Chapter II: Tertiary Structure and Ligand Binding Sites of the Human D_2 Dopamine Receptor

Modified from the manuscript of paper in preparation for submission to the Journal of Molecular Biology by M. Yashar S. Kalani, Nagarajan Vaidehi, Peter L. Freddolino, Maziyar A. Kalani, and William A. Goddard III[*]

Chapter II: Tertiary Structure and Ligand Binding Sites of the Human D_2 Dopamine Receptor

Abstract:

The dopamine neurotransmitter and its receptors play a critical role in such diseases as Parkinson's and schizophrenia. A problem with developing drugs for such diseases is that there are five subtypes of dopamine receptors only one of which should be affected for each disease. Since the binding sites are quite similar, it is difficult to design the subtype specific agonists and antagonists required for therapy with minimal side effects. This task has been particularly difficult since there are no crystal structures for any dopamine receptor or any closely related G-protein coupled receptors (GPCR) because of the difficulty in crystallizing these membrane-bound proteins.

We report here the 3-D structure of the human D_2 dopamine receptor (hD_2DR), predicted from the primary sequence using *ab initio* theoretical and computational techniques. This 3-D structure is validated by predicting the binding site and relative binding affinities of dopamine plus 3 known dopamine receptor agonists (antiparkinsonian) and 8 known antagonists (antipsychotic) in the hD_2DR receptor. The residues in the predicted binding site of these 11 ligands, agree quite well with all mutational studies implicating residues in the binding site of these ligands. The calculated binding energies also compare well with available experimental measurements of dissociation constants (a correlation factor of 0.92 with one outlier).

We find that all agonists (including dopamine) bind to a site located between transmembrane helices 3, 4, 5 and 6, but we find that antagonists for these receptors all fall into two distinct classes which we classify in terms of their binding site. Class I antagonists, such as clozapine, have structural similarity to agonists, and bind between TM helices 3, 4, 5 and 6. Class II antagonists, such as haloperidol, have two large aromatic ring systems linked by an alkyl chain containing an amine group, and occupy a site between TM 2, 3, 4, 5, 6 and 7. We find that agonists bind tightly between TM helices 3 and 5 but that antagonists bind tightly between TM3 and TM6. This may provide useful hints on the processes involved in activation.

We also constructed a homology model of D_2DR based on the high-resolution crystal structure of bovine rhodopsin (BovRh), which we compare to the *ab initio* structure of hD_2DR. Since the sequence identity between hD_2DR and BovRh is only 19%, it is not surprising that the homology model is not useful in predicting the binding site for dopamine, but that it can be used to rationalize experimental information on the ligand-receptor structure. The CRMS difference between the experimental BovRh structure and the ab initio structure is 5.55 Å.

Keywords: G-protein coupled receptors (GPCR), structure-based drug-design, dopamine, antipsychotic, antiparkinsonian, MembStruk, HierDock.

* *To whom correspondence should be sent:* wag@wag.caltech.edu

1.0 Introduction

Biogenic amines (such as epinephrine, dopamine, norepinephrine, tryptophan, and serotonin) play an essential role in the central and peripheral nervous systems. These biomolecules exert their effects by binding to the extracellular surface of the GPCR, which causes changes that lead to activation of a G-protein on the intracellular surface, which in turn leads to a cascade of events in the cytoplasm. These GPCRs consist of an extracellular amino terminus, an intracellular carboxy terminal region, and seven transmembrane (TM) domains, each of which has the form of an a-helix spanning the membrane. Connecting the TM domains are three intracellular (IC) loops and three extracellular (EC) loops.

A particularly well-studied biogenic amine is dopamine (a catecholamine intermediate in the biosynthesis of epinephrine and norepinephrine), whose receptors are important targets for treating schizophrenia (antagonists to D3) and Parkinson's diseases (agonists to D2)[1,2]. There are five known human Dopamine Receptors (DRs) with multiple isoforms for each. These DRs are partitioned on the basis of their pharmacological behavior into two subfamilies:

- D1-like: D1 and D5 show 82% sequence homology. These have a short third intracellular loop (IC3) and a long carboxy terminus.
- D2-like: D2, D3, D4 show 54 to 76% sequence homology (76% homology between D2 and D3 and 54% between D2 and D4). These have a long IC3 loop and a short carboxy terminus,

On the other hand the D1 and D_2 DRs have a sequence homology of only 44%.

Mutational studies have indicated that the IC3 loop is directly involved in G-protein coupling[3], but it is unlikely that these length differences between D1 and D2 affect the interaction the binding of dopamine.

Since all five DR's have the same endogenous ligand, dopamine, the binding sites are expected to be quite similar. This makes it a challenge to design agonists and antagonists specific to only one subtype of the DR's required to prevent cross-reaction of drugs between various DR subtypes. This difficulty is exacerbated greatly because there is no experimental 3-D structure for any DR of any species. Indeed considering all GPCRs of all forms of life, there is a 3-D structure for only a single GPCR, bovine rhodopsin. This is because GPCRs are membrane bound, making it difficult to obtain sufficiently good crystals for experimental xray diffraction studies. Some research groups have built homology molecular models for the D_2-DR based on the structure of bacteriorhodopsin (for example see: Ref. 4-5), or on bovine rhodopsin (for some examples see: Ref 6-7). Unfortunately the sequence similarity of D_2-DR is 19% to bovine rhodopsin, so that these homology models are not accurate enough to be used to design drugs specific to the receptor subtypes. They have been useful in rationalizing the results of biochemical experiments after using the information to refine the homology model.

Because sufficiently accurate experimental structures are not available, we developed computational first principles methods to predict the three-dimensional structures of GPCRs (MembStruk) and to predict the binding site and energy for various ligands to these structures (HierDock). These methods have been validated on bovine rhodopsin [8-9], human □2-adrenergic receptor[10], several mouse olfactory receptors [11-13]. Recently we provided and overview of the binding site of agonists and antagonists in the

human dopamine D_2DR denoted as hD_2DR[14] predicted using these methods. In this paper we report the details of the predicted three-dimensional structure of hD_2DR. The residues involved in recognizing the known agonists and antagonists and their contributions to the binding energy are also described here. The change in binding energy due to computational alanine scanning mutation of these residues are calculated and compared to the experimental change in binding affinities from experimental mutation studies. These results from *ab initio* structures are in excellent agreement with the experimental data on the binding site and ligand affinity of the hD_2DR, validating these predicted structures. In addition, we extracted some new insights about the characteristics of this receptor and the nature of the binding of agonists and antagonists. These results should be of use for design of subtype specific ligands for dopamine receptors and the validation of these computational techniques for this well characterized system opens the door to and the study of other GPCR targets where there is often little experimental knowledge.

We also built a structural model of hD_2DR using homology method [15] with the bovine rhodopsin crystal structure[16] (pdb code: 1F88) as template. The comparison with the *ab initio* 3-D structure provides insight into how much information can be extracted about ligand binding site from homology models.

2.0 Results and Discussion

2.1 Predicted Structure of hD_2DR

The structure prediction method is described in detail in reference 9 and the details relevant to the prediction of structure and ligand binding sites for hD_2DR is given in Methods and Materials section. **Figure 3-1** shows the predicted structure of the

hD_2DR. We refer to this predicted structure as hD_2DR (MS) where MS stands for MembStruk. The TM2ndS procedure[9] was utilized to identify the transmembrane (TM) spanning regions based on a hydrophobic analysis of the sequence. Twenty sequences of proteins with sequence similarity varying from 30 to 80% were used in the multiple sequence alignment using ClustalW program.[17] A seven helical motif was identified (supplementary materials) ranging from 19-29 residues per helix. The highlighted residues represent TM helices, while the intervening sequences are loop regions. The TM predictions for hD_2DR have helix core regions identical with the rhodopsin helices, but the rhodopsin helices are generally a bit longer. It is not clear whether this is because the D_2DR actually has shorter helices or because our prediction methods may leave off the helical termini.

These TM predictions were used to build seven canonical α-helices, whose coordinates were optimized using the MembStruk3.0 procedure described in reference 8. The helices were bundled in explicit bilayer of dilaurylphosphatidylcholine lipid molecules to mimic the biological membrane. The structural factors such as helical bend, helical tilt etc., of the predicted structure of hD_2 DR structure compared to the crystal structure of rhodopsin are summarized in **Table 2-1**.

Table 2-1 shows that the helical bends are very different for hD_2DR (MS) compared to the crystal structure of bovine rhodopsin. There is one disulfide bond between Cys107 and Cys182 in EC2 in the hD_2DR structure. The calculated RMS in the coordinates of the C_α atoms of the hD_2DR (MS) structure compared to the crystal structure of bovine rhodopsin is 5.55 Å and a main chain RMS of 6.24 Å. **Figure 2-1** has an overlay of hD_2DR (MS) to bovine rhodopsin, which shows that the hD_2DR (MS)

structure has several helices moved or tilted slightly outward relative to those in rhodopsin. Most notably, helices 1, 2, and 4 show net translations away from the protein center, while helices 6 and 7 tilt outward on the intracellular end. There are, however, many striking similarities between the two structures, especially in the characteristics of the individual helices. Helix 1 of hD_2DR (MS) follows the same curvature as in bovine rhodopsin, and both structures show a similar kink in the middle of TM 2. Helices 3 and 4 are relatively straight in both structures, although the helices in hD_2DR (MS) show a slightly different tilt. TM5 of hD_2DR (MS) is slightly bent in the center so that both ends are tilted inside of the barrel, a trait not seen in the rhodopsin structure. In both structures TM6 shows a pronounced kink due to a conserved proline residue, causing the ends of the helices to be tilted inwards. The structures differ notably only in helix 7, where the predicted dopamine structure has a much shorter TM region than the rhodopsin, breaking at a proline where the rhodopsin structure only kinks. Helix 7 of hD_2DR (MS) also appears to be translated slightly up (toward the extracellular end) relative to the rhodopsin helix. However we do not observe a peak corresponding to the 8^{th} helix in the hydropathicity profile for hD_2DR.

Comparison of predicted hD2DR structure to experiments: Javitch *et al.* performed a series of substituted cysteine accessibility method (SCAM) experiments that indicates which residues of each helix is water accessible (reviewed in ref. 18). **Figure 2-2** shows a helix-by-helix analysis of the residues implicated as solvent accessible and their location in the predicted hD2DR structure. We find that residues G51, N52, A58, and V59 identified as solvent accessible by SCAM studies are indeed not facing the membrane in the predicted structure as shown in **Figure 2-2a**. The residues implicated

by SCAM in helix2, helix3, helix4, helix5, helix6 and helix 7 are all found to be water accessible in our predicted structure as shown in **Figures 2-2b** to **Figure 2-2g**. SCAM studies identify 13 residues as being present in the water accessible crevice of TM5. Of the 13 residues, 10 are consecutive, from Phe189 to Phe198 and this is possible only if the conformation of TM5 is dynamic. We suggest that this is evidence for TM5 being mobile and able to readily rotate and expose different residues into the cavity Pro89 on helix2 causes a 13^{o} bend in TM2 whereas the bend in bovine rhodopsin is 22^{o}. Asn52 (TM1), Asp80 (TM2) and Asn351 (TM7) make up the sodium-binding pocket, an allosteric site that regulates agonist binding and receptor activation. This site has been shown to be sensitive to both sodium and proton concentrations and is far from the actual binding site of ligand[19]. **Figure 2-3** shows the predicted structure hD2DR (MS), overlaid with all residues identified experimentally as solvent-accessible (and thus interior to the helical bundle) highlighted in yellow. For all helices except TM4, our predicted structure has TM positions and rotations that correspond very well with experiment. Indeed for TM4 the solvent accessibility study led to discrepancies the experimental orientation of TM4 in rhodopsin, indicating possible experimental difficulties.

2.2 Prediction of binding sites and binding energies for 11 agonists and antagonists

Using the HierDock2.0 protocol with the hD_2DR (MS) structure, we predicted the binding sites and binding energies for a library of 11 agonists and antagonists (**Figure 2-4**). The entire hD_2DR (MS) structure was scanned to locate the putative binding region for these agonists and antagonists. The details of this scanning procedure and results are

given in the methods section. The figures for the binding site of all the 11 ligands are given in *http://www.wag.caltech.edu/GPCR/D2DR*

These studies suggest that the antagonists be partitioned into two classes:

1) class I antagonists are bulky and bind in a location between TM3, TM5, and TM6 similar to agonists; this is exemplified by Clozapine.

2) class II antagonists consists of two aromatic moieties connected by a flexible linker chain, one aromatic moiety is in the TM3, TM5, and TM6 region while the other is in the TM3, TM2, TM7 region; this is exemplified by Haloperidol.

2.2.1 Predicted binding site of dopamine: Figure 2-5 shows the predicted binding site of dopamine, the endogenous ligand of the dopamine receptors. We find that the binding site is located between TM3, 4, 5, and 6. Two main contacts dominate the energetic stabilization of dopamine in this docked conformation:

1) a salt bridge between Asp114 in TM3 and the protonated primary amine in dopamine (bidentate, 2.6 Å and 2.8 Å), and stop

2) a network of hydrogen bonds to the TM5 Ser193, Ser194, and Ser197 that bond to the catechol hydroxyls. In hD_2DR (MS) Ser193 and Ser197 on TM5, are both 2.7 Å from the catechol hydroxyls (meta and para respectively). Ser194 is ~ 5 Å away, too far for a direct hydrogen bond, however a rotation of TM5 could allow Ser194 to form strong interactions with the catechol. Even with helix motions we do not expect all three Ser in TM5 to form strong hydrogen bonds to dopamine. Interaction between Ser194 and Ser197 could occur during steps leading to or resulting from activation.

The residues Phe110, Met117, Cys118 (TM3), Phe164 (TM4), Phe189, Val190 (TM5), Trp386, Phe390, and His393 (TM6) provide hydrophobic packing for the ligand (all within 5.5 Å of the ligand). The conserved WXXFF motif is Trp386, Phe389, and Phe390, where Phe389 is 5.7 Å from dopamine.

His393 in TM6 also bonds to the amino end of dopamine. Asp114 can form a bidentate salt bridge, or a salt bridge plus a hydrogen bond to the amino group of dopamine. His393 can also bind to the amino group.

2.2.2 7-OH DPAT: Figure 2-6 shows the predicted binding site of 7-OH DPAT to the receptor. We find that this ligand binds in the same site as dopamine, between TM3, 4, 5, and 6. The protonated amino group of 7-OH DPAT forms a salt bridge to TM3 Asp114 (2.8 Å). The hydroxyl group on the phenyl ring forms two strong hydrogen bonds to TM5, 2.9 Å to Ser193 and 3.1 Å to Ser197. The Ser194 is at 5.4 Å, too long for a direct hydrogen bond.

The remaining residues in the cavity are mainly hydrophobic, providing stabilization for the aromatic and aliphatic rings of the ligand. The residues in close hydrophobic contact (<5.5Å with the ligand are Phe110, Leu113, Met117, Cys118 (all in TM3), Phe164 (TM4), Asn186, Val190 (both in TM5), Trp386, Phe390, His393, Ile394, Ile397 (all TM6), Thr412, Tyr416 (both in TM7).

2.2.3 Apomorphine: Figure 2-7 shows the predicted binding site of apomorphine to the receptor. The ligand binds in the same binding site as dopamine, located between TM3, 4, 5, and 6. The protonated amino group of apomorphine makes a salt bridge to TM3 Asp114 (2.9 Å). The other major polar contacts are two hydrogen bonds to the two-hydroxyl groups in apomorphine to both Ser193 (2.8 Å) and Ser197 (3.0 Å). The contact to Ser194 is ~ 6 Å.

Other residues form a mostly hydrophobic pocket around the ligand, including: Phe110, Met117, Cys118 (all TM3), Trp160, Phe164 (both TM4), Val190 (TM5), Trp386, Phe389, Phe390, His393 (all TM6), Thr412 (TM7). Mutational experiments showed that Trp386A and Phe390A mutants show substantial reduction in binding constants for apomorphine[34], indicating that these residues are important in ligand recognition. We find that they are within 5.0 Å of the ligand.

2.2.4 Bromocriptine: Figure 2-8 shows the predicted binding site of bromocriptine to the receptor. The ligand binds in the same site as dopamine located between TM3, 4, 5, 6, but also interacts with some residues in the EC2 loop. Again there is a salt bridge (2.8 Å) to TM3 Asp114 and hydrogen bonds of 2.8 Å and 3.8 Å to TM5 Ser193 and 197 respectively. There is also a 2.9 Å heteroatom contacts from the bromine to His393 and a second weak hydrogen bond (3.8 Å) from the amide hydrogen of bromocriptine to Thr412).

The residues providing a mostly hydrophobic pocket for the ligand are Phe110, Val111, Cys118, Thr119 and Ile122 (TM3), Trp160 and Phe164 (TM4), Asn186, Phe189, Val190, Tyr192, Val196, Val200 (TM5), Trp386, Phe389, Phe390 His393, Ile394, Ile397 (TM6), Tyr408, Ser409, Thr412 (TM7).

2.2.5 Clozapine: Figure 2-9 shows the predicted binding site of clozapine to the receptor. We classify clozapine as a class I antagonist, since it binds in the agonist binding site located between TM3, 4, 5, and 6. The amino group forms a salt bridge to TM3 Asp114 (2.8 Å), but now as for all antagonists studied, there is at least a single hydrogen bond to the serines of TM5. We find the hydrogen bond to Ser193 (3.2 Å), but with slight changes in the structure the heteroatom could instead form a bond to Ser197. A third favorable interaction, (3.1 Å) is between the N of Trp386 in TM6 and the hydrogen bonded to the ring nitrogen of clozapine. Cys118 is 4.5 Å from the other nitrogen in the clozapine ring. The remainder of the residues provide a mostly hydrophobic pocket for the ligand: Phe110, Leu112, Met117, Cys118 (TM3), Phe164 (TM4), Phe189, Val190 (TM5), and Phe382, Phe389, Phe390 (TM6), Thr412 and Tyr416 (TM7).

2.2.6 Haloperidol: Figure 2-10 shows the predicted binding site of haloperidol to the receptor, which we classify as a class II antagonists. This binding site is located between TM2, 3, 4, 5, 6, and 7, quite different from that of class I antagonists (TM3, 4, 5, 6). Haloperidol contains two aromatic domains connected by a flexible alkyl chain linker that has a protonated amino group. This protonated group forms a salt bridge to TM3

Asp114 (2.8 Å). There are also several other heteroatom contacts too weak to be considered hydrogen bonds. This includes a weak interaction between the chlorine and TM5 Ser197 (3.1 Å). There is also a hydrogen bond between the hydroxyl of haloperidol and the N of TM6 Trp386 (3.8 Å). There is a third weak interaction between the fluorine atom on the second aromatic portion of haloperidol and the N of TM2 Trp90 (3.0 Å). Other residues forming a mostly hydrophobic pocket around the ligand include: Leu94 (TM2), Phe110, Leu113, Met117, Cys118 (all in TM3), Trp160 and Phe164 (both in TM4), Ser193 and Ser194 (both in TM5), Phe389, Phe390, His393 (all in TM6), and Ser409, Thr412, Trp413, and Tyr416 (all in TM7).

2.2.7 Domperidone: Figure 2-11 shows the predicted binding site of domperidone to the receptor. This is also a class II antagonist with its predicted binding site between TM2, 3, 4, 5, 6, and 7. The two aromatic micro domains of the ligand bind between TM2 and 7 and between TM4 and 6 while TM3 provides a salt bridge to the amino group and TM5 provides weak interactions to the heteroatoms of the ligand. The protonated amino group is salt bridged to TM3 Asp114 (2.9 Å). The important heteroatom contacts are to Ser197 (3.3 Å), Cys118 (3.2 Å), Trp90 (3.7 Å) and Thr412 (3.9 Å). Other residues close to the ligand form a hydrophobic pocket for the ligand, including: Val87, Val91, Leu94 (TM2), Phe110, Leu113, Met117, Cys118, Ser121, Ile122 (TM3), Trp160, Phe164 (TM4), Val196, Val200 (TM5), Phe383, Trp386, Phe389, Phe390 (TM6), Ser409, Ala410, Thr412, Trp413, Tyr416 (TM7).

2.2.8 Raclopride: Figure 2-12 shows the predicted binding site of raclopride to the hD2DR receptor. Raclopride binds in the class I antagonist binding site, although its binding site begins to extend to the TM2, 7 aromatic micro domain. The critical contact points are the salt bridge between TM3 Asp114 and the protonated amino group of the ligand (2.8 Å), and interactions with the TM3 Cys118 (3.6 Å) and Ser193 (3.3 Å) and Ser197 (4.0 Å), the latter of which is too long to be considered a hydrogen bond. Other residues in close proximity of the ligand include: Phe110, Met117, Cys118 (TM3), Phe164 (TM4), Phe189, Val190 (TM5), Trp386, Phe389, Phe390, His393, Ile394 (TM6), Ser409 and Thr412 (TM7).

2.2.9 Spiperone: Figure 2-13 shows the predicted binding site of spiperone to the receptor. Spiperone is a class II antagonist, with its aromatic moiety extending to the void between TM2 & 7 and TM4 & 6. The protonated amino group of the ligand is salt bridged to the TM3 Asp114 (2.9 Å). There is a hydrogen bond between the fluorine of the ligand and Ser197 (2.8 Å). Another weak, non-ionic interaction is to Trp386 (3.3 Å) on TM6. Other residues that form a hydrophobic pocket for the aliphatic chain of the ligand include Trp90 (TM2), Phe110, Leu113, Met117, Cys118 (all in TM3), Trp160, Phe164 (TM4), Val190, Ser194 (TM5), Phe389, Phe390, His393 (TM6), Ser409, Thr412, Trp413, Tyr416 (all in TM7).

2.2.10 Sulpiride: Figure 2-14 shows the predicted binding site of sulpiride to the receptor. Sulpiride is a class I antagonist and it binds in the putative agonist binding site with minimal extension into the TM2, 7 aromatic micro domain. Sulpiride has a salt

bridge to the TM3 Asp114 (2.8 Å). Interestingly, the same Asp is also hydrogen bonding to the amide hydrogen of sulpiride (2.8 Å). This interaction is also found in all the sulpiride-like ligands. Although the addition of the amide was introduced to decrease the lipophilicity of these drugs, the addition seems to fair well due to the interactions it has with the Asp114 in TM3. The sulfonamide portion of the ligand interacts well with the sequence of Serines in TM5. There are two hydrogen bonds of 2.8 Å and 3.3 Å to Ser193 and Ser197, respectively. The strength of these hydrogen bonds is, however, a different issue. The remainder of the residues provide a mostly hydrophobic pocket. The residues in close proximity include Phe110, Leu113, Met117, Cys118 (TM3), Trp160, Phe164 (TM4), Phe189, Val190, Ser194 (TM5), Trp386, Phe389, Phe390, His393, Ile394 (TM6), Thr412, Tyr416 (TM7).

2.2.11 YM-09151-2: Figure 3-15 shows the predicted binding site of YM-09151-2 to the receptor. This is a class II antagonist. There is a salt bridge between the protonated amino group and Asp114 (2.8 Å) on TM3. There is also a weaker interaction of ~4 Å to the amide hydrogen of the ligand. There is a very weak interaction between Cys118 and the chlorine of the ligand (3.7 Å). There is one hydrogen bond to Ser193 (2.9 Å). The rest of the ligand is stabilized in the most hydrophobic pocket created by Trp90 (TM2), Phe110, Leu113, Met117, Cys118 (TM3), Phe164 (TM4), Asn186, Phe189, Val190, Ser197 (TM5), Trp386, Phe389, Phe390, His393, Ile394 (TM6), The412, Trp413, Tyr416 (TM7).

2.3 Comparison of predicted binding site of agonists vs. antagonists: Experimental studies (see reference 18 and references therein) have outlined the binding site for agonists and antagonists as shown in **Figure 2-16 to 2-18**. The putative agonist-binding site is located between TM3, 4, 5, & 6, and some residues in the EC2 loop (assuming this loop closes during the process of activation). Class I antagonists, such as Clozapine bind in the putative agonist binding pocket meaning they occupy the void between TM3, 4, 5 and 6. Class II antagonists, such as haloperidol consists of two aromatic domains connected by a linker, which possess a protonated amino group. These antagonists bind in the cavity between TM2, 3, 4, 5, 6, and 7. There are two aromatic microdomains present in hD_2DR.[20] The first aromatic microdomain is located in TM4 and TM6, and is composed of the very symmetric Trp160 (TM4), Phe164 (TM4), Trp386 (TM6), and Phe390 (TM6). This aromatic microdomain stabilizes one of the aromatic rings of the class II antagonists. TM3 provides the salt bridge to Asp114 on TM3, which stabilizes the ligand in place. TM5 provides weak interactions with heteroatom functionalities such as halogens on the rings of class II antagonists. These interactions are weak but important in recognition of the correct aromatic domain for docking into the cavity. The second aromatic microdomain composed of Trp90 (TM2), Phe110 (TM3), Trp413 (TM7), and Tyr416 (TM7) stabilize the second aromatic ring group of the class II antagonists.

2.3.1 Comparison of agonist binding sites: There is very little difference in the putative agonist-binding site of different agonists. The agonists studied, (dopamine, 7-OH DPAT, Apomorphine, and Bromocriptine) all possess protonated amino groups, which are salt bridged to TM3 Asp114. All of these ligands studied form favorable hydrogen bonding interactions to a network of serines on TM5, although this is not an absolute necessity for

agonism[21]. However, it should be noted that we have not studied any agonist without any hydroxyl group on the aromatic ring in this paper. There is, in every case studied, favorable interaction from the first aromatic microdomain located in TM4 and TM6. His393 (TM6) contributes greatly to ligand binding[22], yet it appears that there few agonists utilize this residue effectively. It is important to note that the agonists studied here effectively interact with both Ser193 and Ser197 in TM5. The interactions to Ser194 is too weak to be considered a hydrogen bond (on the order of ~ 5 Å). Due to structural constraints, we observe that all three serines cannot simultaneously interact with the agonists. It appears that there could at most be two hydrogen bonds to the TM5 serines. Mutation studies on these serines show that they are directly involved in ligand binding[23-27]. Although in the predicted structure Ser194 is not participating in any interactions, it is possible that in a slightly different structure, perhaps one resulting from activation, there could be interactions to Ser194 and Ser197 as opposed to Ser193 and Ser197. All agonists studied cause strong coupling of TM3 and TM5. None of the agonists studied block the motion between TM3 and TM6. Based on structural studies of rhodopsin, it has been established that residues in the cytoplasmic side of TM6 undergo large movements for activation[28-29]. The coupling of TM3 and TM5 by agonists causes a decrease in distance between TM3 and TM5 while allowing for motion between TM3 and TM6.

2.3.2 Comparison of antagonists binding site: The antagonists studied have been classified into two categories: class I antagonists (exemplified by clozapine), which bind in the putative agonist binding pocket; and class II antagonists (exemplified by haloperidol), which bind in the cavity between TM2, 3, 4, 5, 6, and 7.

2.3.2a Class I Antagonists: all the clozapine-like antagonists form a salt bridge to the TM3 Asp114 with their protonated amino group. They do not have two aromatic rings and therefore only utilize the first aromatic microdomain between TM4 and 6. Both class I and class II antagonists form only one weak interaction with the serine network on TM5. In our models the interaction may be with Ser193 or Ser197. At first glance, it appears that both the number and strength of the interactions with the TM5 serines may be important for activation. Interestingly, Payne *et al.* have identified unhydroxylated agonists of the dipropyl-aminotetralin series, for the short isoform of human D_2DR. The unhydroxylated agonists however exhibit only medium affinity for the receptor[21]. The critical distinguishing feature of the predicted binding site of an agonist vs. an antagonist appears to be the relative position of the ligand to TM6. Class I antagonists bury their aliphatic domain deep into the conserved TM6 WXXFF motif. The binding of antagonists at the conserved WXXFF region on TM6 would prevent any motion, specifically the hinge motion between TM3 and TM6 that is necessary for activation. It appears that the presence of one hydroxyl/one hydrogen bond donor/acceptor is not an absolute necessity for antagonism. The class I antagonists are further stabilized by Phe110 in TM3. The cluster of aromatic residues WXXFF on TM6 has been shown to be solvent accessible by substituted-cysteine accessibility experiments[30].

2.3.2b Class II Antagonists: the haloperidol-like antagonists salt bridge to TM3 Asp114 with their protonated amino group. They possess two aromatic ring units. In most cases, only one of the aromatic rings is halogenated, and this is the ring that binds to the first aromatic microdomain, with the second ring binding in the second aromatic microdomain[31]. The halogenated ring binds effectively into the cavity between TM4 and

TM6 and forms weak interactions with either Ser193 or Ser197 in TM5. It gains stability from the presence of TM4 Trp160 and Phe164 and TM6 Trp386 and Phe390. In some cases His393 (TM6) may also stabilize the class II antagonists. The non-halogenated aromatic domain is located in the void between TM2 and TM7. It gains stabilization from Trp90 (TM2), Phe110 (TM3), Trp413 (TM7), and Tyr416 (TM7). As is the case with the class I antagonists, class II antagonists prevent motion between TM3 and TM6 by burying their aliphatic portion between TM3 and TM6, thereby preventing interaction of these helices.

2.4 Comparison of calculated binding energies to the experimental inhibition constants: Figure 2-18 shows the calculated binding energies (relative to dopamine), for the 9 ligands with experimental dissociation constants. The measured experimental values of dissociation constants vary over orders of magnitude for a given ligand. Although this may be understandable from the experimental point of view, it is not adequate for comparison to calculated binding energies. However the calculated binding energy always falls within this range of experimental values. To provide a visual idea of the correlation between experiments and theory, we performed a least-square fit to the nearest experimental values but by always taking dopamine into account, and the best correlation factor calculated is 0.92 with bromocriptine as an outlier to this fit. At this point it is most scientific to compare the relative ordering of ligand affinities to the calculated ordering of ligand binding energies. **Table 2-2** shows the binding energy of all ligands studied in the library. Our calculated binding energies agree well with the experimentally determined binding constants for the studied ligands. Our binding cavities

also agree well with the mutation, side chain studies, and activation studies available in the literature.

2.5 Computations of the effect of mutations using the hD_2DR (MS) structure:

We have performed computational alanine scanning studies of the residues in the binding cavity of the agonist dopamine, and the antagonist haldoperidol. The alanine scanning computations were performed by identifying the residues that are within 5Å of the bound ligand, and mutating them to Ala while optimizing the side chain rotamer conformation of other residues within 5Å. The side chain rotamer optimizations were performed using SCREAM, a side-chain replacement program (Kam, Vaidehi and Goddard, in preparation) that uses a side-chain rotamer library of 1478 rotamers with 1.0Å resolution, with the all-atom DREIDING energy function to evaluate the energy for the ligand-protein complex. The side chain conformations of other residues that are in the binding pocket and not mutated to Ala are also optimized using SCREAM. The potential energy of the complex is minimized after the side chain placements using conjugate gradients and the binding energy calculated as described in the Methods section.

2.5.1 Dopamine alanine scanning mutations: (**Table 2-3**) The mutation of 7 residues to alanine has significant effect on the binding of the ligand to the hD_2DR receptor. Mutating each of the residues Phe110 (TM3), Asp114 (TM3), Cys118 (TM3), Ser197 (TM5), Phe390 (TM6), His393 (TM6), and Thr412 (TM7) to Ala reduces the binding energy of dopamine significantly. Phe110 (TM3) has been implicated as an important player in the D_2/D_4 specificity of certain antagonists[31]. It appears that interactions with this aromatic residue may be important in stabilizing antagonists, through possibly a cation/pi interaction with the protonated amino group of the ligand. Experimental studies

on the β2 adrenergic receptor, have established the importance of the conserved TM3 aspartate in ligand recognition[32]. The alanine mutation of this residue effectively removes the receptor's ability to bind both agonists and antagonists. In our study, the alanine mutation of Asp114 also reduces the binding energy of the ligand significantly. Alberts and coworkers have established the importance of Cys114 (TM3) in the D_3 subtype of the human dopamine receptor in recognizing agonists[33]. We also identify the corresponding residue, Cys118 (TM3), in the D_2 subtype as being critical for dopamine recognition and binding. Alanine scanning results by two groups[23-24] implicate Ser197 (TM5) as being important in agonist binding. Similar analysis using the computational alanine scanning mutation also highlights the importance of Ser197 in ligand binding. Cho et al. have performed alanine mutation studies and determined inhibition constants for the binding of both dopamine and Apomorphine (another agonist) to the mutated D_2DR.[34] In their studies, Cho et al. show the critical nature of Phe390 (TM6). The alanine mutation of Phe390 completely removes any specific binding of both dopamine and Apomorphine to D_2DR. The same group has determined that the Phe389Ala mutation is also important for ligand binding, although significantly less important than the Phe390Ala mutation. In our studies we do not find that the Phe389Ala mutation has a significant effect on binding, but do find that Phe390 is important for both dopamine and Apomorphine binding to the receptor. Lundstrom and coworkers, studying the D_3 subtype of the human dopamine receptor, have identified the importance of both the histidines on TM6 and the Thr on TM7 in agonist binding [35]. The predicted model also highlights the importance of these residues in dopamine binding. Both residues contribute positively to the binding energy of dopamine and other agonists in the active site.

Despite the importance of the above-mentioned residues in binding, mutation of 4 residues appears to improve binding of dopamine to hD_2DR. The alanine mutation of Val111 (TM3), Met117 (TM3), Trp160 (TM4), and Phe164 (TM4) improve the binding of dopamine by about 3-5 kcals/mol for Val111, Trp160, and Phe164 and by about 13 kcal/mol for Met117. No similar alanine scanning mutations have been reported for any of these residues. It appears that in essentially every case the alanine mutation improves binding due to the removal of van der Waals clash between the native residue and the ligand. In the case of large agonists such as Apomorphine this van der Waals clash becomes exceedingly important.

2.5.2 Haloperidol Alanine Scanning Mutations: Haloperidol or haldol is a class II antagonist that occupies the water accessible void between TM2, 3, 4, 5, 6, and 7. Haldol also forms a tight salt bridge to the TM3 Asp and contains only minimal contacts to the TM5 Ser network (the nature of the contacts with the TM5 Ser cluster is not only weak but only one of the Ser is involved in a hydrogen bond). **Table 2-4** shows the calculated the change in binding energy caused by alanine substitution of all residues within the 5 Å binding site of haldol, where positive represents an unfavorable change and a negative value represents a beneficial change.

Eight residues were identified as being important for antagonist binding by the computational alanine-scanning algorithm. These residues were picked as those residues that contributed 2+ kcal/mol to the binding of the ligand. Trp90 (TM2), Asp114 (TM3), Cys118 (TM3), Phe164 (TM4), Ser193 (TM5), Phe390 (TM6), His393 (TM6), Ser409 (TM7) are the residues which provide the favorable contacts that stabilize haldol in the ligand binding site. Trp90 (TM2) has been identified as being present in the active site of

antagonists[31]. Val91 (TM2) was also identified by Simpson et al. to be important in the D2/D4 specificity of ligands. In our studies, we find that Val91 (TM2) does not seem to contribute positively or negatively to the binding of haldol although the mutation of Val91 to Phe should contribute to better stability of the class II antagonists to the D2 receptor. Asp114 (TM3) has been shown to be important in antagonist binding in the work by Strader and coworkers[32] performed on the □2 adrenergic receptor**.** This residue is essential to antagonist binding and its alanine mutation causes complete removal of ligand binding. Cys118 (TM3) has been extensively studied and utilized in the design of new antagonists to D_2DR. Studies by Javitch *et al.* have implicated this residue in binding and recognition of both the class I and II antagonists[36]. We also identify this residue as being critical to the binding of the ligand to the D_2DR. Phe164 (TM4) is a critical contact for the aromatic moieties of the class I and class II antagonists. Although we have not identified any experimental studies involving the direct mutation and calculation of binding affinities of ligands to the mutant, we predict that this mutation should decrease binding of ligands to the D2 receptor. Ser197 has been determined to be important for antagonist binding[37]. Both Ser 193 and Ser 197 have been shown to affect ligand binding in our calculations. As is the case with the agonist-binding site, the antagonist binding site utilizes the aromatic residues in TM6 for the stabilization. Studies by Javitch *et al.* (1998) and Cho *et al.* (1995) have highlighted the importance of Trp386 and Phe389 on TM6 in the binding of ligands.[30,34] Although we identify both residues as being essential to binding, we find that Phe390 (TM6) is the most important member of the aromatic microdomain. Indeed Cho *et al*. have shown that the mutation of this residue completely removes binding. Javitch *et al*. (1998) have also demonstrated the importance of His393

in recognizing antagonists. We find similarly that the alanine mutation of His393 has a large deleterious affect on ligand binding. Ser409 has not been reported as being important for the binding or recognition of antagonists. We suggest this residue as a possible mutation for the experimentalists.

As is the case with the agonist-binding site, the mutation of some residues in the antagonist binding site to alanine actually improves binding. For the case of haloperidol, Val115 (TM3), Met117 (TM3), Ile122 (TM3), and Trp160 (TM4). Two of these residues, Val115 and Trp160 have been studied experimentally and have been shown to be present in the binding cavity of antagonists by SCAM studies[38]. Although these residues are indeed present in the binding site of both class I and II antagonists, it appears that the van der Waals contact between these residues and the aromatic rings of the ligands may actually destabilize the ligand in the receptor. Met117 and Ile122 can be directly tested by mutations to Ala.

2.5.3 Alanine scanning discussion:

The alanine scanning mutations show an excellent agreement with the experimentally determined residues that interact with both agonist and antagonists. For the agonist-binding pocket, we identify Asp114 (TM3) and the Serine network in TM5 (mainly Ser193 and Ser197) as being the major contributors to the binding of the ligand. Other residues such as His393 (TM6) and Thr412 (TM7) are consistently present in the binding site of other agonists and seem to contribute favorably to binding. Several non-polar residues are also present in the binding site of all agonists; these residues include Phe110 (TM3), Trp160 (TM4), Phe164 (TM4), Trp386 (TM6), and Phe390 (TM6). Although these residues seem to contribute to the binding of agonists, the mutation of Phe110

(TM3) and Trp160 (TM4) actually improves binding in most cases. Agonists chemically contact all TM5 Ser. The conserved TM3 Cys118 seems to contribute very little to the binding of the agonists, although better design may take advantage of the presence of the Cys and could yield better binding agonists for this and other dopamine receptors. Dopamine and the agonists studied do not extend into the TM2 and TM7 void and therefore do not utilize any residues in this region to a great amount. We have identified only one single loop residue Asn186 (EC2) as being involved in the binding of agonists. This is mainly attributed to our inability to model the loops at this time. Further refined studies may highlight the importance and contribution of other EC2 loop residues to agonist binding.

The antagonist binding pocket and the residues in contact with antagonists are consistent with those experimentally determined. There are two classes of antagonists: the class I antagonists (occupy the void usually covered by dopamine meaning the water accessible crevice between TM3, 4, 5 and 6), and class II antagonists (occupy the void between TM2, 3, 4, 5, 6, and 7). As is the case in the agonist binding pocket, Asp114 (TM3) is the counter ion for the charged amino group. The network of TM5 Ser make very little contacts with both class I and class II antagonists. Unlike the agonists which interact with essentially all three Ser in TM5, the antagonists make good chemical contacts to only one TM5 Ser, either Ser193 or Ser197. Agonists usually bear a hydroxyl in a position which can interact with the Ser network, whereas antagonists use halogens which posses half as strong a hydrogen bond relative to a hydroxyl group. Experimental observations have shown that minimal contact to TM5 Ser may be the mechanism of antagonism used by these drugs. The main difference between class I and class II

antagonists may is in the aliphatic/aromatic residues that they contact. Class I antagonists reside in the putative dopamine binding pocket and will only benefit from the first aromatic micro domain in TM4 and 6 (Trp160 (TM4), Phe164 (TM4), Trp386 (TM6) and Phe390 (TM6)). Class II antagonists, use both the first and second aromatic clusters (the second aromatic cluster consists of Trp90 (TM2), and Tyr416 (TM7)). Whereas tricyclic agonists bind rather well into the D2 receptor, the tricyclic antagonists (such as clozapine) do not bind very favorably. This is attributed to the Van der Waals clashes with several key aliphatic residues such as Phe110 (TM3). The reason for poor binding of the antagonists seems to be that they are restricted into a binding mode that causes them to clash with the aliphatic, whereas tricyclic agonists are rather free to adopt a favorable binding mode in the cavity. The binding mode of the tricyclic agonists and antagonists is perpendicular to each other ensuring that only the agonist can make the critical salt bridge (TM3) and hydrogen bonds (TM5), when perpendicular to the plane of the transmembrane helices, where as the antagonist must orient more parallel to the transmembrane helices in order to bind properly to the TM3 Asp and TM5 Ser. As is the case with agonists, most alanine mutations actually deter binding of ligands to this receptor. There are however, residues in the antagonist binding cavity that when mutated actually are beneficial for binding, the most important residue being Trp160 (TM4).

3.0 Homology Modeling:

To compare the first principles predicted hD_2DR (MS) structure to the model for hD_2DR based on homology to bovine rhodopsin, we have predicted the homology model of hD_2DR using the homology modeling program MODELER[39] based on a ClustalW alignment of the sequence to bovine rhodopsin. We have compared this model with the

published Vriend homology model, and the D_2 structure obtained via the Membstruk procedure. Our studies highlight the shortcomings and failures of both rhodopsin based models and the strengths of the Membstruk procedure and homology models based on the Membstruk structures.

3.1 D_2 Homology Model Based on Bovine Rhodopsin Crystal Structure: Based on the crystal structure of bovine rhodopsin (pdb code: 1F88) we derived a homology-based model for hD_2DR using MODELER method. A simple sequence alignment of the human D_2 sequence to the bovine rhodopsin sequence using ClustalW multiple sequence alignment and subsequent homology modeling based on the crystal structure produces a structure with unreliable and abnormal bulges and bends in the helices. A comparison of the sequence shows a large gap in the alignment that coincides with the TM5 bulge present in the structure. Additionally, a simple alignment to the rhodopsin sequence places the highly conserved TM5 serines in the second extracellular loop. Clearly, homology modeling for sequences with low sequence similarity needs experimental results as inputs to refine the model. Thus homology modeling is not sufficient for structure prediction of GPCR unless experimental information is directly used for the models. However refined homology models have been derived by many research groups (Teeter, Mosberg, Trumpp-Kallmeyer, Javitch, Weinstein) for hD_2DR. Since there are several research groups involved in homology based modeling for D_2DR, we might have inadvertently omitted referring to some of them. In the next section we show the function prediction studies we have conducted using the homology model of hD2DR modeled by Vriend and is available at the web site http://www.gpcr.org.

3.2 D_2 Homology Vriend Model Based on the Rhodopsin X-ray Structure and Sequence Alignment*:* The Vriend model was utilized for further modeling studies. Since only the structure of the TM domains was provided in the database, the program WhatIf was utilized to add the extracellular and intracellular loops using the same procedure used for the hD2DR (MS) structure prediction. The potential energy of all the atoms of the homology model based on Vriend's prediction of TM domain was minimized to 0.1kcal/mol/Å in RMS in force using conjugate gradients. The D_2 Vriend model, referred to as hD_2DR (Vriend), is a more reasonable structure with a root mean square deviation of 8.87 Å in the C_{α} atoms and 9.23 Å over all equivalent atoms versus the bovine rhodopsin structure. The Vriend structure does not have any bulges or unrealistic kinks. It must be noted that akin to most homology structures, those of GPCRs involve a great deal of manual intervention to derive a model: i.e., the rotation of the TM helices are done based on the location of conserved residues and distance restraints derived from experiment and not based on any physical interpretations of the structure (for an example, see reference 18). Comparison of the homology model obtained by using Modeler with automatic sequence alignment from clustalW to the hD_2DR (Vriend) shows a RMSD of 12.02 Å for the C_{α} atoms and 12.60 Å over all equivalent atoms. To further validate the utility of the hD_2DR (Vriend) homology structure, the structure was utilized in docking studies. Scanning of the receptor, however, resulted in the identification of a site that is inconsistent with the experimentally determined agonist/dopamine binding site. The HierDock procedure used for identifying the binding sites identifies a site in the extracellular loops 1 and 2 as the best agonist sites. **Figure 2-19** below shows a graphical representation of the scanning results in the homology model compared to those in the

Membstruk model. It is clear from this figure that the most favorable binding sites in the homology model are in the loops, not in the helical cavities as in the Membstruk model. The experimental agonist-binding site is identified as the worst possible binding region in the hD_2DR (Vriend) model.

The differences in binding site prediction between the homology model and the Membstruk model are most likely due to the differences in packing between the structures. The use of the rhodopsin backbone does not create a model with the correct packing for the dopamine receptor and hence the cavity is too tight for the docking of dopamine. An analysis of the structures shows a 5.34 Å difference in the Cα and a 6.19-Å difference in overall equivalent atoms between the hD_2DR (Vriend) and hD_2DR (MS) structures. Clearly a scanning of the receptor structure to determine the putative binding region illustrates that the homology-modeled structure could not be used in the automated prediction of the binding site and binding residues without prior knowledge of the residues involved in binding.

The calculated binding energies of the 10 pharmaceutical ligands to hD_2DR(Vriend) model do not correlate with experiment. Based on the above inhibition constants, haloperidol and spiperone are predicted to have similar binding energies to this receptor. Despite this experimental observation, the binding energies between the two complexes differ by as much as nearly 17 kcal/mol. For the sake of fairness, however, apomorphine and bromocriptine are known to experimentally have similar binding affinities and the structure does reproduce this well; however, analysis of further cases such as the relative binding affinities of raclopride and spiperone highlights the inconsistencies of this structure. In comparison, the MS structure predicts correct relative

binding affinities for all but one ligand in the study. The binding cavity is strikingly consistent with the experimentally determined binding site, however most modeled structure of this type are built with constraints from experiments. Thus a simple homology structure based on sequence alignments and using the crystal structure of bovine rhodopsin as the backbone does not produce a reasonable model for structure-function studies due to the presence of large gaps, bulges, and kinks. This is certainly true for more complex GPCRs such as the peptide and lipid receptors. A homology structure based on modified alignments (based on conserved residues) and manual rotations may not have predictive value although it could be used for generating a model to explain a set of experimental observations.

4.0 Conclusion: Using the MembStruk and HierDock structure and function prediction methods, we have predicted the structure of the hD_2DR and the ligand binding sites and binding energies. The predicted binding sites are in agreement with the mutation results. With the exception of bromocriptine, the calculated binding energies fall within the range of experimental dissociation constants measured for these ligands. The structure of the binding site of agonists and antagonists led to further predictions of residues that can be directly tested out by mutation studies by experimental research groups. Comparison with homology models based on bovine rhodopsin crystal structure indicates that the homology models often do not have ab initio predictive capacity. Although this does not mitigate the effectiveness of such studies for interpreting biochemical data, these shortcomings indicate the importance of ab initio approaches for obtaining novel structural data about this class of receptors.

5.0 Materials and Methods

5.1 Choice of forcefields (FF): All calculations for the protein used the DREIDING FF[40] with charges from CHARMM22[41] unless specified otherwise. The non-bond interactions were calculated using Cell Multipole Method[42] in MPSim[43]. The ligands were described with the DREIDING FF using Gasteiger charges[44]. For the lipids we used the DREIDING FF with QEq charges[45]. Some calculations were done in the vacuum (e.g., final optimization of receptor structure to approximate the low dielectric membrane environment). Most calculations treated the solvent (water) using the Analytical Volume Generalized Born (AVGB) approximation to Poisson-Boltzmann solvation model[46].

5.2 MembStruk Structure Prediction Method: The MembStruk procedure version MembStruk3.0 used to predict the three dimensional structure of hD_2DR is described in detail in references 8-9. Here we detail the steps that are relevant to the prediction of hD_2DR. The various steps of the MembStruk procedure are as follows:

The seven TM boundaries of the hD_2DR was predicted using TM2ndS[9] procedure. Twenty sequences of D_2DR across many species were aligned using multiple sequence alignment program CLUSTALW[17].

This alignment was used to predict the TM regions using TM2ndS. The predicted TM regions of the human D2 dopamine receptor are shown in **Scheme 2-1**. It is seen that the seven TM helices in hD_2DR are of different length and also are different in length from the corresponding TM helices of rhodopsin. We built 7 canonical a-helices, and then constructed the TM seven helical barrel with the helical axes positioned based on the 7.5 Å three-dimensional density map of frog rhodopsin[47].

(a) We then performed optimization of the translational orientation of the canonical helices by using the hydrophobic center algorithm described in reference 9. The maximum hydrophobic centers of the seven helices are residue 17 for TM1, residue 13 for TM2, residue 11 for TM3, residue 11 for TM4, residue 13 for TM5, residue 15 for TM6 and residue 16 for TM7. These hydrophobic centers were fitted to a plane and thus an optimum of relative translational orientation of the helices was obtained.

(b) The rotational orientation of the canonical helices was also optimized using the multisequence hydrophobicity moments, of the middle third of each helix about their maximum hydrophobic center. These analyses yielded a clear consensus on which residues should contact the membrane and which should face the receptor interior.

(c) The canonical helices were optimized with NEIMO torsional dynamics[48-49] or Cartesian dynamics (described with the DREIDING FF and Charmm22 charges), for 500 ps at 300 K constant temperature and picked the minimum energy conformation from the dynamics. This step optimizes the kinks and bends in the helices.

(d) The helical bundle now has helices with bends and kinks. The rotational orientation of these non-canonical helices was further optimized using both the procedure in step c) followed by energy based optimization called “Rotmin” described in reference 9. Steps c), d) and f) is a part of systematic search algorithm for optimum translational and rotational orientation and these steps aid in getting over large barriers for structure optimization.

(e) The optimized TM barrel structure was then equilibrated by immersing it in a bilayer barrel of dilauroylphosphatidyl choline and the full system was optimized with rigid body quaternion molecular dynamics (MD), treating each molecule as a rigid body for 50ps at constant temperature of 300K using MPSim code.

(f) The interhelical loops were built using WHAT IF[50] and disulfide bonds were formed between Cys 107 in TM3 and Cys 182 in extracellular loop 2. This full system was then optimized with conjugate gradient minimization technique to 0.1kcal/mol/Å RMS in force.

5.3 Prediction of Ligand Binding Sites and Binding Energies:

5.3.1 Ligand Structure Preparation: The 10 ligands shown in **Figure 2-4** were built with chemdraw and the two dimensional structure was converted to three dimensional structures in cerius2. The pharmaceuticals consist of agonists, dopamine, 7OH-DPAT, apomorphine, bromocriptine, and antagonists like clozapine, domperidone, haloperidol, etc. We have categorized the antagonists studied here into two classes: 1) class I, clozapine-like bulky antagonists and 2) class II antagonists that have two aromatic or ring moieties conencted by a flexible linker with a protonated amine group like haloperidol. Hydrogens were added with Gasteiger charges assigned also using the concord software. We then minimized the potential energy of each ligand using conjugate gradients to a RMSD in force of 0.1kcal/mol/Å.

5.3.2 Function Prediction: HierDock protocol is a hierarchical strategy of ranging from coarse-grain docking to fine-grain optimization for docking ligands in proteins. This method has been tested for various GPCRs[8-14], membrane proteins[51] and globular proteins[52-56]. This protocol has been described in detail in these references. In here we use

the version of HierDock2.0 described in reference 8. In brief the various steps of HierDock protocol version 2.0 is as follows:

The HierDock ligand screening protocol follows a hierarchical strategy for examining ligand binding conformations, and calculating their binding energies. The steps are as follows:

a) First we carry out a coarse grain docking procedure to generate a set of conformations for ligand binding in the receptor. Here we use Dock 4.0[57] to generate and score 1000 configurations, of which 10% (100) were selected using a buried surface area cutoff of 90% and using energy scoring from Dock4.0, for further analysis. The options used in Dock4.0 are flexible ligand docking with torsion drive and allowing four bumps.

b) The 100 best conformations selected for each ligand from step a) are subjected to all-atom minimization keeping the protein fixed but the ligand movable. The solvation of each of these 100 minimized structures was calculated using the Analytical Volume Generalized Born (AVGB) continuum solvation method[46]. Then the 10 best structures based on the potential energy of the ligand in the protein, were selected from these 100 structures for the next step.

c) Next we optimize the structure of the receptor/ligand complex allowing the structure of the protein to accommodate the ligand. This is essential to identify the optimum conformations for the complex. The all-atom receptor/ligand energy minimization was performed on the 10 structures from the previous step. Using these optimized structures, we calculate the binding energy (BE) using the equation

$$BE = PE \text{ (ligand in protein)} - PE \text{ (ligand in solvent)} \quad (1)$$

as the difference between the energy of the ligand in the protein and the energy of the ligand in water. The energy of the ligand in water is calculated using DREIDING FF and the SGB or AVGB continuum solvation method[46].

d) Next we select from the five structures from step 3, the one with the maximum number of hydrogen bonds between ligand and protein. For this structure we use the SCREAM side chain replacement program to reassign all side chains for the residues within 4Å in the binding pocket [this uses a side-chain rotamer library (1478 rotamers with 1.0Å resolution) with the all-atom DREIDING energy function to evaluate the energy for the ligand-protein complex]. The binding energy of all the 5 optimized complexes is calculated.

5.3.3 Locating the Putative Binding Site: To locate the binding site of dopamine, other agonists and antagonists, we scanned the entire D2DR structure without any knowledge of the binding site. The molecular surface of the entire receptor structure was mapped using autoMS utility of DOCK4.0[57]. Spheres were generated to fill up the void regions of the entire receptor using sphgen utility of Dock4.0. The program "Pass"[58] was then used to locate plausible centers of large void regions in the receptor. The spheres that are within 5.0Å of these centers are gathered for docking of ligands. For D2DR we obtained 9 regions where we applied the ScanBindSite protocol for each region, with the following docking steps:

5.3.3a ScanBindSite: Figure 2-19 shows the results of the "ScanBindSite" procedure for D2DR for all the 11 ligands studied here. It is clear from the table that for dopamine, agonists and class I antagonists region 1 is dominantly better in energy than any other region. For class II long antagonists it is seen that regions 1, 4 and 11 are similar in energy. A tolerance of 30kcal/mol is used since the minimization of the ligand with the fixed receptor structure is done for a fixed number of 50 steps.

5.3.3b Prediction of binding sites and binding energies: We used HierDock protocol steps a) to d) to dock the agonists and class I antagonists region 1. For the class II antagonists, we merged the spheres in regions 1, 4 and 11 and thinned the density of spheres to about 150. This resulted in 77 number of spheres across regions 1, 4 and 11. HierDock protocol steps a) to d) were applied to these regions and the best 5 bound structures for each ligand was chosen.

5.3.3c Refinement of the bound structures: The binding site of the best-bound structures for each ligand was further refined using the following procedure. The docked structures were fully minimized for 5000 steps or 0.1 RMS deviations. Residues in the 5.5 Å vicinity of the ligand were replaced with alanine. Conjugate gradient minimization was carried out for 5000 steps or 0.1 kcal/mol/Å RMS deviations to relax the ligand in the active site. This would allow the ligand to optimize in the putative binding cavity. The side chain rotamers of the residues were replaced using SCREAM side chain placement program and the ligand/receptor complex was again minimized in energy for 5000 conjugate gradient steps or 0.1kcal/mol/Å RMS deviations. The binding energies were calculated suing equation (1).

6.0 Acknowledgements. This research was partially supported by NIH-BRGRO1-GM625523, NIH-R29AI40567, NIH-HD36385, and the computational facilities were provided by a SUR grant from IBM and a DURIP grant from ARO. The facilities of the Materials and Process Simulation Center are also supported by DURIP-ARO, DURIP-ONR, DOE (ASCI ASAP), NSF (CHE and MRI), MURI-ARO, MURI-ONR, General Motors, ChevronTexaco, Seiko-Epson, Beckman Institute, and Asahi Kasei.

7.0 References:

1. Jackson D. M., Westlind-Danielsson A. *Pharmacol. Ther.* 64, 291-369 (**1994**).

2. Strange P.G. *Adv. Drug Res*. 28, 313-351 (**1996**).

3. Kobilka B.K., Kobilka T.S., Daniel K. *Science* 240, 1310 (**1988**).

4. Teeter M.M., Froimowitz M., Stec B., DuRand C.J. *J. Med. Chem*. 37, 2874-2888 (**1994**).

5. Trumpp-Kallmeyer S., Hoflack J., Bruinvels A., Hibert M. *J. Med. Chem.* 35, 3448-3462 (**1992**).

6. Neve K.A., Cumbay M.G., Thompson K.R., Yang R., Buck D.C., Watts V.J., DuRand C.J., Teeter M.M. *Mol. Pharmacol*. 60, 373-381 (**2001**).

7. Varady J., Wu X., Fang X., Min J., Hu Z., Levant B., Wang S. *J. Med. Chem*. 46, 4377-4392 (**2003**).

8. Vaidehi N., Floriano W., Trabanino R., Hall S., Freddolino P., Choi E.J., Zamanakos G., Goddard, W.A. *Proc. Natl. Acad. Sci. USA* 99, 12622-12627 (**2002**).

9.Trabanino R., Hall S.E., Vaidehi N., Floriano W., Goddard, W.A. *Biophys. J.*, in press (**2004**).

10. Freddolino P., Kalani M.Y., Vaidehi N., Floriano W., Hall S.E., Trabanino R., Kam V.W.T., Goddard, W.A. *Proc. Natl. Acad. Sci. USA* 101, 2736-2741 (**2004**).

11. Floriano W. B., Vaidehi N., Goddard W. A., III, Singer M. S., Shepherd G. M. *Proc. Natl. Acad. Sci. USA* 97**,** 10712–10716 (**2000**).

12. Floriano W.B. Vaidehi N., Goddard W.A., III, *Chemical Senses*, in press (**2004**).

13. Hall S.E., Vaidehi N., Floriano W.B., Goddard W.A., III, *Chemical Senses*, in press (**2004**).

14. Kalani M.Y., Vaidehi N., Hall S.E., Trabanino R., Freddolino P., Kalani M.A., Floriano W.B., Kam V., Goddard W.A., III *Proc. Natl. Acad. Sci. USA* 101, 3815-3820 (**2004**).

15. As described in reference 5.

16. Palczewski K., Kumasaka T., Hori T., Behnke C. A., Motoshima H., Fox B. A., Le Trong I., Teller D. C., Okada T., Stenkamp R. E. *Science* 289**,** 739–745 (**2000**).

17. Higgins D., Thompson J., Gibson T., Thompson J.D., Higgins D.G., Gibson T.J. *Nucleic Acids Res.* 22, 4673-4680 (**1994**).

18. Shi L., Javitch J. *Annu Rev. Pharmacol. Toxicol*. 42, 437-67 (**2002**).

19. Javitch J.A., Ballesteros J.A., Chen J., Chiappa V., Simpson M.M. *Biochemistry*, 38, 7961-7968 (**1996**).

20. Javitch J. A., Ballesteros J. A., Weinstein H., Chen J. A. *Biochemistry* 37, 998-1006 (**1998**).

21. Payne S.L., Johansson A.M., Strange P.G. *J. Neurochem.* 82, 1106-1117 (**2002**).

22. Woodward R., Daniell S.J., Strange P.G., Naylor L.H. *J. Neurochem.* 62, 1664–1669 (**1994**).

23. Wilcox R.E., Huang W.H., Brusniak M.Y., Wilcox D.M., Pearlman R.S. *J. Med. Chem.* 43, 3005-3019 (**2000**).

24. Mansour A., Meng F., Meador-Woodruff J.H., Taylor L.P., Civelli O., Akil H. *Eur. J. Pharmacol.* 227, 205–214 (**1992**).

25. Cox B. A., Henningsen R. A., Spanoyannis A., Neve R. L. and Neve K. A. *J. Neurochem.* 59, 627–635 (**1992**).

26. Strange P. G. *Trends Pharmacol. Sci*. 17, 235–271 (**1996a**).

27. Strange P. G. *Trends Pharmacol*. *Sci.* 17, 346 (**1996b**).

28. Greasley PJ, Fanelli F, Rossier O, Abuin L, Cotecchia S. *Mol. Pharmacol.* 61, 1025-1032 (**2002**).

29. Hubbell WL, Altenbach C, Hubbell CM, Khorana HG. *Adv Protein Chem*. 63, 243-290 (**2003**).

30. Javitch J.A., Ballesteros J.A., Weinstein H., Chen J. *Biochemistry* 37, 998–1006 (**1998**).

31. Simpson M.M., Ballesteros J.A., Chiappa V., Chen J., Suehiro M., Javitch, J.A. *Mol. Pharmacol.* 56, 1116–26 (**1999**).

32. Strader C.D., Sigal I.S., Register R.B., Candelore M.R., Rands E., Dixon R.A. *Proc. Natl. Acad. Sci. USA* 84, 4384–88 (**1987**).

33. Alberts G.L., Pregenzer J.F., Im W.B. *Br. J. Pharmacol.* 125, 705–710 (**1998**).

34. Cho W., Taylor L.P., Mansour A., Akil H. *J. Neurochem.* 65, 2105–2115 (**1994**).

35. Lundstrom K., Turpin M.P., Large C., Robertson G., Thomas P., Lewell X.Q. *J. Recept. Signal Transduct. Res.* 18, 133–150 (**1998**).

36. Javitch J.A., Fu D., Chen J. *Mol.Pharmacol.* 49, 692–698 (**1996**).

37. Javitch J.A., Fu D., Chen J. *Biochemistry* 34, 16,433–16, 439 (**1995**).

38. Javitch J. A., Shi L., Simpson M. M., Chen J., Chiappa V., Visiers I., Weinstein H., Ballesteros J. A. *Biochemistry* 39, 12190-12199 (**2000**).

39. Mayo S. L., Olafson B.D., Goddard III, W.A. *J. Phys. Chem.* 94, 8897-8909 (**1990**).

40. Sali A., Potterton L., Yuan F., van Vlijmen H., Karplus M. *Proteins* 23**,** 318 (**1995**).

41. MacKerell A.D., Bashford D., Bellott M., Dunbrack R.L., Evanseck J.D., Field M.J., Fischer S., Gao J., Guo H., Ha S., Joseph-McCarthy D., Kuchnir L., Kuczera K., Lau F.T.K., Mattos C., Michnick S., Ngo T., Nguyen D.T., Prodhom B., Reiher W.E., Roux B., Schlenkrich M., Smith J.C., Stote R., Straub J., Watanabe M., Wiorkiewicz-Kuczera J., Yin D., Karplus M. *J. Phys. Chem. B* 102, 3586-3616 (**1998**).

42. Ding H. Q., Karasawa N., Goddard III, W. A. *J. Chem. Phys.* 97, 4309 (**1992**).

43. Lim K-T, Brunett S., Iotov M., McClurg R.B., Vaidehi N., Dasgupta S., Taylor S., Goddard III, W.A. *J. Comput. Chem.* 18, 501-521 (**1997**).

44. Gasteiger J., Marsili M. *Tetrahedron* 36, 3219-3228 (**1980**).

45. Rappé A.K., Goddard III, W.A. *J. Phys. Chem.* 95, 3358-3363 (**1991**).

46. Zamanakos G., Physics Doctoral Thesis, California Institute of Technology, (**2001**).

47. Schertler G.F.X. *Eye* 12, 504-510 (**1998**).

48. Jain A., Vaidehi N., Rodriguez G. *J. Comp. Phys.* 106, 258-268 (**1993**).

49. Vaidehi N., Jain A., Goddard III, W.A. *J. Phys. Chem.* 100, 10508-10517 (**1996**).

50. Vriend G. *J. Mol. Graph.* 8, 52-56 (**1990**).

51. Datta D., Vaidehi N., Floriano W. B., Kim K. S., Prasadarao N. V., Goddard W. A., III *Proteins Struct. Funct. Genet.* 50**,** 213–221 (**2003**).

52. Datta D., Vaidehi N., Xu X., Goddard W. A., III *Proc. Natl. Acad. Sci. USA* 99**,** 2636–2641 (**2002**).

53. Zhang D., Vaidehi N., Goddard W. A., III, Danzer J. F., Debe D. *Proc. Natl. Acad. Sci. USA* 99**,** 6579–6584 (**2002**).

54. Wang P., Vaidehi N., Tirrell D. A., Goddard W. A., III *J. Am. Chem. Soc.* 124**,** 14442–14449 (**2002**).

55. Kekenes-Huskey P. M., Vaidehi N., Floriano W. B., Goddard W. A., III *J. Phys. Chem. B* 107**,** 11549–11557 (**2003**).

56. Floriano W. B., Vaidehi N., Zamanakos G., Goddard W. A., III *J. Med. Chem.* 47**,** 56–71 (**2004**).

57. Ewing T.A., Kuntz I.D. *J. Comput. Chem.* 18, 1175-1189 (**1997**).

58. Brady P., Stouten F.W. *J. Comp. Molec. Des.* 14, 383-401 (**2000**).

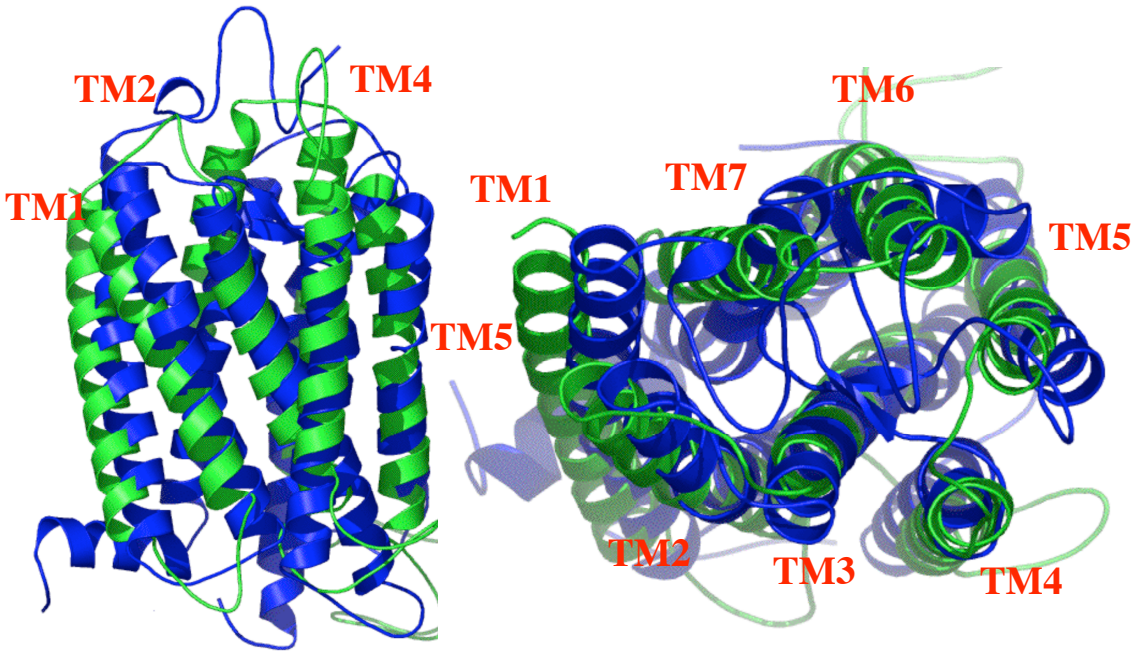

Figure 2-1. Structure for human D_2 dopamine receptor predicted using MembStruk (green) overlaid upon the 2.8-Å crystal structure of BovRhod (blue).

TM Helix	Helical Bend	Plane Tilt	HPM Angle	HPM Mag.	HPM Fit	Plane CM Dist.	Plane CM Angle	CM Fit
1	9.4	34.7	10.7	9.5	0.0	15.7	0.0	1.5734
2	12.9	32.7	137.3	6.8	0.0	10.1	35.3	0.2371
3	4.3	15.6	64.1	2.8	0.0	4.1	126.0	2.6519
4	5.5	4.3	72.5	9.3	0.0	14.6	126.5	-2.5712
5	9.7	12.3	-100.0	8.4	0.0	14.3	189.2	2.0554
6	19.6	14.3	-54.7	10.5	0.0	11.7	245.9	1.7931
7	0.0	17.3	-3.5	6.1	0.0	10.0	306.9	-5.7397

Table 2-1. Important structural parameters of the human dopamine receptor.

Ligand	Experiment	Theory	B.E.	D114	S193	S197
7OHDPAT	Ag	Ag	27.42	X	X	X
Apomorphine	Ag	Ag	31.23	X	X	X
Bromocriptine	Ag	Ag	61.99	X	X	X
Clozapine	At	At	14.72	X	X	
Domperidone	At	At	27.02	X		X
Dopamine	Ag	Ag	0	X	X	X
Haloperidol	At	At	27.69	X		X
Raclopride	At	At	22.47	X	X	
Spiperone	At	At	23.52	X		X
Sulpiride	At	At	24.72	X	X	X
YM-09151-2	At	At	23.28	X	X	

Table 2-2. The binding energy (kcal/mol) relative to dopamine of the ligands used in this study. Based on the criteria of one hydrogen bond to the sequence of TM5 serines for antagonism and two hydrogen bonds for agonism, we have correctly classified the ligands as agonists or antagonists. Sulpiride is the only antagonists which has interactions with both Ser193 and Ser197 but is an antagonist; this observation suggests that not only number but also strenghts of the hydrogen bonds are important factors.

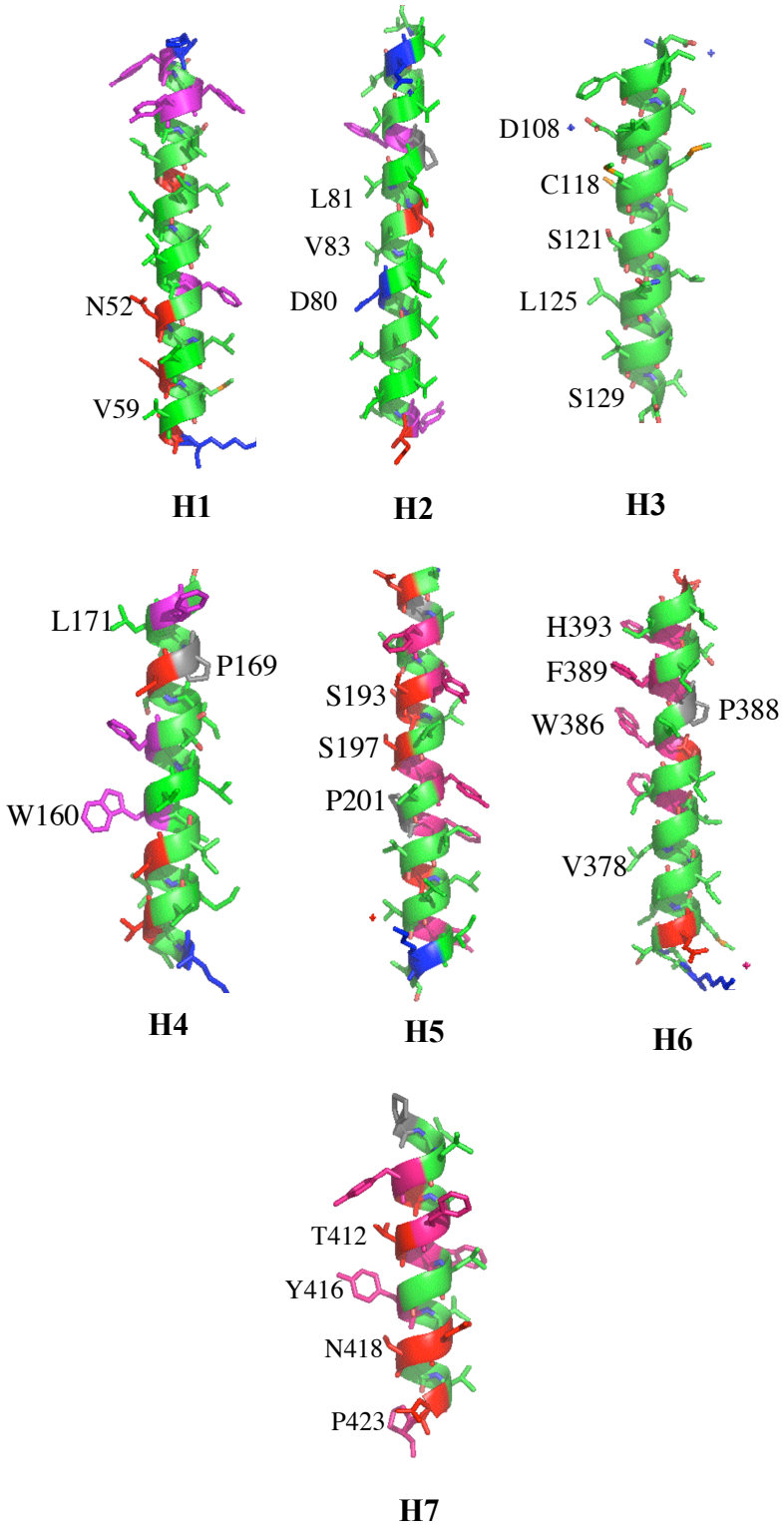

Figure 2-2a-g. A helix-by-helix analysis of the hD2DR (MS). Charged residues (blue), aromatics (purple), and polar (red) residues are highlighted. The right side of each helix is in contact with the ligand, while the right hand side is facing the water accessible crevice. The third transmembrane helix is not in direct contact with the lipid and is not color-

coded according the other helices in direct contact with the lipid bilayer. SCAM studies (as reviewed by Shi and Javitch (reference 18)) identify the presence of the following residues in each helix in the water accessible crevice: **H1**: G51, N52, A58, and V59; **H2**: D80, L81, V83, V87, W90, and V91; **H3**: D108, I109, F110, V111, V115, C118, S121, I122, L125, and S129; **H4**: W160, P160, and L171; **H5**: P201, S193, S197; **H6**: I377, V378, W386, F389, F390, T392, H393, I394, and I397; **H7**: L407, F411, T412, W413, Y416, N418, N422, P423.

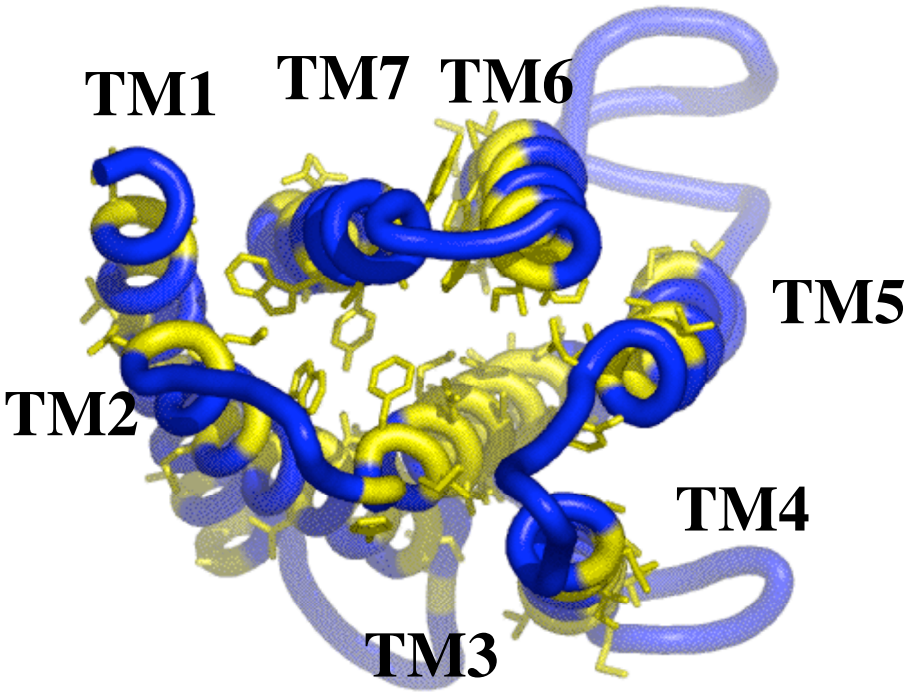

Figure 2-3. This figure shows the predicted structure hD2DR (MS), overlaid with all residues identified experimentally as solvent-accessible (and thus interior to the helical bundle) highlighted in yellow. For all helices except TM4, our predicted structure has TM positions and rotations that correspond very well with experiment. Indeed for TM4 the solvent accessibility study led to discrepancies the experimental orientation of TM4 in rhodopsin, indicating possible experimental difficulties.

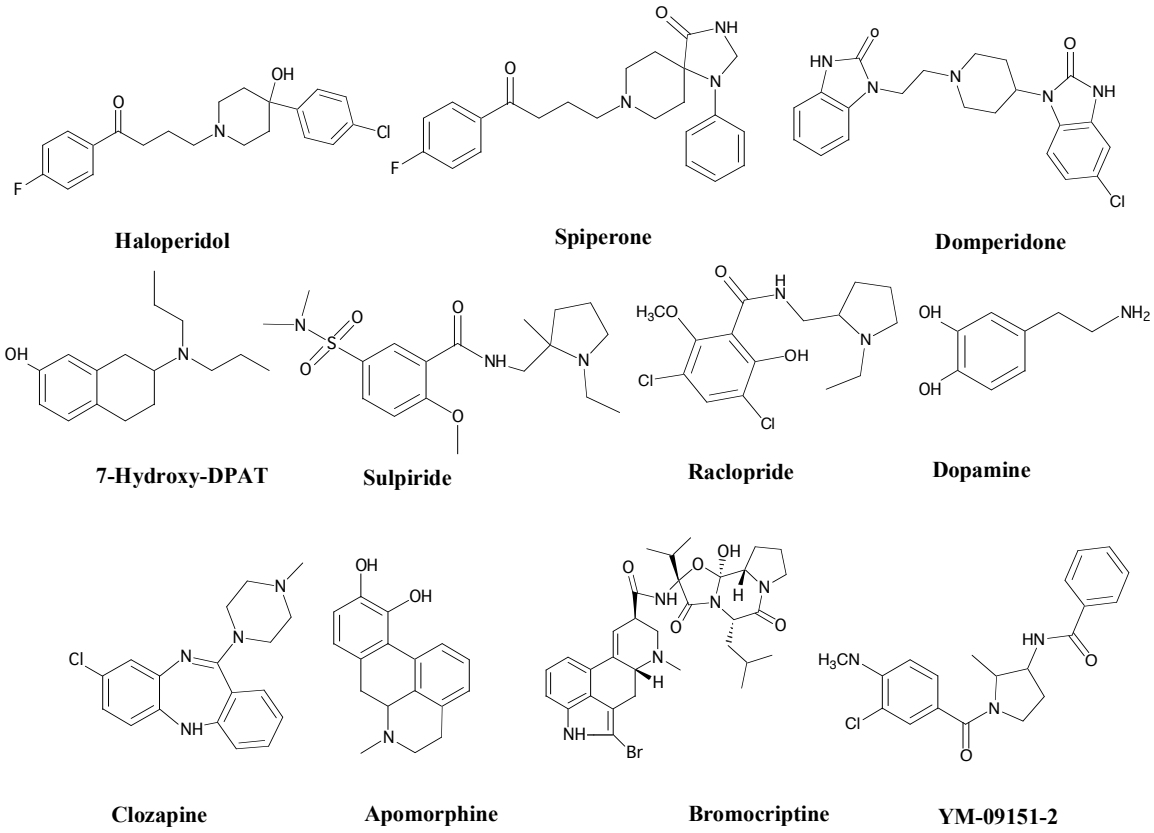

Figure 2-4. The two-dimensional figures of 11 agonists and antagonists used in this study.

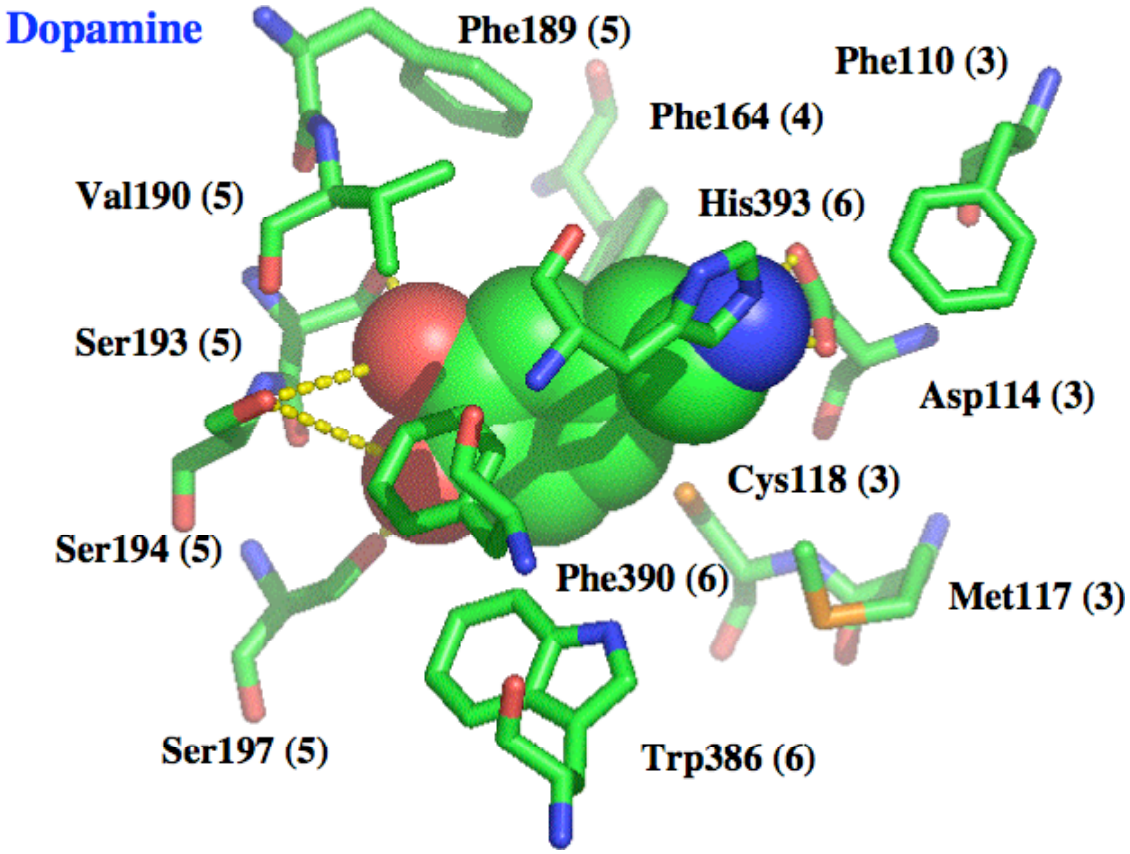

Figure 2-5. The predicted binding site of dopamine, the endogenous ligand of the dopamine receptors. We find that the binding site is located between TM3, 4, 5, and 6.

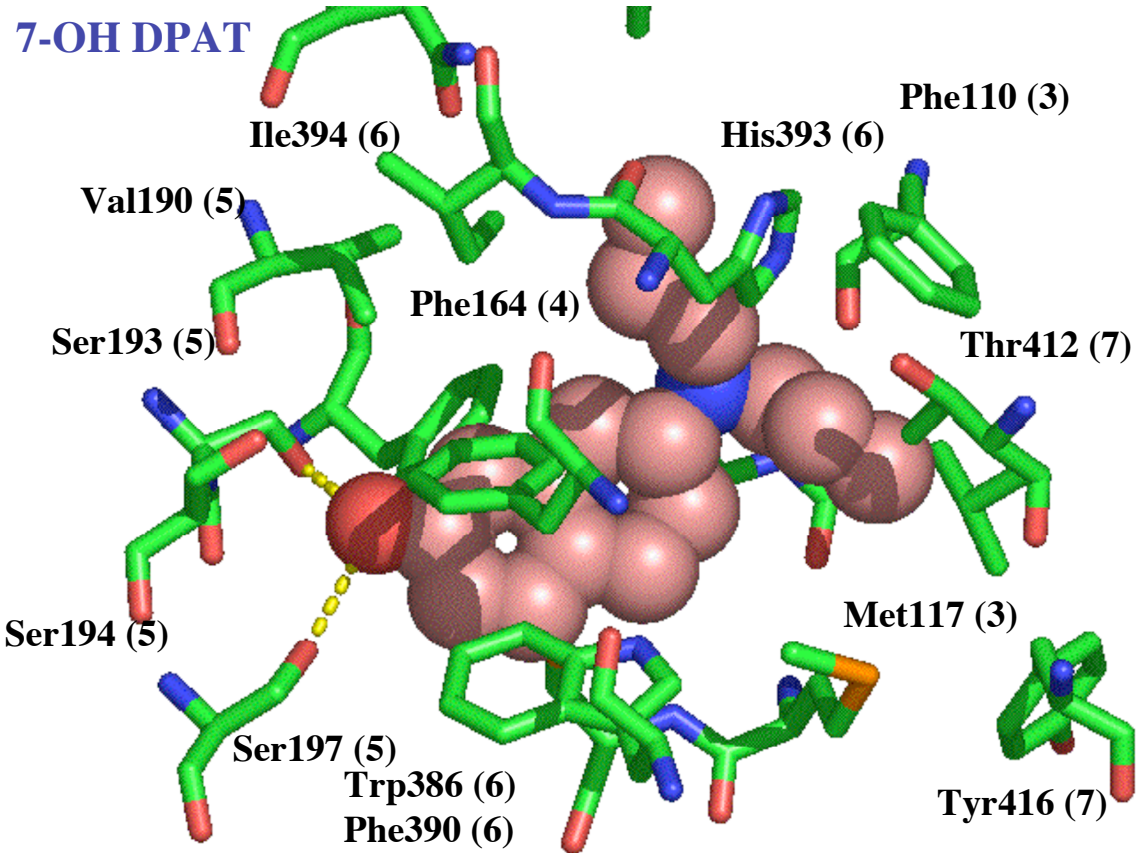

Figure 2-6. The predicted binding site of 7-OH DPAT to the receptor; again the binding site is located between TM3, 5, and 6 with minor contacts to TM4 and 7.

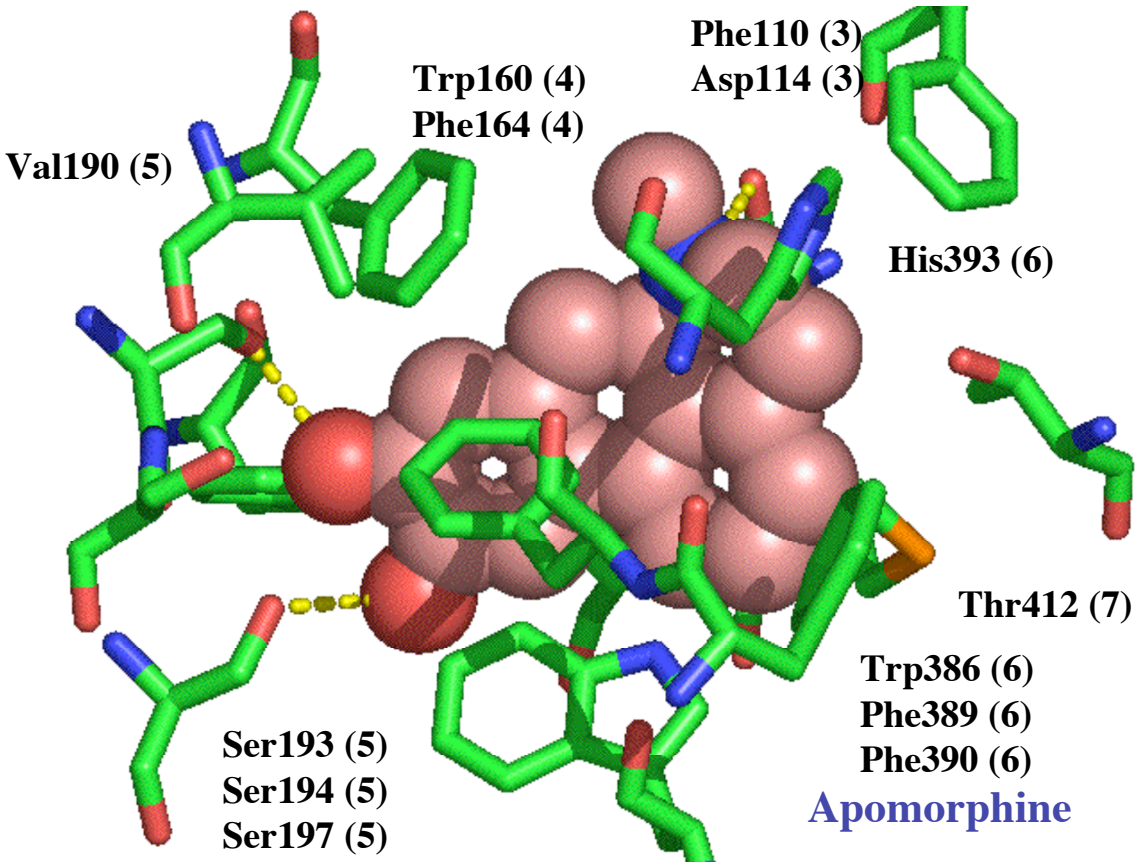

Figure 2-7. The predicted binding site of apomorphine to the receptor.

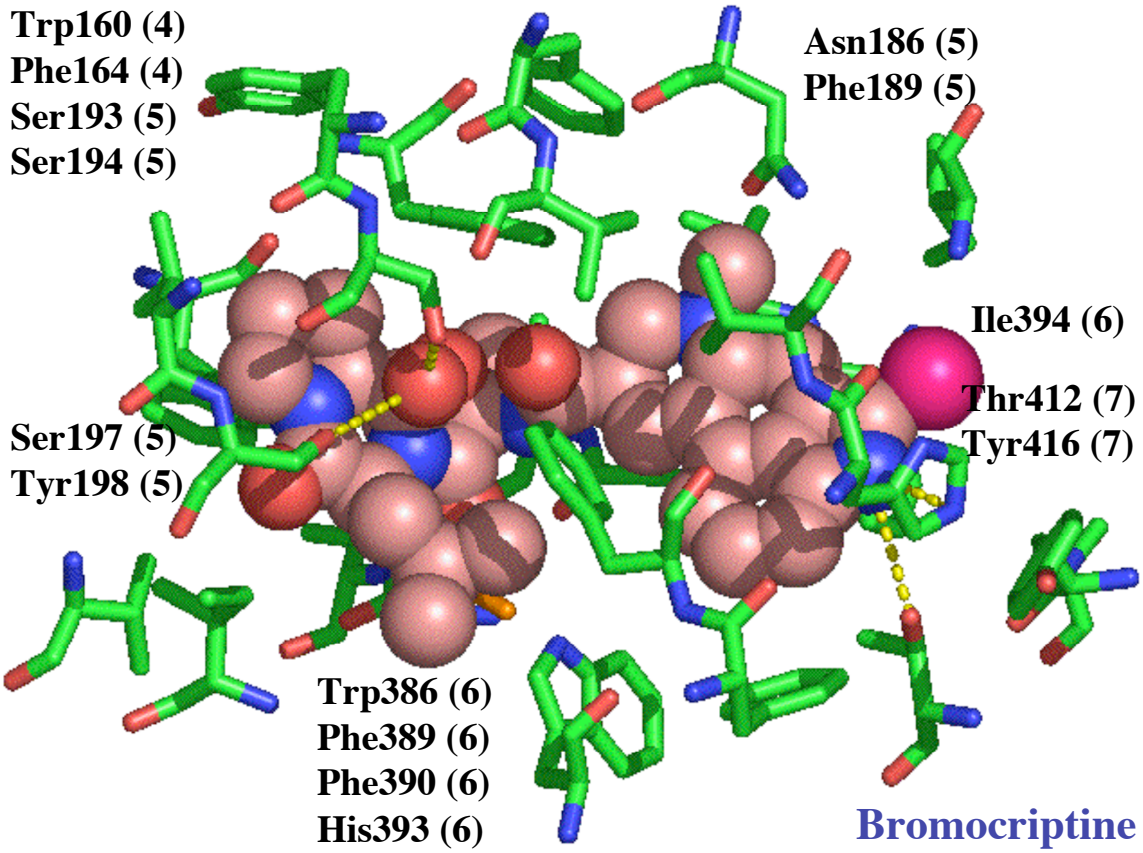

Figure 2-8. The predicted binding site of bromocriptine to the receptor.

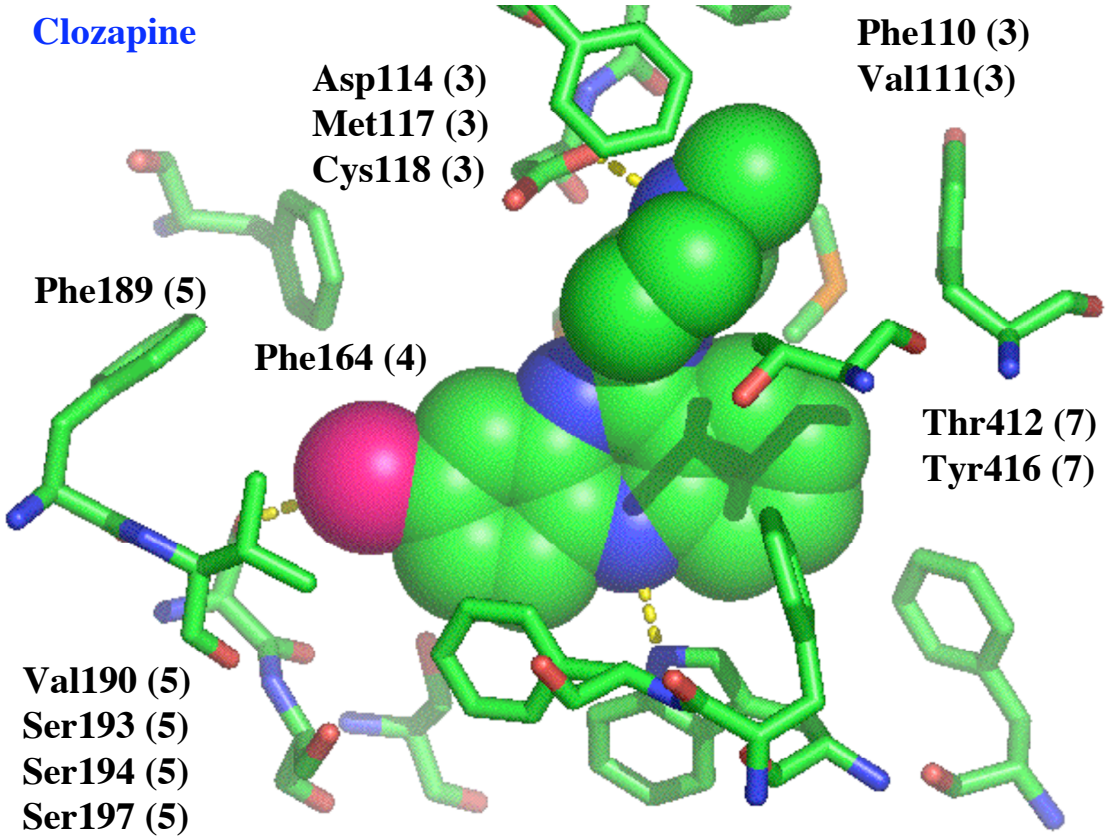

Figure 2-9. The predicted binding site of clozapine to the human D2 dopamine receptor. We classify clozapine as a class I antagonist, since it binds in the agonist binding site located between TM3, 4, 5, and 6.

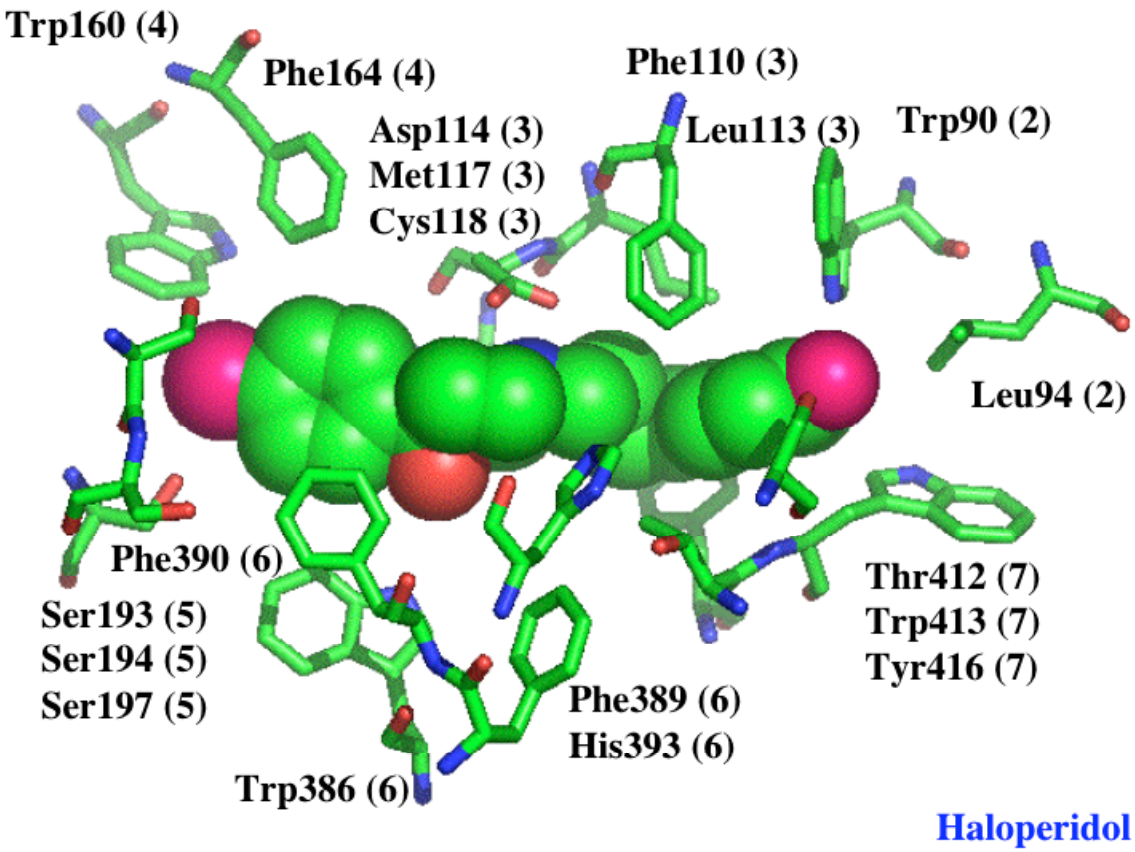

Figure 2-10. The predicted binding site of haloperidol to the receptor, which we classify as a class II antagonists. This binding site is located between TM2, 3, 4, 5, 6, and 7, quite different from that of class I antagonists (TM3, 4, 5, 6).

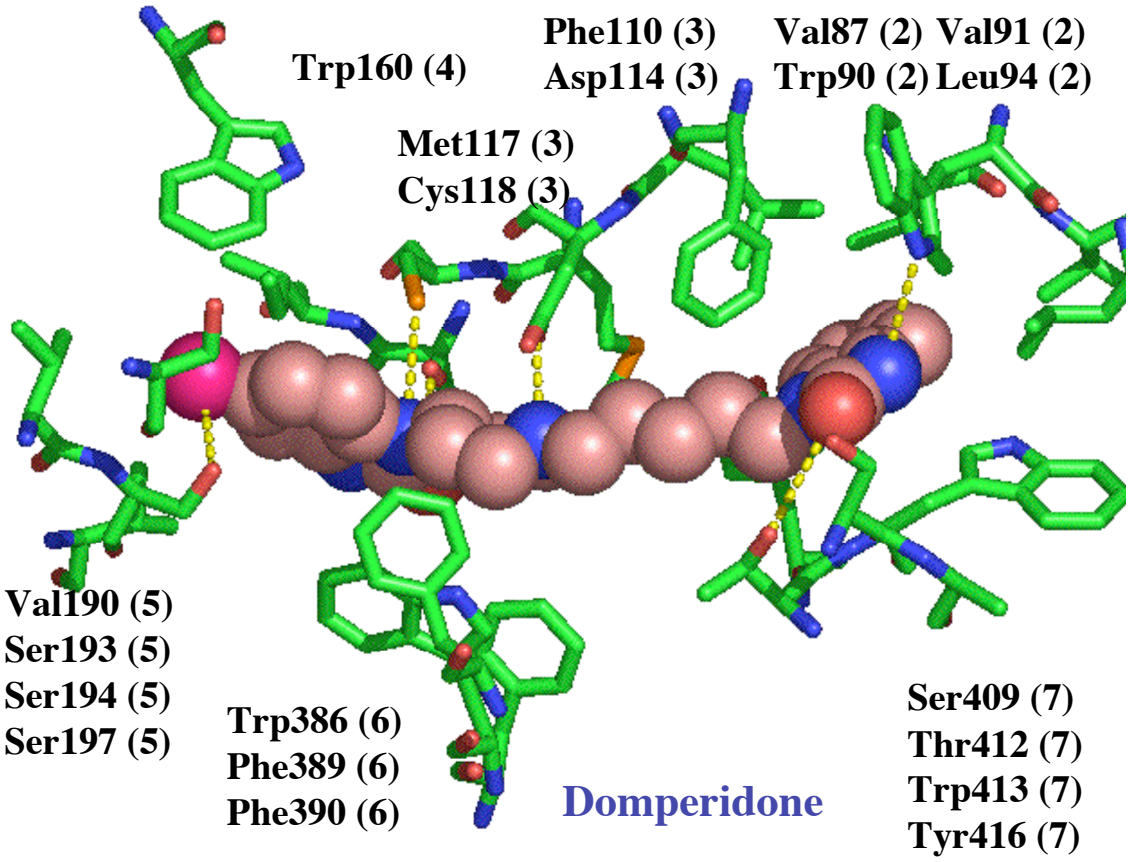

Figure 2-11. The predicted binding site of domperidone to the receptor.

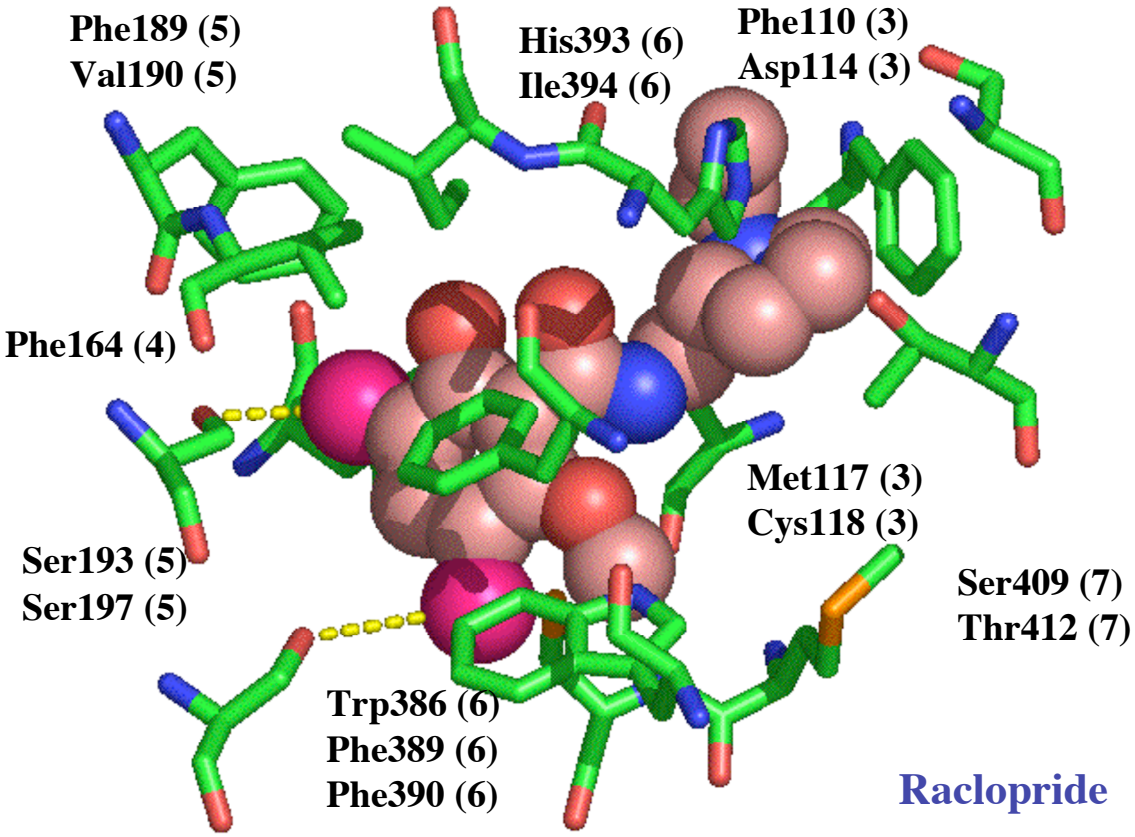

Figure 2-12. The predicted binding site of raclopride to the receptor.

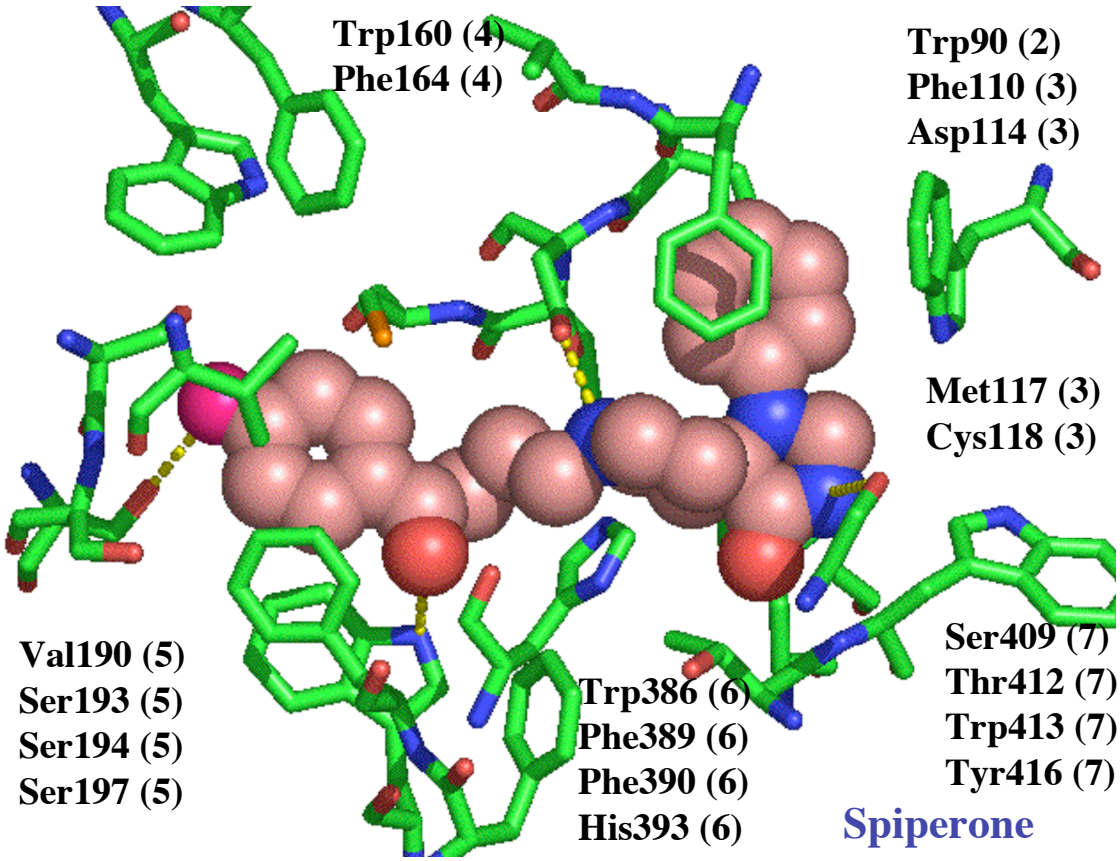

Figure 2-13. The predicted binding site of spiperone to the receptor.

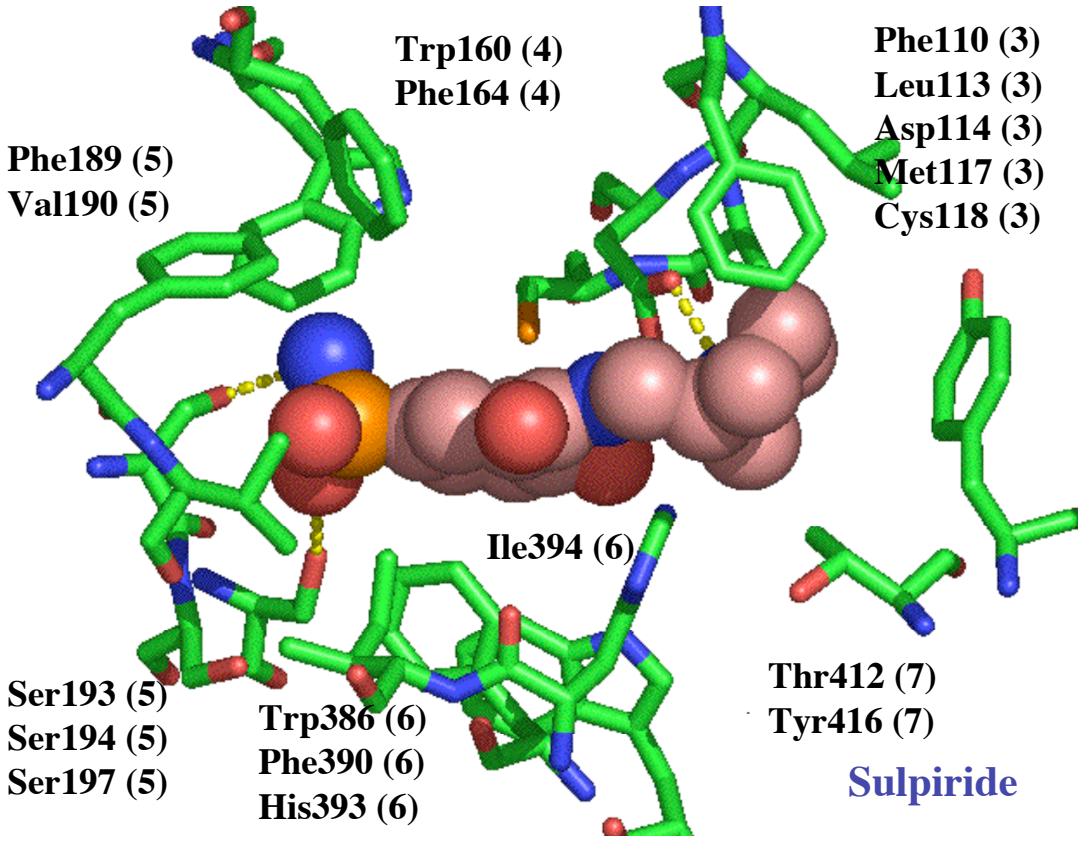

Figure 2-14. The predicted binding site of sulpiride to the receptor.

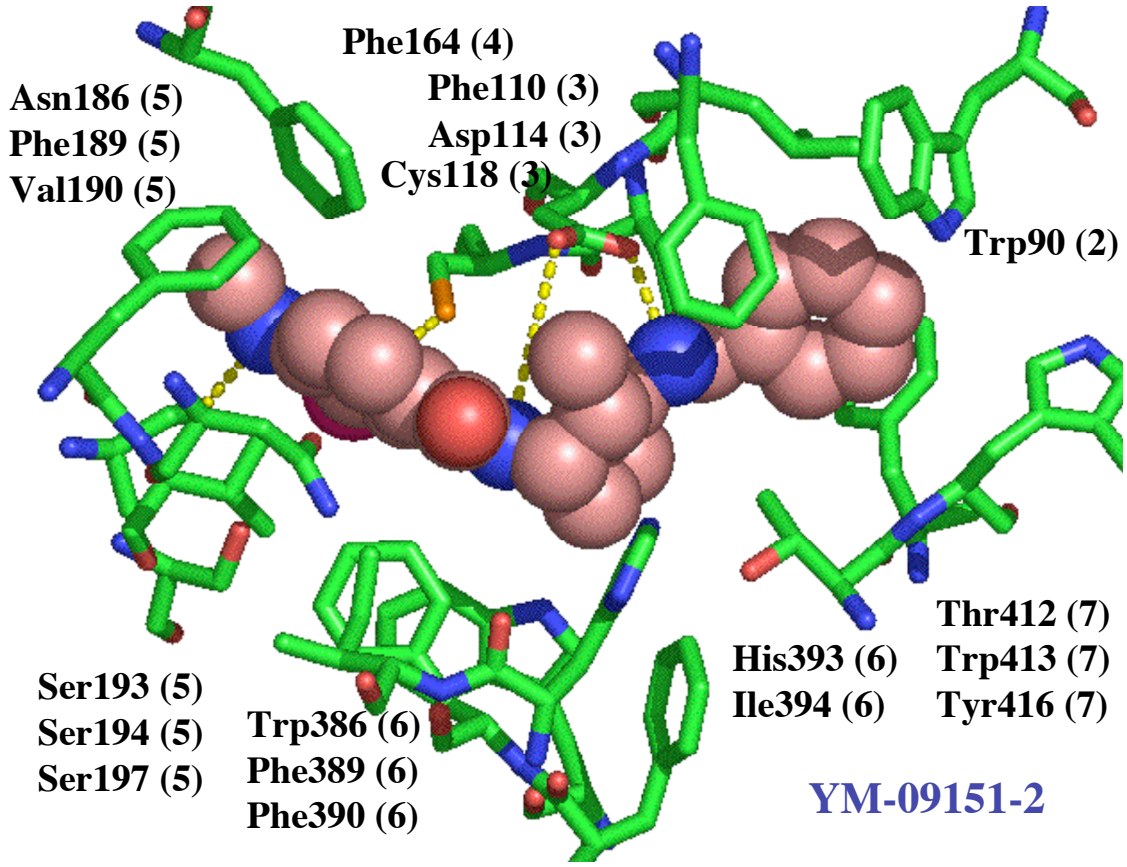

Figure 2-15. The predicted binding site of YM-09151-2 to the receptor.

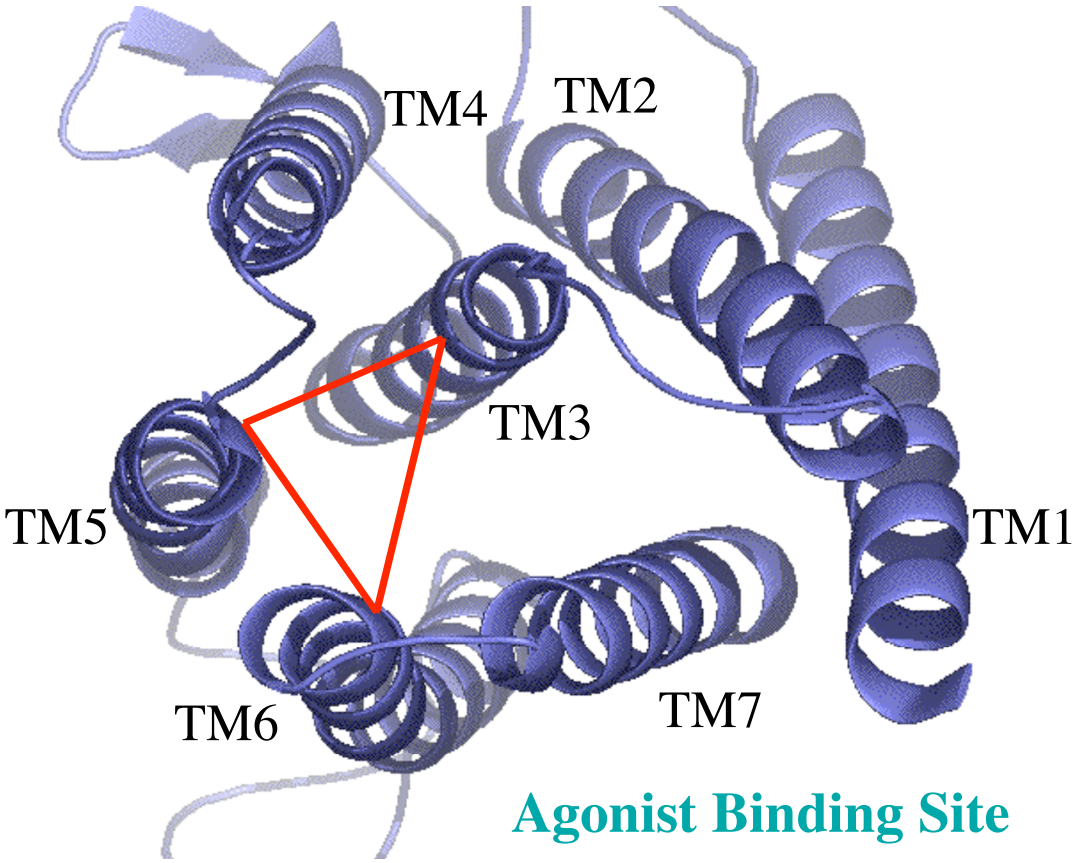

Figure 2-16. The predicted agonist (i.e., dopamine) binding site of the receptor.

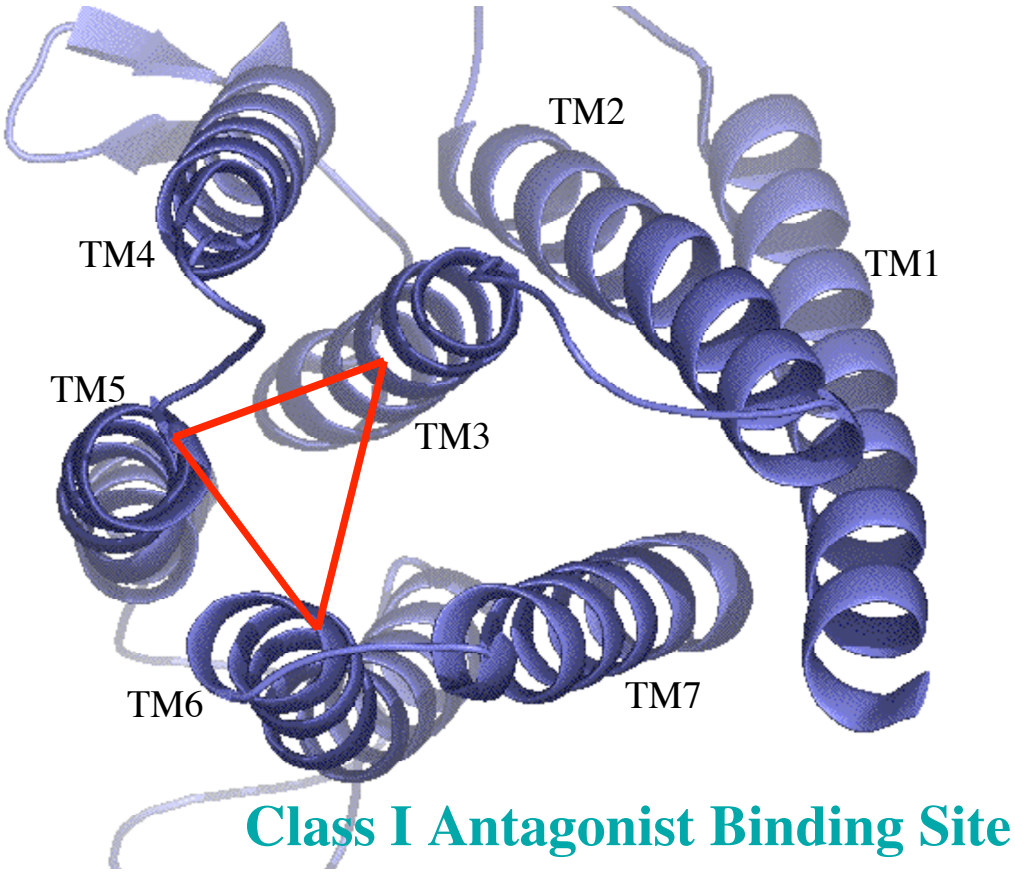

Figure 2-17. The predicted class I antagonist (clozapine-like) binding site of the receptor.

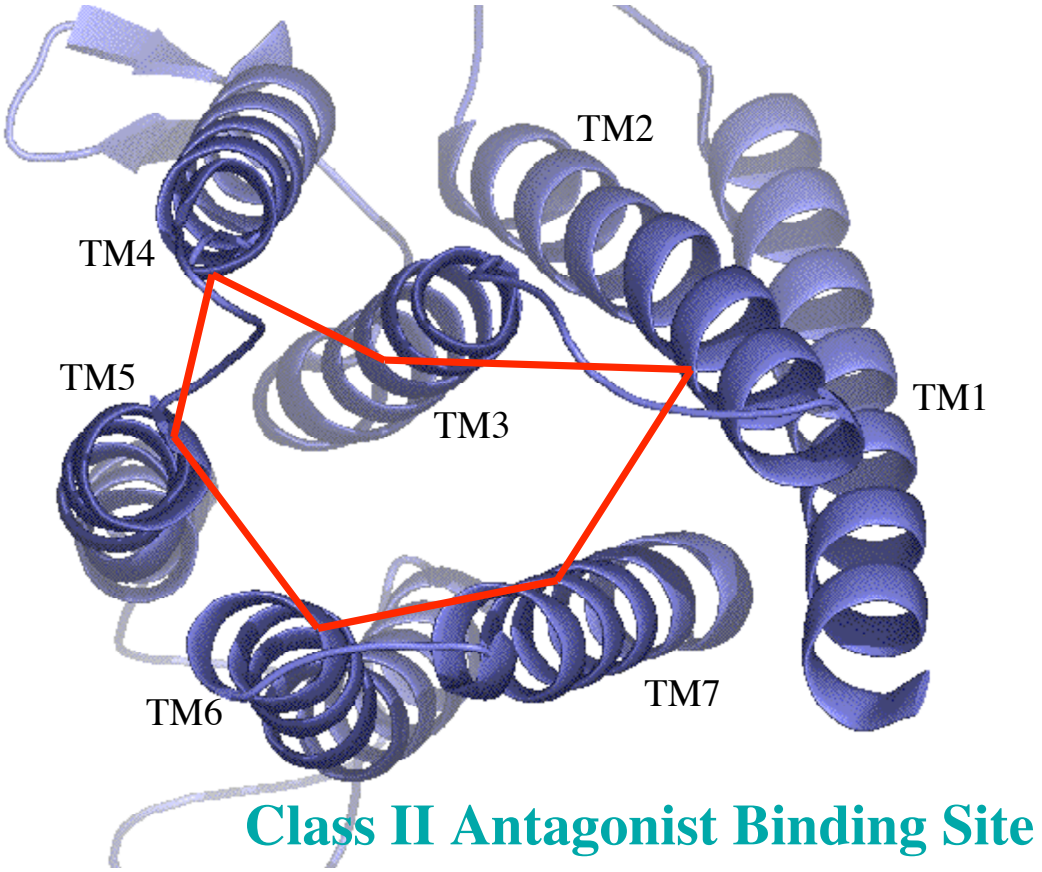

Figure 2-18. The predicted class II antagonist (haloperidol-like) binding site of the receptor.

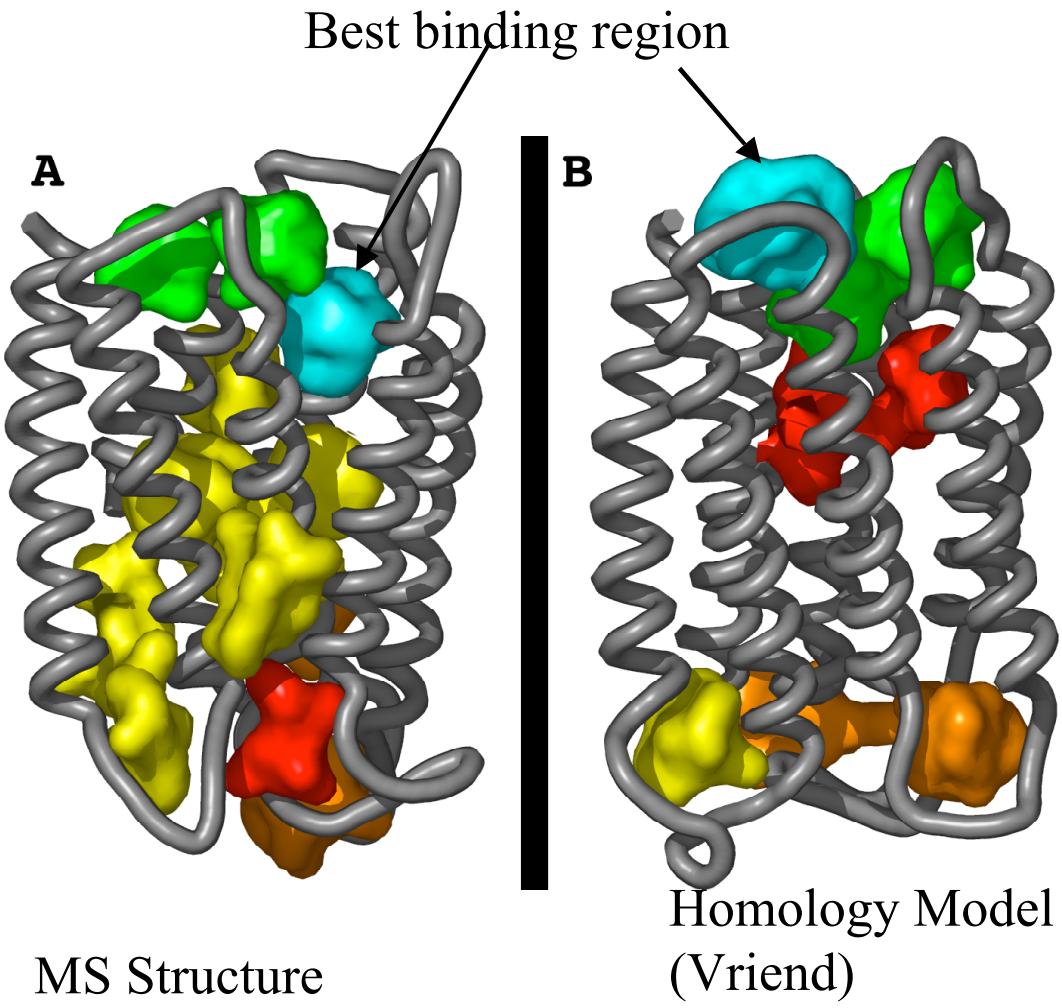

Figure 2-19. Side by side comparison of scanning results for D2DR Membstruk (A) and homology model (B). Sites are color-coded based on the energy ranges obtained during scanning for ligands in them (-100 – 100=green, 0 - 350=yellow, 200 - 500=orange, 100-900=red). The best identified site for dopamine in each case is colored cyan, and falls into the 'green' energy range. Note that in the Membstruk model a cavity within the bundle is predicted as the dopamine-binding site, but this same site in the homology model is quite unfavorable. Proteins are displayed extracellular side up, with helix I on the far left.

Dopamine

Residue #	Original Res.	Change in E	Experimental Data:
114	D	117.7	
393	H	7.4	
110	F	7	F110L
118	C	5.5	C118K
412	T	3	
197	S	2.3	S197A
390	F	2.1	F390A
194	S	0.6	
189	F	0.6	F189Y/C
192	Y	0.4	
394	I	0.2	
389	F	0.2	F389A
196	V	0.1	
386	W	0.1	W386C
115	V	0.1	V115C
190	V	0.1	
198	F	0	F198C
195	I	-0.1	
191	V	-0.1	
193	S	-0.4	
164	F	-3.3	
111	V	-3.5	
160	W	-4.9	W160C
117	M	-12.8	

Table 2-3. The change in relative binding energies for several alanine mutants for dopamine relative to the wild type binding energies.

Haloperidol

Residue #	Original Res.	Change in E
114	D	123.2
164	F	10.1
393	H	8.6
90	W	8.6
118	C	4.7
390	F	3.6
409	S	2.3
193	S	2.1
412	T	1.8
113	L	1.7
416	Y	1.7
197	S	1.2
386	W	1.2
111	V	0.7
42	T	0.6
190	V	0.6
413	W	0.5
194	S	0.5
189	F	0.4
94	L	0.4
394	I	0.2
110	F	0.2
196	V	0.1
417	V	0.1
87	V	0.1
408	Y	0
91	V	0
410	A	-0.2
122	I	-2.3
115	V	-2.4
117	M	-3.5
160	W	-14.5

Table 2-4. The change in relative binding energies of several alanine mutants for haldol relative to the wild type binding energies.

Ligand	Binding Energy (kcal/mol)
7-OH DPAT (++)	24.0
Apomorphine (+++)	11.7
Bromocriptine (+++)	12.8
Clozapine (+)	-10.6
Domperidone (ND)	44.7
Dopamine (+)	-8.9
Haloperidol (++++)	26.5
Raclopride (+++)	-22.4
Spiperone (++++)	9.7
(-)-Sulpiride (++)	29.7

Table 3-5. The binding energy of 10 pharmaceutical ligands to the D_2 DR(Vriend) homology model. The more negative the binding energy the better binder is the ligand. Plus signs (+) in the brackets correspond to the following legend: ++++, Inhibition constant (*K*i) <0.5 nM; +++, 0.5 nM < *K*i < 2- 5 nM; ++, 5 nM < *K*i < 50 nM; +, 50 nM < *K*i < 500 nM; +/-, 500 nM < *K*i < 5 mM; 7-OH-DPAT, 7- hydroxy-dipropylaminotetralin; ND, Not Determined. Data from Missale *et al.,* 1998.

MDPLNLSWYDDDLERQNWSRPFNGSDGKADRPHYNYYATLLTLLIAVIVFGNV
LVCMAVSREKALQTTTNYLIVSLAVADLLVATLVMPWVVYLEVVGEWKFSRIH
CDIFVTLDVMMCTASILNLCAISIDRYTAVAMPMLYNTRYSSKRRVTVMISIVWV
LSFTISCPLLFGLNNADQNECIIANPAFVVYSSIVSFYVPFIVTLLVYIKIYIVLRRRR
KRVNTKRSSRAFRAHLRAPLKGNCTHPEDMKLCTVIMKSNGSFPVNRRRVEAAR
RAQELEMEMLSSTSPPERTRYSPIPPSHHQLTLPDPSHHGLHSTPDSPAKPEKNGH
AKDHPKIAKIFEIQTMPNGKTRTSLKTMSRRKLSQQKEKKATQMLAIVLGVFIIC
WLPFFITHILNIHCDCNIPPVLYSAFTWLGYVNSAVNPIIYTTFNIEFRKAFLKILHC

Scheme 2-1. The predicted transmembrane regions of the human D_2 dopamine receptor. Residues in red correspond to the transmembrane helices, while the residues in black represent the N & C termini, and the loops connecting the transmembrane domains.