interaction energies (kcal/mol) for the known ligands, reported for each of the three methods used. Negative values correspond to overall attractive forces and positive values correspond to overall repulsive forces.

	Nicotine	Cocaine	Galantamine	Morphine (1)	Morphine (4)
MP2		-20.61	-9.65	46.88	44.41
SVWN	-10.60	-26.18	-21.25	13.97	19.86
B3LYP	16.12	7.32	8.85	99.81	86.18

Of all of the known ligands studied, cocaine exhibited the strongest overall attractive (negative) interaction energies, with a MP2 overall interaction energy of -20.61 kcal/mol. Galantamine also had a high affinity for the Aromatic Box, with a MP2 overall interaction energy of -9.65 kcal/mol. It is interesting to note that the interactions with TYR bottom at the base of the active site bound less strongly to nicotine and cocaine (MP2 2.28 and -2.76 kcal/mol, respectively) when compared with the other residues, but TYR bottom had the most attractive interaction energy with galantamine, with a MP2 value of -7.52 kcal/mol.

Morphine always had the highest repulsive values, regardless of its conformation. Morphine was re-docked and its interaction energies were recalculated three times without improvement over the first (Tables 3.1, 3.2, 3.3) (MP2 total interaction energy of 46.10 vs. 54.89 kcal/mol). Overall, it was concluded that the larger drug molecules with multiple rings (cocaine, galantamine) bind more effectively to the Aromatic Box than smaller molecules (nicotine). Morphine was considered in all subsequent studies, keeping in mind that its interaction energy results were less than favorable.

Although MP2 is the standard method for this study, SVWN shows a similar trend in the interaction energy results. Cocaine still had the highest binging affinity (-26.18 kcal/mol), followed by galantamine (-21.25 kcal/mol) and nicotine (-10.60 kcal/mol). The specific ligand/residue pairs that had the most attractive MP2 interaction energies also

showed a similarly high relative attraction in the SVWN interaction energies (e.g. cocaine/TRP front: SVWN -9.24; MP2 -7.59). However, B3LYP does not mimic the trend of MP2 as well as SVWN. Cocaine still had the most attractive interaction energy (7.32 kcal/mol), but it was shown as having a repulsive force. In continuing to use cocaine/TRP front as an example, the high affinity shown in SVWN and MP2 was not shown in B3LYP; it had the most repulsive force for this complex at 3.81 kcal/mol. These results are as expected, since this active site is dominated by induction and dispersion weak interactions, for which B3LYP is not the best method.[10] However, it is important to note that both SVWN and B3LYP calculations took significantly less time to complete.

Table 3.1: MP2 interaction energies (kcal/mol) for each morphine conformation. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

Residue	Morphine (1)	Morphine (2)	Morphine (3)	Morphine (4)
TRP Left	-1.43	-4.09	-2.74	0.28
TRP Front	29.88	15.67	-4.56	-3.76
TYR Back	10.80	70.59	51.35	49.57
TYR Right	1.05	10.73	9.60	-0.06
TYR Bottom	14.59	-0.25	-0.38	0.08
TOTAL:	54.89	92.66	53.27	46.10

Table 3.2: SVWN interaction energies (kcal/mol) for each morphine conformation. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

Residue	Morphine (1)	Morphine (2)	Morphine (3)	Morphine (4)
TRP Left	-4.51	-3.76	-1.72	-1.82
TRP Front	17.17	4.26	-6.80	-6.60
TYR Back	0.88	47.04	32.05	33.27
TYR Right	-2.26	4.88	4.29	-2.87
TYR Bottom	2.69	-0.31	0.01	-2.12
TOTAL:	13.97	52.12	27.84	19.86

Table 3.3: B3LYP interaction energies (kcal/mol) for each morphine conformation. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

Residue	Morphine (1)	Morphine (2)	Morphine (3)	Morphine (4)
TRP Left	5.83	1.20	-0.15	6.30
TRP Front	47.37	23.93	1.12	5.60
TYR Back	22.10	81.02	60.22	65.55
TYR Right	6.34	16.25	16.10	6.53
TYR Bottom	18.18	1.01	0.47	2.20
TOTAL:	99.81	123.41	77.76	86.18

3.3 *Electrostatic Potential Maps*

In order to further understand the electronic nature of the Box, electrostatic potential maps of the empty Aromatic Box and the Box in complex with nicotine, cocaine, morphine, and galantamine were calculated and examined (Figure 4). These images show the partial charge distribution, which can be interpreted to aid in drug design. By seeing where the “pockets” of positive and negative charge are located, decisions can be made about which functional groups to include in drug candidates.

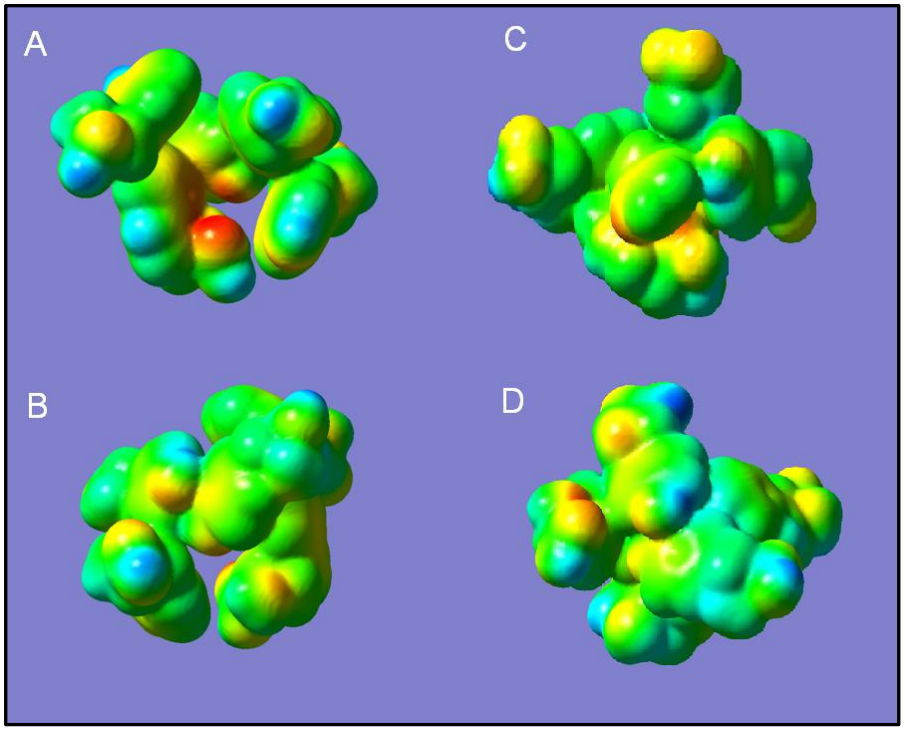


Figure 4: Empty Aromatic Box, top view (A), empty Aromatic Box, bottom view (B), Aromatic Box in complex with nicotine, top view (C), Aromatic Box in complex with nicotine, bottom view (D). All electrostatic potential maps were created using HF/3-21g.

The two areas that were designated as targets for ligand binding after the electrostatic potential maps (ESPs) were generated were the negative pocket inside the base of the Aromatic Box and the ring of nearly alternating charges surrounding the entrance to the binding pocket (Figure 4-A). Nicotine shows adequate binding by stabilizing many of the pockets of negative and positive charge.

Once again, morphine showed less than desirable results (Figure 5). There are many areas where strong partial positively charged pockets are in close proximity with both similarly charged and oppositely charged pockets. This creates imbalance by way of a lack of stability in the complex. The visual instability seen in morphine's ESP maps is the manifestation of the destabilizing positive interaction energies seen in Tables 3.1, 3.2, and 3.3.

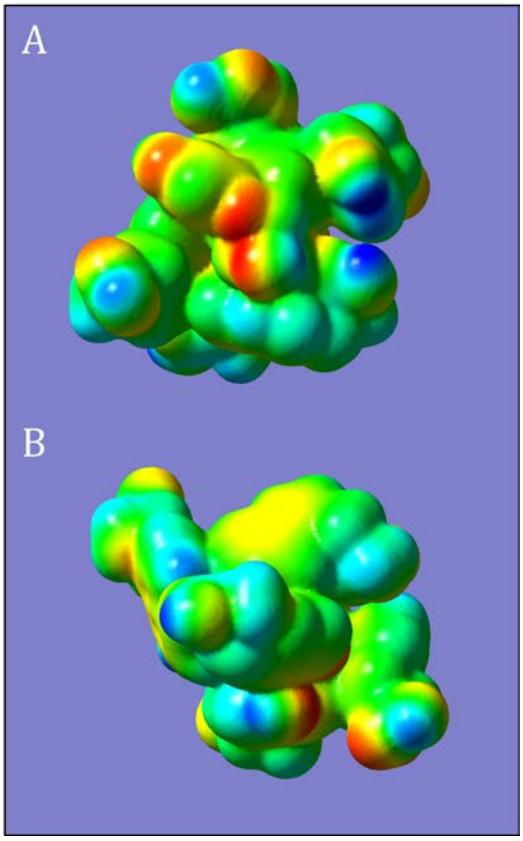


Figure 5: Aromatic Box in complex with morphine (conformation 4), top view (A) and the Aromatic Box in complex with morphine (conformation 4), bottom view (B). All electrostatic potential maps were created using HF/3-21g.

3.4 *Drug Design*

Based upon the ESP 3D renderings of the active site, several drug candidates were created to take advantage of the electron-rich deep interior of the Aromatic Box as well as the ring of electron-deficient patches at the entrance to the Aromatic Box. Polycyclic structures with broad back ends served as the basic structures so that the entire Aromatic Box would be filled. The conjoined aromatic rings span the entrance to the Aromatic Box and were designed to stabilize the ring of charges through their π-clouds. The candidates that were designed are shown in the order of their creation in Figure 6.

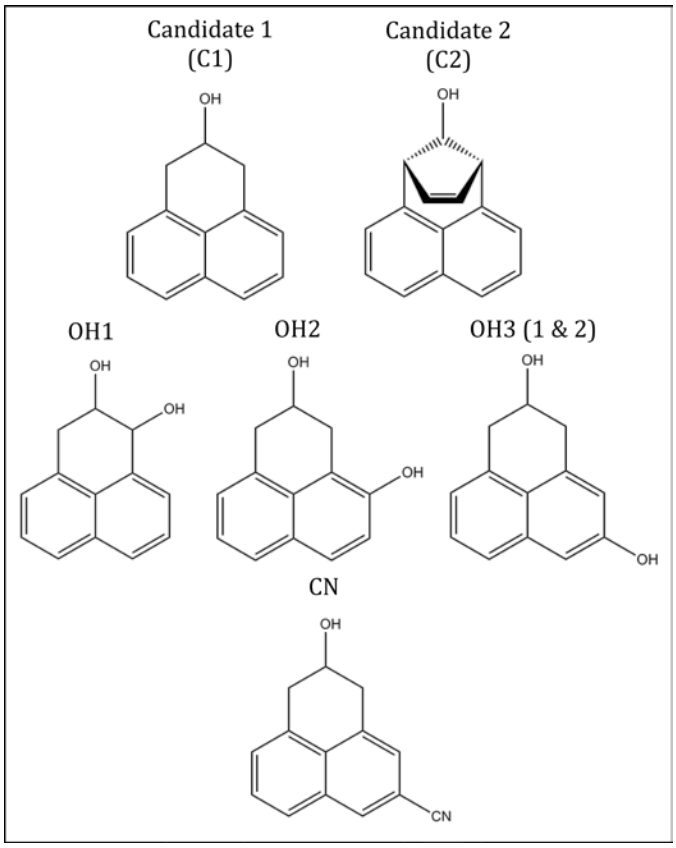
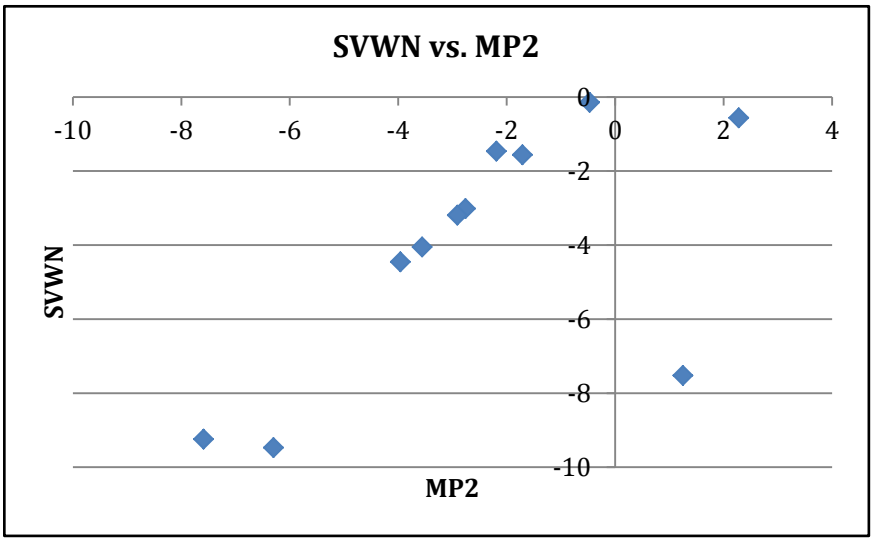


Figure 6: Structures of novel drug candidates studied. Each row represents a new generation of candidates.

The interaction energies for the candidates were calculated in the same way as those for the known ligands with one exception. MP2 was not used due to the extensive amount of time and supercomputing resources necessary to perform these calculations. However, all previous data that had both MP2 and SVWN data points for the same structure were compared in Figure 7, and it was concluded that SVWN adequately represented the trend of MP2 via a linear relationship with few outliers. Stereotypically, the SVWN DFT method overestimates interaction energies. However, SVWN is known to balance this discrepancy when it is used with a smaller basis set. Because the

interaction energies in this study were calculated using the 6-311+G* basis set, SVWN is a good method to replicate the trend of MP2 while using less computing resources.

Figure 7: Comparison between MP2 and SVWN for all calculations within this study for which these two data points exist. Includes interaction energies of the known ligands.



The interaction energies for the drug candidates are shown in Tables 4.1 and 4.2. Hydroxyl groups were used as a starting functional group for to their amphipathic qualities in order to take advantage of the pocket of negative charge at the base of the active site, shown in Figure 3-A. The known ligand with the most attractive interaction energy was cocaine, with a total MP2 interaction energy of -20.61 kcal/mol and total SVWN interaction energy of -26.18 kcal/mol. Candidates 1 and 2 were evaluated first in order to determine the direction of drug design. Candidate 1 did not show an overall improvement over cocaine, binding with a total SVWN interaction energy of -15.62 kcal/mol. However, Candidate 2 was significantly worse than both Candidate 1 and cocaine with a total repulsive SVWN interaction energy of 111.56 kcal/mol. This erratic

number is largely due to the ligand's interaction with TYR back, which was extremely repulsive 120.87 kcal/mol. Steric hindrance from Candidate 2's extra ring is likely responsible for this nonsensical number, as it leaves very little space between the residues of the Aromatic Box and the ligand itself (Candidate 2 is as close as 0.91 Å to TYR back). Candidate 1 was selected as the base for functional group modifications.

Table 4.1: SVWN interaction energies for each drug candidate. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

Residue	C1	C2	OH1	OH2	OH3_1	OH3_2	CN
TRPleft	-5.89	-1.89	-5.61	-4.94	-11.62	-5.63	-6.95
TRPfront	-9.19	-1.96	-9.34	-11.70	-13.85	-11.98	-6.91
TYRback	-0.12	120.87	-2.66	-3.31	-2.82	-4.42	-11.18
TYRright	-0.30	-2.25	-10.20	-14.60	-2.82	-2.22	-4.18
TYRbottom	-0.12	-3.22	-0.89	-0.32	-2.05	-10.27	-2.10
TOTAL	-15.62	111.56	-28.70	-34.87	-33.16	-34.52	-31.32

Table 4.2: B3LYP interaction energies for each drug candidate. Negative values correspond to attractive forces and positive values correspond to repulsive forces.

Residue	C1	C2	OH1	OH2	OH3_1	OH3_2	CN
TRPleft	1.37	0.81	2.87	2.21	0.02	2.14	2.58
TRPfront	1.81	-0.34	-0.50	-1.42	1.89	-0.59	1.07
TYRback	-0.90	169.56	0.47	1.44	1.02	0.43	-1.94
TYRright	0.14	2.02	-0.85	-5.52	1.02	4.00	1.80
TYRbottom	-0.09	-0.50	-0.12	-0.32	0.77	-3.64	1.43
TOTAL	3.15	171.52	1.97	-3.62	4.71	2.34	4.93

Each of the subsequent candidates showed an improvement in binding (more negative total binding energy) over cocaine. The OH1 candidate showed the least amount of improvement, with a SVWN total binding energy of -28.70 kcal/mol. The OH2 and OH3 (2) candidates were the best, with total binding energies of -34.87 and -34.52 kcal/mol, respectively. Overall, the addition of multiple hydroxyl groups produced significantly improved binding results over any of the known ligands. This is especially pointed when considering galantamine, which had a total MP2 interaction energy of -9.65

and a total SVWN interaction energy of -21.25. Based upon this portion of the study, the candidates are strong contenders to replace galantamine's pharmacological use in targeting this protein.

3.5 *Correlating QSAR Molecular Properties*

In order to further evaluate whether these candidates are viable for pharmacological use, QSAR calculations were performed. QSAR calculations quantitatively evaluate several molecular properties that relate to the reactivity and solubility of the drug *in vivo*. These molecular properties were correlated with potency measured by pEC50. The EC50 value of a substance refers to the plasma concentration that yields 50% of the desired full effect of the drug or other substance when used *in vivo*.[23] pEC50 is often used as a way to measure the potency of a drug; the lower the pEC50 value, the more potent the drug. The pEC50 values for known ligands are correlated with the molecular property values from the QSAR calculations to create an equation in the form of Equation 1, where x, y, and z are the molecular properties used to calculate a regression in Microsoft Excel, through which the a, b, and c coefficients can be found as well. This study used three terms (three molecular properties), but any number may be used.

(1)

$$pEC50 = ax + by + cz + ...$$

The properties studied included the dipole moment, which gives information about binding interactions, the HOMO and LUMO, which represents the energy needed for electrons to move, and the ratio of solvation energies, which shows the ability of the molecule to pass through the blood and cell membranes. Aggregate pEC50 values were

compiled from a variety of literature sources [24-33]. These values were correlated with the molecular properties from the QSAR data in Table 5 (Figure 8) and linear regressions were found for each of the properties across all the known ligands (Table 6). These regression equations were ranked by their R^2 values, denoting the strength of their correlation with the pEC50 values of the known ligands. Multi-variable equations for pEC50 were created using the molecular properties with the highest R^2 values (Table 7). The exception to this is the ratio of solvation energies, which did not have the highest R^2 value (0.42); it was used due to the importance of this quality in the solubility of drug candidates in blood as well as their ability to cross the blood-brain barrier. Each equation trial used a different combination of QSAR molecular properties, and some trials omitted morphine's data due to the ligand's inconclusive data earlier in this study. The coefficients of each equation were found using Microsoft Excel's regression function, specifying which known ligands and molecular properties to use.

Table 5: Calculations of a variety of molecular properties—solvation energy (Hartrees) in water and chloroform, dipole moment (Debye), HOMO (Hartrees), LUMO (Hartrees), and the solvation ratio—for each of the known ligands (nicotine, cocaine, galantamine, and morphine) and drug candidates. Calculations were done in the gas phase, implicit water, and implicit chloroform solvents.

	E_{solv} (H_2O)	**E_{solv} (Chl)**	**Dipole Moment**	**HOMO**	**LUMO**	**Solvation Ratio**
Nic	-6.46	-4.26	4.28	-0.19	-0.07	0.66
Coc	-13.17	-8.67	2.00	-0.22	-0.08	0.66
Gala	-26.21	-18.96	5.26	-0.11	-0.05	0.72
Morp	-20.41	-13.18	5.55	-0.18	-0.04	0.65
C1	-11.66	-7.42	2.90	-0.19	-0.06	0.64
C2	-13.17	-8.67	2.00	-0.20	-0.08	0.66
OH1	-15.60	-10.01	3.35	-0.20	-0.08	0.64
OH2	-17.06	-11.08	2.39	-0.19	-0.07	0.65
OH3_1	-17.14	-11.22	2.86	-0.19	-0.07	0.65
OH3_2	-17.14	-11.24	2.59	-0.19	-0.07	0.66
CN	-15.22	-10.38	8.55	-0.23	-0.10	0.68

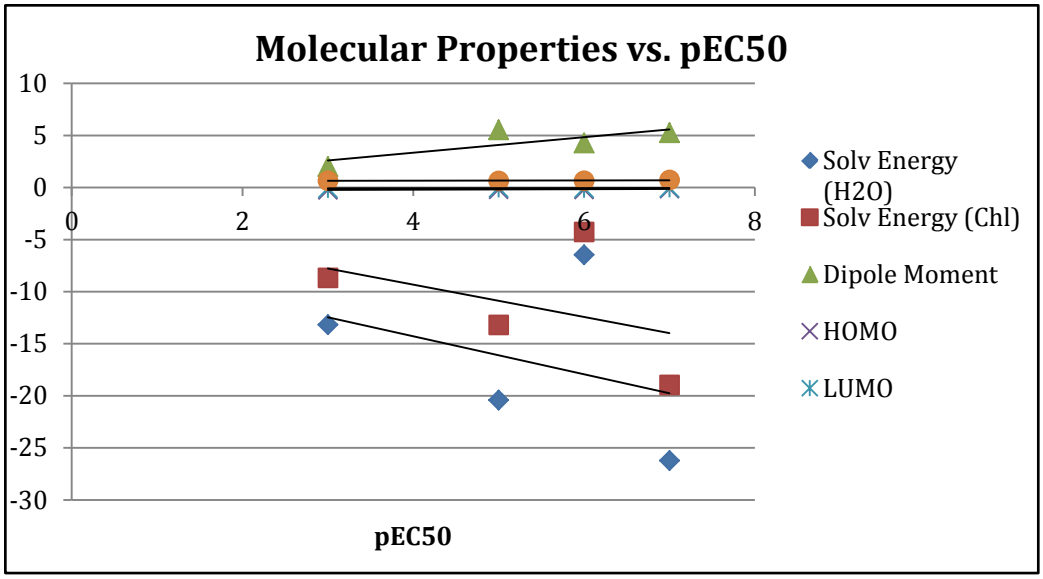


Figure 8: Graph correlating QSAR calculations of molecular properties to the literature values for the aggregate pEC50 for the known ligands.

Table 6: Linear regression equations from Figure 7.

Molecular Property	Linear Fit Equation	R^2 Value
Solvation Energy in H_2O	y=-1.8269x-6.9748	0.13186
Solvation Energy in Chloroform	y=-1.5498x-3.1308	0.17716
Dipole Moment	y=0.7446x+0.3623	0.62687
HOMO	y=0.0229x-0.2926	0.71158
LUMO	y=0.0068x-0.0943	0.43246
Solvation Ratio	y=0.0134x+0.6014	0.42738

Table 7: QSAR equations predicting pEC50. Equations were generated in Microsoft Excel, using the R2 values of the linear regressions from Figure 7 and Table 6 to guide which molecular properties were used in the pEC50 equations.

Trial	Ligands Used	pEC50 Equation
QSAR Trial 1	Nicotine Cocaine Galantamine	=3.37 - 4.6(Solv. Ratio) + 1.32(Dipole Moment)
QSAR Trial 2	Nicotine Cocaine Galantamine Morphine	= 23.22 - 16.06(Solv. Ratio) + 41.63(HOMO)
QSAR Trial 3	Nicotine Cocaine Galantamine Morphine	= -13.05 + 0.70(Dipole Moment) + 22.80(Solv. Ratio)

Table 7 shows the first few formulae created to predict the candidates' pEC50. Efficacy of the equations is based upon how close the equation comes to predicting the correct aggregate pEC50 value for the known ligands. Equation 3 was the most successful, coming between 0.12 and 1.03 away from the literature values (Table 8). Using Equation 3, the CN-modified candidate is the most similar to galantamine (7.89 and 7.12, respectively, literature value 7), but all of the other candidates came closer to cocaine's pEC50 value (3.36, literature value 3). These studies show that all of the candidates presented in this study may be considered for further studies regarding this active site and α_7 nAChR active sites.

Table 8: Comparison between aggregate EC50 and pEC50 values for each of the known ligands and the predicted pEC50 values for each of the known ligands and candidates, calculated from the equations in Table 7 and values in Table 5.

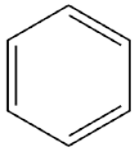
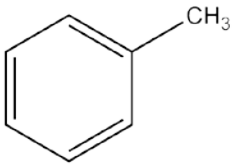
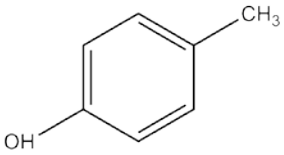
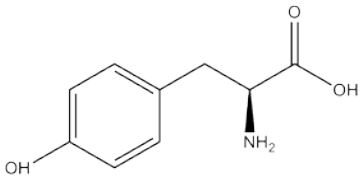
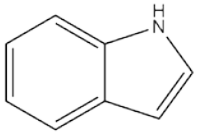
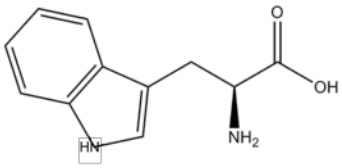
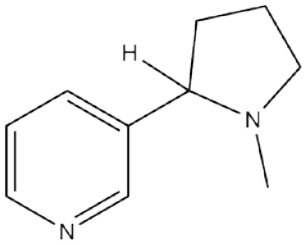
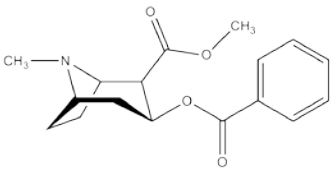
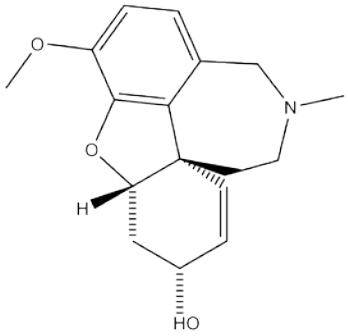
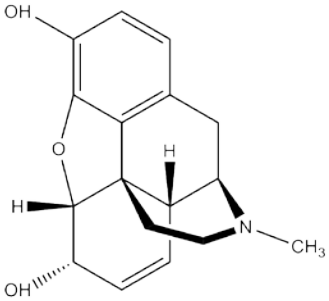
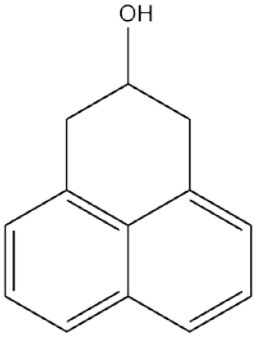
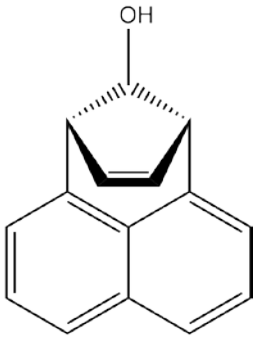
Ligand	Aggregate EC50	pEC50	Equation 1	Equation 2	Equation 3
Nicotine	1	6	5.99	4.73	4.97
Cocaine	1000	3	2.99	3.63	3.36
Galantamine	0.1	7	6.98	7.11	7.12
Morphine	10	5	7.72	5.53	5.56
C1	--	--	4.27	5.20	3.49
C2	--	--	2.99	4.50	3.36
OH1	--	--	4.84	4.60	3.92
OH2	--	--	3.53	5.05	3.43
OH3_1	--	--	4.13	4.86	3.88
OH3_2	--	--	3.77	4.84	3.71
CN	--	--	11.52	2.50	7.89

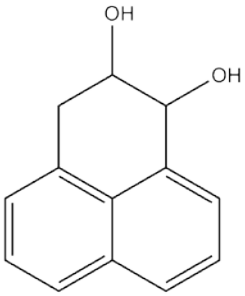
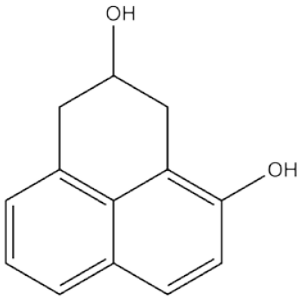
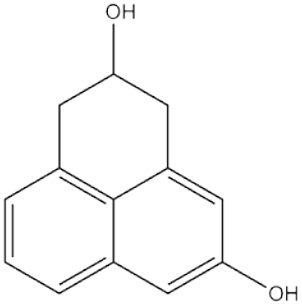
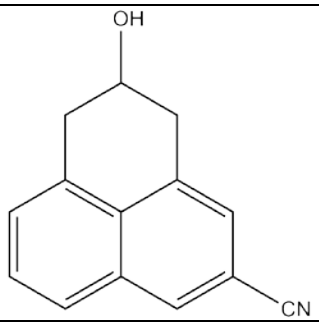
4. Physiological Implications and Conclusions

Based upon the above computational studies, novel drug candidates have been designed that have increased affinity for the Aromatic Box in terms of interaction

energies. Potency in terms of pEC50 was quantitatively predicted, which also yielded promising results. Taking all of these factors into account, all of the candidates except Candidate 2 may be considered for pharmacological use. In the future, it would be helpful to the field for these candidates to be synthesized and evaluated for solubility in water (to simulate blood) and chloroform or similar solvent (to simulate membranes) in order to determine experimentally the pharmacological significance of this work. Due to the addictive nature and detrimental consequences of the abuse of the known ligands, toxicity and potency of the candidates should also be determined experimentally in future studies.

APPENDIX I: STRUCTURES

Benzene	Toluene	p-Cresol
		
Tyrosine	**Indole**	**Tryptophan**
		
Nicotine	**Cocaine**	**Galantamine**
		
Morphine	**C1**	**C2**
		

C1OH1	C1OH2	C1OH3 (1 & 2)
		
	C1CN	
		

REFERENCES

1. National Institutes of Health, National Institute on Aging. http://www.nia.nih.gov/alzheimers/publication/alzheimers-disease-fact-sheet. (Accessed April 16, 2012).
2. U.S. National Library of Medicine. http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001762/. (Accessed April 16, 2012).
3. Ross, G. W.; Petrovitch, H. Current evidence for neuroprotective effects of nicotine and caffeine against Parkinson's disease. Drugs Aging. 2001; 18(11): 797-806.
4. Quik, Maryka; Bordia, Tanuja; O'Leary, Kathryn. Nicotinic receptors as CNS targets for Parkinson's disease. *Biochem Pharmacol*. **2007**. 74(8) 1224-1234.
5. Akdemir, Atilla; Rucktooa, Prakash; Jongejan, Aldo; van Elk, Rene; Bertrand, Sonia; Sixma, Titia K.; Bertrand, Daniel; Smit, August B.; Leurs, Rob; de Graaf, Chris; de Esch, Iwan J.P. Acetylcholine binding protein (AChBP) as a template for hierarchical in silico screening procedures to identify structurally novel ligands for the nicotinic receptors. *Bioorganic & Medicinal Chemistry*. **2011**. 19: 6107-6119.
6. Smit, August B.; Brejc, Katjusa; Syed, Naweed; Sixma, Titia K. Structure and Function of AChBP, Homologue of the Ligand-Binding Domain of the Nicotinic Acetylcholine Receptor. *Ann. N.Y. Acad. Sci.* **2003**. 998: 81-92.
7. Grimster, Neil P.; Stump, Bernhard; Fotsing, Joseph R.; Weide, Timo; Talley, Todd T.; Yamauchi, John G.; Nemecz, Ákos; Kim, Choel; Kwok-Yiu Ho; Sharpless, Barry K.; Taylor, Palmer; Fokin, Valery. *J. Am. Chem. Soc.* **2012**.
8. (CITE)
9. Patrick, G. *Medicinal Chemistry.* BIOS Scientific Publishers Limited: Oxford, 2001; pp 143-150.
10. Zhao, Yan; Truhlar, Donald G. Density Functional Theories of Biological Importance. *J. Chem. Theory Comput.* **2007**. 3(1): 289-300.
11. Vosko, S.H.; Wilk, L.; Nusair, M. Can J Phys 1980, 58, 1200.
12. Slater, J. C. Quantum Theory of Molecular and Solids, Vol. 4: The Self-Consistent Field for Molecular and Solids; McGraw-Hill: New York, 1974.
13. Becke, A.D. J Chem Phys 1993, 98, 5648.
14. Boese, A.D.; Handy, N.C. J Chem Phys 2000, 114, 5497.
15. Boese, A.D.; Doltsinis, N.L.; Handy, N.C.; Sprik, M. J Chem Phys 1999, 112, 1670.
16. A. D. Becke, "A new mixing of Hartree-Fock and local density-functional theories," *J. Chem. Phys.*, **98** (1993) 1372-77.
17. Ulens, C.; Akdemir, A.; Jongejan, A.; Van Elk, R.; Bertrand, S.; Perrakis, A.; Leurs, R.; Smit, A.B.; Sixma, T.K.; Bertrand, D.; De Esch, I.J. Use of acetylcholine binding protein in the search for novel alpha7 nicotinic receptor ligands. In silico docking, pharmacological screening, and X-ray analysis. *J. Med. Chem.* **2009**. 52:2372
18. ArgusLab 4.0.1 Mark A. Thompson Planaria Software LLC, Seattle, WA. <http://www.arguslab.com>.
19. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene,

X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

20. J. Gauss and C. Kremer, *Chem. Phys. Lett.* **150**, 280 (1988)
21. E. A. Salter, G.W. Trucks, and R.J. Bartlett, *J. Chem. Phys.* **90**, 1752 (1989).
22. Hofto, Laura R.; Van Sickle, Karina; Cafiero, Mauricio. Modeling Intercalation Through the Sandwich-Type Interactions Between Benzene and 14 Polyaromatic Molecules: DFT and Ab Initio Results. *International Journal of Quantum Chemistry.* **2007**. 108(1): 112-118.
23. U.S. Food and Drug Administration. http://www.fda.gov/ohrms/dockets/ac/00/slides/3621s1d/sld036.htm. (Accessed Apr 16, 2012).
24. Sullivan, J.P.; Donnelly-Roberts, D.; Briggs, C. A.; Anderson, D. J.; Gopalakrishnan, M.; Xue, I.C.; Piattoni-Kaplan, M.; Molinari, E.; Campbell, J. E.; McKenna, D.G.; Gunn, D. E.; Lin, N.H.; Ryther, K.B.; He, Y.; Holladay, M.W.; Wonnacott, S.; Williams, M.; Arneric, S.P. ABT-089 [2-methyl-3-(2-(S)-pyrrolidinylmethoxy)pyridine]: I. A potent and selective cholinergic channel modulator with neuroprotective properties. *J Pharmacol Exp Ther*. **1997**. 283(1):235-46.
25. Livett, B.G.; Boksa, P.; Dean, D.M.; Mizobe, F.; Lindenbaum, M.H. Use of isolated chromaffin cells to study basic release mechanisms. *J Auton Nerv Syst*. **1983**. 7(1):59-86.
26. Papke, Roger L.; Webster, Christopher; Lippiello, Patrick M.; Bencherif, Merouane; Francis, Michael M. The Activation and Inhibition of Human Nicotinic Acetylcholine Receptor by RJR-2403 Indicate a Selectivity for the ⟨4®2 Receptor Subtype. *Journal of Neurochemistry*. **2000**. 75:204-216. Online.
27. The Binding Database. http://www.bindingdb.org/bind/searchby_r1l.jsp?constrain=0&reactant1=nicotine&tag=r1l&loMW=&hiMW=&loKI=&hiKI=&loIC=&hiIC=&lodG=&hidG=&anDor=and&submit=Search. Accessed April 15, 2012.
28. Wiley, John et. al. *The Biology of Nicotine Dependence*, Google eBook ed.; Ciba Foundation Symposium, 1990; pp 63-95. Online.
29. Sharp, Burt M.; Yatsula, M.; Fu, Yitong. Effects of Galantamine, a Nicotinic Allosteric Potentiating Ligand, on Nicotine-Induced Catecholamine Release in Hippocampus and Nucleus Accumbens of Rats. *Journal of Pharmacology and Experimental Therapeutics.* **2004**. 309(3): 1116-1123. Online.
30. Silman, Israel; Soreq, Hermona; Anglister, Lili. *Cholinergic Mechanisms: Function and Dysfunction*. Taylor & Francis, 2004; pp 301-303.
31. Samochocki, Marek; Höffle, Anja; Fehrenbacher, Andreas; Jostock, Ruth; Ludwig, Jürgen; Christner, Claudia; Radina, Martin; Zerlin, Marion; Ullmer, Christoph; Pereira, Edna F. R.; Lübbert, Hermann; Albuquerque, Edson X.; Maelicke, Alfred. Galantamine Is an Allosterically Potentiating Ligand of Neuronal Nicotinic but Not of Muscarinic

Acetylcholine Receptors. *Journal of Pharmacology and Experimental Therapeutics*. **2003**. 305(3): 1024-1036. Online.

32. The Binding Database. http://www.bindingdb.org/bind/searchby_r1l.jsp?constrain=0&reactant1=cocaine&tag=r1l&loMW=&hiMW=&loKI=&hiKI=&loIC=&hiIC=&lodG=&hidG=&anDor=and&submit=Search. Accessed April 16, 2012.
33. Grant, R.L.; Acosta, D. Comparative toxicity of tetracine, proparacaine and cocaine evaluated with primary cultures of rabbit corneal epithelial cells. *Exp Eye Res*. **1994**. 58(4):469-78.
34. MM