Chapter III: First Principles Predictions and Validation of the Binding of Pharmaceutical Antagonists to Human D2 Dopamine Receptor

Modified from the manuscript of paper in preparation for submission to the Journal of Medicinal Chemistry by M. Yashar S. Kalani, Nagarajan Vaidehi, Peter L. Freddolino, Maziyar A. Kalani, and William A. Goddard III^{*} (Caltech Team) and James Hendrix, Felix Sheinerman, Isabelle Morize, and Abdelazize Laoui (Aventis Pharma Team).

Chapter III: First Principles Predictions and Validation of the Binding of Pharmaceutical Antagonists to Human D₂ Dopamine Receptor

Abstract:

Designing subtype specific drugs for G-protein coupled receptors (GPCRs) is a challenge since there are often many receptors for the same endogenous ligand (e.g., five dopamine receptors) and there are no three-dimensional structures available for any human GPCR. Recently we used the MembStruk first principles computational method to predict the 3D structure of human D_2 dopamine receptor (h D_2DR) and used the HierDock first principles ligand docking method to predict the binding site and binding energy for dopamine and ten agonists and antagonists. We found that the predicted binding sites of agonists and antagonists agree well with the previously reported mutation experiments. Herein we now report the results of a blind test in which Caltech predicted the binding site and binding energies for nine antagonists to hD₂DR for which only Aventis had access to experimental data. After Caltech disclosed the calculated binding site and binding energies of the nine antagonists, the Aventis team provided the experimental inhibition constants for comparison, as reported here. The good correlation of the calculated binding energies to the measured inhibition constants for this blind test further validates the accuracy of the 3D structure of hD₂DR predicted using the first principles MembStruk method and of the accuracy of the HierDock first principles method for predicting binding site and energy. This suggests that these first principles methods

should be valuable for designing antagonists and agonists for other G-protein coupled receptors.

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

1.0 INTRODUCTION

Dopamine receptor antagonists have been developed to block hallucinations and delusions that occur in schizophrenic patients, whereas dopamine receptor agonists are effective in alleviating the hypokinesia of Parkinson's disease. However, blockade of dopamine receptors can induce side effects similar to those resulting from dopamine depletion, and high doses of dopamine agonists can cause psychoses. The therapies of disorders resulting from dopamine imbalances are thus associated with severe side effects. One of the challenges in designing pharmaceutical agents has been to discover selective dopaminergic drugs devoid of adverse effects. Thus it is important to understand how the agonists of one subtype of dopamine receptor affect the other subtypes. These predictions are tedious with just the sequences of receptors or by even trial and error methods, since the sequences are very similar. Hence predicting the three dimensional structure and function of dopamine receptors and the binding sites of dopamine is of utmost value in designing subtype specific agonists and antagonists.

Dopamine receptors belong to the large family of seven helical transmembrane (TM) proteins called the G-protein coupled receptors (GPCRs). GPCRs have a seven helical transmembrane (TM) motif connected by intracellular (IC) and extracellular (EC) loops. Homology structure prediction method based on the crystal structure of bovine rhodopsin in combination with experimental mutation results on dopamine receptors have been used to rationalize experimental ligand binding measurements (for some examples see references 1-4). These methods very often lack the predictive value and are not reliable for GPCRs with very low sequence similarity to bovine rhodopsin. We have developed first principles methods to predict structure and function for GPCRs⁵⁻⁷. These methods have been validated for predictions in bovine rhodopsin⁷, $\beta 2$ adrenergic

receptor⁸ and olfactory receptors^{9,10}. In a previous publication we have used these methods to predict the three dimensional structure of the human D2 dopamine receptor $(hD_2DR)^{11}$. In the same publication we have validated the structure by predicting the binding site and binding energies of 10 agonists and antagonists to hD_2DR .

In this publication we report the results of a *blind* prediction test by the Caltech team on a set of nine compounds for which the Aventis pharma team had data on the relative binding to human D2. The task was to predict the binding site and to calculate the binding energies for nine antagonists using the three-dimensional structure of hD₂DR predicted previously using MembStruk. The Aventis team had carried out measurements of inhibition constants for these nine antagonists, but they provided the Caltech team with just the two dimensional drawings of the nine compounds, as shown in Figure 3-1. The Caltech team used HierDock to predict the binding site and the binding energy for each of the nine Aventis antagonists. The binding sites for all nine antagonists are predicted to be located between TM helices 2, 3, 4, 5, 6, and 7. The binding energies calculated for the nine compounds were provided by Caltech to Aventis and are compared here. The sequence of binding constants is in the same order and correlation of the predicted binding energies to the measured inhibition constants correlates well ($R^2 = 0.75$). This excellent agreement with the experimental results further validates the predicted threedimensional structure of hD_2DR . This demonstrates the utility of the first principles methods for predicting the structures of and binding to other GPCRs for which there is little or no experimental information.

The details of the HierDock computational method as applied to predicting the binding site of the nine Aventis compounds, and the experimental methods used for

measuring the inhibition constants, are summarized in Section 2. Section 3 summarizes the results and discussion followed by the conclusions in Section 4.

2.0 METHODS

All results in this paper use the Dreiding Force Field¹² for the protein and ligands with CHARMM22¹³ charges for the protein and Gasteiger¹⁴ charges for the ligands.

- 2.1 **Prediction of the 3D Structure of hD_2DR:** The three dimensional structure of hD_2DR predicted using MembStruk¹¹ was used in this work. The details of the method are given elsewhere⁷.
- 2.2 Prediction of the binding sites and binding energies of the nine Aventis antagonists in hD2DR: The binding sites and binding energies of the nine Aventis compounds in the hD₂DR structure were predicted using HierDock. The previous studies on D2 scanned the entire hD₂DR structure to determine the putative binding sites for 11 well-studied agonists and antagonists for hD₂DR. The predicted binding site for known dopamine agonists is located between TM helices 3, 4, 5, and 6, while the best antagonists bind to a site involving TM helices 2, 3, 4, 6, and 7 with minimal contacts to TM5. In this paper we applied the HierDock2.0 protocol⁶ to predict the binding region for all nine Aventis antagonists to hD₂DR, but concentrating on the region involving TM helices 2, 3, 4, 6, and 7 that is recognized by the Class II Dopamine antagonists.

2.2.1 The HierDock Protocol: The HierDock ligand docking protocol follows a hierarchical strategy for examining ligand binding conformations and calculating their binding energies. The steps of HierDock 2.0^6 used here were as follows:

- First we used a coarse grain docking procedure to generate a set of conformations for ligand binding in the receptor. Here we use Dock 4.0¹⁵ to generate and score 500 ligand conformations, of which 10% (50) were selected for further analysis (using a buried surface area cutoff of 85% and using energy scoring from Dock4.0 with flexible ligand docking using torsion drive and allowing four bumps).
- The 50 best conformations selected for each ligand from step 1 were minimized (all-atom conjugate gradients) while keeping the protein fixed. For each of these 50 minimized structures, the solvation energy was calculated using the Analytical Volume Generalized Born (AVGB) continuum solvation method¹⁶. The binding energies (BE) were calculated using

$$BE = PE (ligand in protein) - PE (ligand in solvent)$$
(1)

where PE is the potential energy. Five structures based both on binding energies and buried surface areas were selected from the 50 structures for the next step.

- 3. For the best 5 ligand-protein structures, the structure of the receptor/ligand complex was optimized (conjugate gradient minimization, allowing the structure of the protein to accommodate the ligand). Using these optimized structures, the binding energy was calculated 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 was calculated using the DREIDING FF and the AVGB continuum solvation method¹⁶.
- 4. From the five structures from step 3, we selected the one with the maximum number of hydrogen bonds between ligand and protein. We then used the SCREAM side

chain replacement program to reassign all side chain conformations for all residues within 4Å of the ligand [SCREAM 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, and the best is taken as the final bound state.

3.0 RESULTS AND DISCUSSION

Experimental SAR data and interpretation of the ligand-binding site:

The aryl piperazine and aryl piperidine groups are well known for their ability to bind at the dopamine D2 receptor. Therefore, it is not surprising to find that most of the compounds exemplified in this study are potent D2 ligands. It was surprising, however, that three compounds (1, 3, and 4) are weak at the D2 receptor with $>1 \mu$ M activity. We attributed this lack of activity to the presence of the 6-trifluoro methyl group on the benzo[b]thiophene heterocycle that imparts steric bulk that appears to interfere with critical interactions needed for binding. When the CF₃ group of the benzo[b]thiophene is replaced by a smaller fluoride, as with compounds 2 and 9, the D2 activity is significantly improved. The groups appended to the amide or urea also seem to play an important role in D2 binding. While all the groups explored were relatively hydrophobic, the compounds with aliphatic groups on the amide, such as 5 and 6, exhibited only modest D2 affinity. In contrast, where the amide was bound to an aromatic group, such as the biphenyl (7), the piperzinyl pyrimidine (8) and the chloropyridine (9), the D2 affinity was very high. The exception was seen where compounds have a straight chain, four-carbon linker between the piperazine and amide and only modest D2 affinity was observed. It may be that the straight chain linker allows the molecule to extend outside the D2 pocket reducing affinity. The other potent D2 ligands all have constrained linkers that likely make these compounds more compact and a better fit into the receptor.

Theoretical predictions and interpretations of the ligand-binding site:

The predicted three-dimensional structure of hD_2DR with the binding site of the best inhibitor is shown in **Figure 3-2**.

3.1 Predicted Binding Energies of the 9 Aventis compounds

Table 3-1 shows the calculated binding energies of the 9 Aventis compounds (hereafter denoted by a number 1-9 in bold face) along with the experimental inhibition constants. In both cases the order is 7 (best), **8**, **9**, **2**, **6**, **5**, **1**, **4**, **3** (worst). Since these predictions were made prior to any knowledge of the experimental results, this is a dramatic validation of the accuracy of the tertiary structure of the D2 receptor and of the HierDock procedure. Although the sequence is the same, the groupings are somewhat different. The calculated binding energies show 7 as the best binder followed by compounds **8**, **9**, and **2**. Compounds **6**, **5**, **1**, and **4** are medium binders while **3** is the worst binder. According to experiment the good binders are **7**, **8**, and **9** while **2**, **5**, and **6** are medium binders and compounds **1**, **4**, and **3** are weak binders. **Figure 3-3** shows the plot of calculated binding energies versus the natural logarithm of ki values. The correlation is good with a correlation factor of 0.75.

We should point out that compounds 3 and 4 represent special cases, since they lead to diasteromers. As indicated in Figure 3 and table 2 we have done all four cases

with results that depend significantly on the stereochemistry. In the data provided to Caltech by Aventis, the stereochemistry second stereocenter was not been specified. Assuming that the compounds used in this experiment were racemic mixtures at this second center led to an R^2 of 92% in the correlation of theory to experiment, but using the assignment provided by Aventis, the correlation drops to 75%. Note that our predictions suggest that other choices for the stereochemistry of 3 and 4 would move them from weak to medium binders.

3.2 Predicted binding site of the nine Aventis compounds:

Next we will analyze the predicted binding site to determine how each residue in the binding pocket contributes to the binding energy. Here we will analyze the details of the predicted binding site for a strong binder 7, a medium affinity binder 6, and a weak binder 3.

3.2.1 Binding site of the best binder, compound 7: **Figure 3-2** shows the predicted location of the best binding ligand 7 (both in theory and experiment) in the hD₂DR structure. The binding site for this ligand is located between TM helices 3, 4, 5, 6, and 7 with some contributions from residues in TM1 and 2. Figure 5 shows the residues within 5Å of the binding site of 7. We find that this ligand makes contact with residues in the aromatic microdomain located between TM2 and TM7 and also in the aromatic microdomain located between TM4 and TM6. Indeed this binding site is similar to that for such class II antagonists as haloperidol and domperidone¹¹. The residues shown in **Figure 3-4** are color coded to indicate their interaction energies with the ligand. In discussing the residues we will often append in parenthesis the number of the TM helix to which a residue belongs.

The amino group of the ligand 7 forms a salt bridge (2.8 Å) to Asp114 on TM3.

The F atom on the aromatic moiety interacts with Ser193 (3.6 Å) but the strength of the F---H hydrogen bond is lower than the O-----H hydrogen bond.

This antagonist makes no tight contact with Ser194 or Ser197 on TM5. This agrees with the binding of antagonists such as clozapine, haloperidol, and domperidone and it contrasts with the binding sites for agonists such as dopamine, bromocriptine, and apomorphine which make contact to both Ser193 and Ser197 (and sometimes Ser194) on TM5^{11,17}.

The halogenated aromatic ring also interacts well with Trp386 (6) (3.0 Å) and Cys118 (3) (3.1 Å) (these residues are also important for haloperidol and domperidone).

Trp90 (2) has a 2.8Å contact with the amide carbonyl of the ligand (this residue is also important for haloperidol and domperidone).

The non-derivatized aromatic rings of this ligand are strongly stabilized by the second aromatic microdomain composed of Trp90 (2), Phe110 (3), Trp413 (7), and Tyr416 (7).

Other residues in the cavity providing a mainly a hydrophobic pocket for the ligand with each one making a small contribution to the ligand interaction energy include Asn35 (1), Ala38 (1), Thr39 (1), Thr42 (1), Leu43 (1), Val87 (2), Trp90 (2), Val91 (2), Leu94 (2), Glu95 (2), Phe110 (3), Leu113 (3), Asp114 (3), Met117 (3), Cys118 (3), Phe164 (4), Phe189 (5), Val190 (5), Ser193 (5), Ser194 (5), Ser197 (5), Trp386 (6), Phe389 (6), Phe390 (6), His393 (6), Ile394 (6), Ser409 (7), Thr412 (7), Trp413 (7), Leu414 (7), Tyr416 (7), and Val417 (7).

3.2.2 Predicted binding site of medium affinity binder, compound 6: The calculated binding energies for compounds **1**, **2**, **5** and **6** indicate medium affinity binders to hD_2DR with predicted binding energies of 52-59 % that of **7**. Experimentally the medium affinity binders are **6**, **2**, and **5**. We will examine ligand **6**.

The residues within 5Å of the ligand **6** are shown in **Figure 3-5**. Compound **6** is a medium affinity binder with a binding affinity 59% of **7**. The important elements of the bonding here are

The binding site of **6** contains a good salt bridge with Asp114 (3) (2.9 Å).

Most strong binding ligands have a functional group that interacts favorably with serines on TM5, but compound **6** has a methyl derivative positioned towards TM5 so that it does not form hydrogen bonds with any of the serines in TM5.

The amide functionality of 6 is in close contact with the Tyr416 (7).

The aliphatic portions of the ligand are stabilized by the network of aromatic residues in TM2, 4, 6, and 7.

The following residues are present in the binding cavity: Thr42 (1), Val79 (2), Ala80 (2), Leu82 (2), Val83 (2), Leu86 (2), Trp90 (2), Phe110 (3), Leu113 (3), Asp114 (3), Met117 (3), Cys118 (3), Ser121 (3), Phe164 (4), Val190 (5), Ser193 (5), Ser194 (5), Ser197 (5), Trp386 (6), Phe390 (6), His393 (6), Ser409 (7), Thr412 (7), Trp413 (7), Tyr416 (7), and Val417 (7).

Compound **1** is predicted to have a medium binding affinity, while experimentally it is a weak binder. We believe that the problem in compound **1** is an inadequate treatment of amine contacts in the Dreiding forcefield: Our calculated structure has Asp114 of hD₂DR forming a bi-dentate contact with the tertiary amine of the ligand, whereas most other structures form a mono-dentate contact. Quantum mechanics calculations on this system show that a mono-dentate contact is more favorable, with nothing gained from the bi-dentate interaction. However in the forcefield the interaction is dominated by the Coulombic energy, which for the bi-dentate contact is twice as strong as for the mono-dentate ligand. This suggests training the force field to properly treat such interaction.

3.2.3 Predicted Binding site of the weak binder compound 3: The predicted binding site of compound **3** is located again between TM helices 3, 4, 5, and 2, 6, 7. The binding energy of **3** is only 25 % of **7**, making it significantly weaker binding than the other compounds, in good agreement with the experimental measurement as a weak binder. **Figure 3-6** shows the residues within 5Å of the binding site. The analysis of the bonding shows that Compound **3** is a class II antagonist:

The amino group forms a salt bridge to Asp114 of TM3.

The trifluoro derivatized ring is docked in the first aromatic microdomain located between TM4 and TM6. The trifluoro group interacts with the TM5 serines with one close contact to Ser193 at 2.7 Å. Antagonists usually interact well with only one serine in TM5. However the trifluoro group of this ligand leads to fair interactions with both Ser193 (3.1 Å) and Ser197 (4.2 Å) on TM5.

Trp160 (4), Phe164 (4), Trp486 (6), and Phe390 (6) make up the microdomain necessary for stabilizing the aromatic ring system.

The ring also has some interaction with Cys118 (3).

The other ring system of **3** interacts with the second aromatic microdomain composed of Trp90 (2), Trp413 (7) and Tyr416 (7).

The cyclopropyl portion of the ring is stabilized by Phe110 (3).

The remaining residues in the cavity provide a mostly hydrophobic pocket for the ligand. These residues include: Thr42 (1), Val79 (2), Ala80 (2), Leu82 (2), Val83 (2), Leu86 (2), Trp90 (2), Phe110 (3), Leu113 (3), Asp114 (3), Met117 (3), Cys118 (3), Trp160 (4), Phe164 (4), Phe189 (5), Val190 (5), Ser193 (5), Ser194 (5), Ser197 (5), Trp386 (6), Phe390 (6), His393 (6), Ser409 (7), Thr412 (7), Trp413 (7), Tyr416 (7).

There are no other specific contacts to the ligands.

3.2.5 Comparison of High, Medium, and Low affinity binding sites: All ligands in this study are predicted to bind in the same region, occupying the upper third of the transmembrane region and filling both the cavity between TMs 1-2-3-7 and the cavity between 3-4-5-6. This class II binding pattern is characteristic of such well-characterized dopaminergic antagonists as haloperidol, which occupies both the agonist binding site (3-4-5-6) and the smaller 1-2-3-7 cavity¹¹. All ligands in this study have the following characteristics:

They exhibit a contact between the ligand tertiary amine and Asp 114 (3),

They form contacts with serine S193 on TM5 and sometimes S194, but never S197,

They interact with one aromatic microdomain in the 3, 4, 5, 6 cavity and another aromatic microdomain in the 1, 2, 3, 7 cavity.

The binding strengths are determined primarily by the quality of a few specific interactions between protein and ligand. Thus weak binding compound **3** has no strong

specific contacts with the protein, medium binder compound **6** is strengthened by a very good hydrophobic contact with F110 and marginally better interactions with TMs 2 and 7, while strong binder compound **7** is further aided by a hydrogen bond with W90 (2) and better fit into the site, with less internal tension in the bound conformation.

4.0 Summary and Conclusions:

Starting with the first principles predicted structure of hD_2DR Caltech used the first principles HierDock method to predict the binding site and energy for nine proprietary compounds studied experimentally by Aventis. The predicted binding affinities of the compounds in this blind test are in the same sequence as the experimental inhibition constants with a correlation factor of 0.75 (working under the assumption that compounds **3** and **4** are diastereomers). These results combined with the previous studies of 11 agonists and antagonists (in which the binding energies and binding sties are consistent with experimental binding and mutation data) allow us to conclude that the predicted 3-D structure of hD_2DR is sufficiently accurate to predict function.

In addition, the detailed analysis of the predicted binding sites revealed a number of specific interactions that play an important role in determining the relative binding for these ligands. This understanding of the atomistic origins of the observed differences in binding should be useful in designing new ligands with improved efficacy and specificity.

Similar studies are being carried out on other GPCRs that often exhibit cross reactivity with antagonists to D2 (e.g., D₃DR and α 1A adrenergic receptor) are expected to provide hints on how to design subtype specific ligands. This study suggests that computational ligand screening on predicted tertiary structures of GPCRs can be

expected to identify the ligands having better binding affinities, prior to experimental screening. It is not established that the binding energies are sufficiently accurate to dispense with experiment but the predicted binding affinities accompanied with the analysis of ligand binding site might be expected to reduce substantially the number of experiments, greatly reducing the need for cell culturing and experimental synthesis. The detailed picture of receptor-ligand interactions derived from the computational studies should be useful in modify already promising compounds to optimize performance.

5.0 References:

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

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

3. 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).

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

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

6. 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**).

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

8. 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).

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

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

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

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

13. 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**).

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

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

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

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



Figure 3-1: The nine proprietary Aventis compounds. Compounds 7 and 9 have been deleted due to pending patents.



Figure 3-2: The predicted binding site for compound **8** (the best binder) to hD_2DR .



Figure 3-3: A plot of the calculated binding energy versus the experimentally determined inhibition constant for the 9 proprietary Aventis compounds.



Figure 3-4: The predicted binding site for compound 7 (a good binder) to hD₂DR.



Figure 3-5: The predicted binding site for compound 6 (a medium binder) to hD_2DR .



Figure 3-6: The predicted binding site for compound 3 (a poor binder) to hD₂DR.

Ligands	Calc.BE (kcal/mol)	Experimental ki (nM)
cpd1	37.47	1460
cpd2	47.42	51.3
cpd3	-8.44	10000
cpd4	4.74	10000
cpd5	37.60	163
cpd6	42.20	42.6
cpd7	71.62	0.05
cpd8	50.00	0.0857
cpd9	49.44	4.3

Table 3-1: The calculated binding energy of the 9 proprietary Aventis compounds and comparison to the experimentally determined inhibition constants (Ki).