The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists

M. Yashar S. Kalani, Nagarajan Vaidehi, Spencer E. Hall, Rene J. Trabanino, Peter L. Freddolino, Maziyar A. Kalani, Wely B. Floriano, Victor Wai Tak Kam, and William A. Goddard III*

Materials and Process Simulation Center, Beckman Institute, California Institute of Technology, Pasadena, CA 91125

Contributed by William A. Goddard III, January 6, 2004

Dopamine neurotransmitter and its receptors play a critical role in the cell signaling process responsible for information transfer in neurons functioning in the nervous system. Development of improved therapeutics for such disorders as Parkinson's disease and schizophrenia would be significantly enhanced with the availability of the 3D structure for the dopamine receptors and of the binding site for dopamine and other agonists and antagonists. We report here the 3D structure for the long isoform of the human D2 dopamine receptor, predicted from primary sequence using first-principles theoretical and computational techniques (i.e., we did not use bioinformatic or experimental 3D structural information in predicting structures). The predicted 3D structure is validated by comparison of the predicted binding site and the relative binding affinities of dopamine, three known dopamine antagonists (antipsychotic), and seven known antagonists (antipsychotic) in the D2 receptor to experimentally determined values. These structures correctly predict the critical residues for binding dopamine and several antagonists, identified by mutation studies, and give relative binding affinities that correlate well with experiments. The predicted binding site for dopamine and agonists is located between transmembrane (TM) helices 3, 4, 5, and 6, whereas the best antagonists bind to a site involving TM helices 2, 3, 4, 6, and 7 with minimal contacts to TM helix 5. We identify characteristic differences between the binding sites of agonists and antagonists.

With the implication of G protein-coupled receptor (GPCR) in many diseases (1, 2), the need to solve the high-resolution 3D structure of this class of integral membrane proteins to enable structure-based drug design is an important problem in structural biology. Despite the importance of solving the structure of the GPCRs, the only experimental 3D structure available for a GPCR is bovine rhodopsin. This lack of structures is because the GPCRs are bound to the membrane, making it difficult to express in sufficient quantities for crystallization.

To provide structural and ligand binding information on GPCRs, we have been developing first-principles computational techniques for predicting the 3D structure of GPCRs using only the amino acid sequence (MembStruk) and for predicting binding site and binding energy of various ligands to GPCRs (HierDock). Using these techniques, we have reported the structure of olfactory receptors (3, 4), bovine rhodopsin (4, 5), and other GPCRs (4).

Dopamine neurotransmitter plays a critical role in cellular signaling processes responsible for information transfer in neurons functioning in the nervous system (6, 7). Dopamine receptors (DR) belong to the superfamily of GPCRs, and to date there are five reported sequences for the human DR with multiple isoforms for each. The DRs may be subdivided based on their pharmacological behavior into the D1-like and the D2-like subfamilies, and these are ideal targets for treating schizophrenia and Parkinson’s disease; therefore, development of improved remedies would be significantly enhanced with the availability of the 3D structure for the DR and of the dopamine-binding site. We report here the 3D structure of the long isoform of human D2 DR, hereafter referred to as D2DR, predicted from primary sequence using the MembStruk (3–5) and HierDock (3) first-principles theoretical and computational techniques. The structure is validated by predicting the binding site and relative binding affinities of dopamine, three known dopamine agonists, and seven known antagonists. The predicted binding site for dopamine and agonists is located between transmembrane (TM) helices 3, 4, 5, and 6, whereas the best antagonists bind to a site involving TM helices 2, 3, 4, 6, and 7 with minimal contacts to TM helix 5. The predicted binding sites contain the critical residues for binding dopamine and several antagonists that have been identified by mutation studies and give relative binding affinities that correlate well with experiment.

Computational Methods

All calculations for the protein used the DREIDING force field (FF) (8) with charges from CHARM82 (9) unless specified otherwise. The non-bond interactions were calculated by using the cell multipole method (10) in MPSim (11). The ligands were described with the DREIDING FF using Gasteiger charges (12). For the lipids we used the DREIDING FF with QEq charges (13). All of the calculations treated the solvent (water) using the analytical volume generalized born approximation to the Poisson–Boltzmann solvation model (14).

MembStruk Structure Prediction Method for D2DR. The MembStruk procedure (MembStruk3.5), used to predict the 3D structure of D2DR, is described in ref. 5. Here we detail the steps that are relevant to the prediction of D2DR. The various steps of the MembStruk3.5 as applied to D2DR are as follows.

*To whom correspondence should be addressed. E-mail: wag@wag.caltech.edu.

Abbreviations: GPCR, G protein-coupled receptor; FF, force field; TM, transmembrane; DR, dopamine receptor.

© 2004 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0400100101
the multisequence hydrophobicity moments of the middle third of each helix about their maximum hydrophobic center. This analysis yielded a clear consensus on which residues should contact the membrane and which residues should face the receptor interior.

**Optimization of helical bends and kinks.** The kinks and bends in each canonical helix was optimized with Newton–Euler inverse mass operator (NEIMO) torsional dynamics (18, 19) or Cartesian dynamics (described with the Dreiding FF and Charmm22 charges) for 500 ps at 300 K constant temperature, and we picked the minimum energy conformation from the dynamics. The helical bundle now has helices with bends and kinks. The rotational orientation of these noncanonical helices was further optimized by using both the procedure in step 4 followed by energy-based optimization called “Rotmin” described in ref. 5. In the energy-based rotational orientation optimization, we performed systematic rotations on all of the seven helices over 5° at a time and over a small grid of rotation angles from 50° to 50°. At each rotation angle the side chains were added by using the side chain placement program SCWRL5.0 (20) followed by potential energy minimization using conjugate gradients to 0.1 kcal/mol per Å rms deviation in force per atom. Steps 3, 4, and 6 are part of a systematic search algorithm for optimum translational and rotational orientation, and these steps allow conformational search and aid in getting over large energy barriers between various conformations of the TM barrel.

**Equilibration of the optimized TM barrel.** The optimized TM barrel structure from the previous step 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, treating each molecule as a rigid body for 50 ps at a constant temperature of 300 K using the MPSim code (11).

**Loop addition and final optimization.** The interhelical loops were built by using WHATIF (21), and disulfide bonds were formed between Cys-107 in TM3 and Cys-182 in extracellular loop 2. This full system was then optimized with the conjugate gradient minimization technique to 0.1 kcal/mol per Å rms in force.

**Prediction of Ligand Binding Sites and Binding Energies.**

**Ligand structure preparation.** The 10 ligands shown in Fig. 1, were built with CHEMDRAW, and the 2D structure was converted to 3D structures by using POLYGRAF software. These ligands consist of agonists dopamine, 7-hydroxydipropylaminotetralin, apomorphine, and bromocriptine and antagonists like clozapine, domperidone, haloperidol, etc. We have categorized the antagonists...
studied here into two classes: (i) class I, clozapine-like bulky antagonists; and (ii) class II antagonists that have two aromatic or ring moieties connected by a flexible linker with a protonated amine group as in haloperidol. Hydrogens were added, and Gasteiger charges were assigned to all of the ligands with appropriate protonated states. We then minimized the potential energy of each ligand using conjugate gradients to an rms deviation in force of 0.1 kcal/mol per Å.

**Function prediction method for D2DR.** HierDock protocol is a hierarchical strategy ranging from coarse grain docking to fine grain optimization for docking ligands in proteins and determining their putative binding site. This method has been tested for various GPCRs (3–5, 22), outer membrane protein A (23), and globular proteins (24–28). (The protocol has been described in detail in these references.) The version of HierDock used in this study is described in ref. 4. In brief, the various steps of HierDock version 2.0 protocol is as follows:

1. First we carried out a coarse grain docking procedure to generate a set of conformations for ligand binding in the receptor. Here we used DOCK4.0 (29) to generate and score 500 conformations, of which 50 (10%) were selected by using a buried surface area cutoff of 90% and 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 with the protein.

2. The 50 best conformations selected for each ligand from step 1 were subjected to all-atom minimization keeping the protein fixed but the ligand movable. The solvation energy of each of these 50 minimized conformations was calculated by using analytical volume generalized born continuum solvation method (14). Then the five best scoring conformations based on the potential energy of the ligand in the protein were selected from these 50 conformations for the next step.

3. Next we optimized the structure of the receptor/ligand complex allowing the structure of the protein to accommodate the ligand. This was 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 calculated 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 potential energy (PE) of the ligand in the protein and the potential energy of the ligand in water. The energy of the ligand in water is calculated by using DREIDING FF and analytical volume generalized born continuum solvation method.

4. Next we selected, from the five structures from step 3, the one with the maximum number of hydrogen bonds between ligand and protein. For this structure we used the Scream side chain replacement program (V.W.T.K., N.Y., and W.A.G., unpublished observations) to realign all side chains for the residues within 5 Å in the binding pocket. (This uses a side-chain rotamer library of 1,478 rotamers with 1.0-Å resolution, with the all-atom DREIDING energy function to evaluate the energy for the ligand–protein complex.) The binding energy of all of the five optimized complexes was calculated by using Eq. 1.

**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 by using autoMS utility of DOCK4.0. Spheres were generated to fill up the void regions of the entire receptor using the sphen utility of DOCK4.0. The program PASS (30) was then used to locate plausible centers of large void regions in the receptor. The spheres that were within 5.0 Å of these centers were clustered for docking of ligands. For D2DR we obtained nine regions where we applied the only steps 1 and 2 of the HierDock protocol. The best energy conformation for each ligand in each region was selected based on (i) location of the binding site (those in the intracellular loop or completely in the extracellular loop regions were omitted for D2DR), (ii) buried surface area (90% cutoff value was used), and (iii) interaction energy of the ligand with the protein. If more than one region scanned had similar energies within a certain tolerance factor, then both those regions were chosen as putative binding regions. A tolerance of 30 kcal/mol was used because the minimization of the ligand with the fixed receptor structure was done for only a fixed number of 50 conjugate gradient steps. Using this procedure we found similar putative binding sites for agonists and class I antagonists whereas, the predicted binding sites for class II antagonists were different. Class II antagonists had three contiguous regions with the best bound conformation having ligand–receptor interaction energies within 30 kcal/mol of each other. Hence, for these long antagonists we merged the spheres of these three regions together as the putative binding site. Because the number of spheres in the merged region was high (>300), the sphere density was thinned to ~100 spheres. Details of these energies will be published elsewhere.

**Prediction of binding sites and binding energies.** Once the putative binding sites were determined for agonists and antagonists, all of the agonists and antagonists were docked into their respective putative binding regions using HierDock protocol steps 1–4, and the five best bound structures for each ligand was chosen. Herein we performed iterative docking: the structure with side chains replaced and minimized was used in the generation of a new sphere set, and the target ligands were redocked to this optimized binding structure by performing HierDock procedure steps 1–4. This iterative HierDock procedure optimizes the side chain conformation of the residues in the binding site for ligand binding.

**Results and Discussion**

Fig. 2 shows the 3D structure of the human D2 DR predicted by using MembStruk3.5 and the predicted binding site of dopamine determined by using HierDock. During the scanning procedure to determine the binding site, only one site was found to be favorable for binding dopamine and other agonists studied. Thus, the predicted binding site of dopamine is located in the top third of the 7-TM barrel involving TM domains 3–6. The amino acid residues within 5.5 Å of the dopamine binding site in D2DR are shown in Fig. 3. We found the following residues to be essential for binding of dopamine in the human D2 receptor:

1. Asp-114 in TM3. The carboxyl group of the aspartate forms a tight salt bridge (2.6 Å) with the primary amino group of dopamine. This residue is conserved over all five human DRs, as well as in all human biogenic amine receptors. Mutation studies have implicated this residue in the direct binding of the dopamine to D2DR (31).

2. Ser-193 and Ser-197 in TM5. These residues hydrogen-bond to the metahydroxyl (2.7 Å) and parahydroxyl (2.7 Å) groups, respectively, of the catechol ring of dopamine, playing an essential role in recognizing dopamine. These two residues are conserved over all five human DRs and have been shown to be involved in direct binding of ligand through mutation studies (31). Ser-194 is also conserved over all five human DRs, and mutation studies indicate that it is involved in binding of the ligand. In our structure, Ser-194 is hydrogen-bonded to the backbone nitrogen of residue 192 rather than dopamine. However, Ser-194 might serve as an alternate to Ser-193 in hydrogen bonding to the metahydroxyl group of the catechol for the slightly modified structure of the receptor that might result from activation.
3. Phe-110, Met-117, Cys-118 (TM3), Phe-164 (TM4), Phe-189, Val-190 (TM5), Trp-386, Phe-390, and His-394 (TM6) form a mostly hydrophobic pocket for dopamine. We find Trp-386 and Phe-390 of the conserved WXXFF motif on TM6 to be within the 4.5-Å binding pocket, but Phe-389 is 7.2 Å away from dopamine. Thus, the predicted binding site (involving TM domains 3–6) is in excellent agreement with the residues determined experimentally to be involved in binding (31–38). This provides a good validation of the predicted structure for D2DR and of the predicted binding site of dopamine. We also carried out full HierDock predictions for the binding sites or three other known agonists: apomorphine, 7-hydroxydipropylaminotetralin, and bromocriptine, and we found very similar active sites. In particular, all dopamine agonists bind tightly to both Ser-193 and Ser-197 in TM5 in addition to Asp-114 in TM3 just as in dopamine. These predicted structures can be further tested by selective mutation of some of the additional residues we have identified in the active site.

To further validate the predicted structure for D2DR, we used the HierDock procedure to predict the binding conformation and the binding energy for seven well-studied antagonists. We found two classes of antagonists.

Class I antagonists (exemplified by clozapine) occupy the region between TM3, TM4, TM5, and TM6 (the agonist binding pocket). Thus, clozapine makes (i) a 2.8-Å salt bridge to Asp-114 (TM3); (ii) a hydrogen bond to Ser-193 (TM5) (3.2 Å) but not to Ser-194 or Ser-197; (iii) heteroatom interactions with Trp-386 (TM6) (3.1 Å); and (iv) a mostly hydrophobic pocket shown in Fig. 4 formed by Val-87 and Trp-90 (TM2), Phe-110, Leu-113, Val-115, Met-117, and Cys-118 (TM3), Phe-164 (TM4), Phe-189, Val-190, Ser-194, and Ser-197 (TM5), Phe-382, Trp-386, Phe-389, and Phe-390 (TM6), and Thr-412, Trp-413, Tyr-416, and Ser-419 (TM7).

![Fig. 2. Predicted binding site of dopamine (shown in spheres) in the predicted structure of human dopamine D2 receptor.](image)

![Fig. 3. Residues within 5.5 Å of the dopamine binding site in the human dopamine D2 receptor. The numbers shown in parentheses are the TM helix to which the residue belongs. The distance between Asp-114 and the primary amine group in dopamine is 2.6 Å. Ser-193 on TM5 makes a 2.7 Å hydrogen bond with the metahydroxyl group, and Ser-197 makes a 2.7 Å hydrogen bond to the para hydroxyl group of dopamine.](image)

![Fig. 4. Residues within 5.5 Å of clozapine bound to human D2DR.](image)
Class II antagonists, exemplified by haloperidol, occupy the region between TM2, TM3, TM6, and TM7, with minimal contacts to TM4 and TM5. Thus, haloperidol makes (i) a 2.8-Å salt bridge to Asp-114 (TM3); (ii) a hydrogen bond (3.2 Å) to Ser-197 (TM5) (3.2 Å) but not to Ser-193 or Ser-194; (iii) heteroatom interactions with Trp-386 (TM6) at 3.8 Å and Trp-90 (TM2) at 3.0 Å; and (iv) a mostly hydrophobic pocket (as shown in Fig. 5) provided by Val-87, Val-91, and Leu-94 (TM2), Phe-110, Leu-113, Val-115, Met-117, and Cys-118 (TM3), Trp-160 and Phe-164 (TM4), Phe-189, Val-190, and Val-196 (TM5), Trp-386, Phe-389, Phe-390, and His-393 (TM6), and Ser-409, Thr-412, Trp-413, Tyr-416, and Val-417 (TM7). Other class II antagonists with very similar binding sites include spiperone and sulpiride. Interestingly, many class II antagonists have been shown to crossreact with other amine receptor and other class A GPCRs. In some GPCRs, CCR1 for example, the conserved TM3 Asp has replaced with a longer-chain Glu. From distant restraints and structural considerations it appears that a longer linker alkyl chain in the ligand would push the protonated amino group of the ligand closer to the TM2 and TM7, and this longer linker alkyl chain ligand would bind better and be specific to GPCRs with Glu in this position.

The residues we found to be important for binding of agonists and antagonists are consistent with all available data from substituted cysteine accessibility method, radiolabeling, and mutation experiments (31) on these receptors. All seven agonists make a tight contact to Asp-114 in TM3, but none forms strong contacts to both Ser residues in TM5, in contrast with our observation that all agonists have strong coupling to Asp in TM3 and both Ser in TM5. This finding suggests that strong coupling between TM3 and TM5 is essential for dopaminergic transduction (consistent with experiments showing that hydrogen bonds to both conserved serines of TM3 and TM5 are essential for dopamine activation) (31–38). The predicted structures for antagonists lead to a weakening of the coupling between TM3 and TM5 and (particularly for class II antagonists) prevent motion between TM3 and TM6.

Based on the criteria that a salt bridge to TM3 and two hydrogen bond contacts to TM5 are both essential for activation, whereas the salt bridge and one hydrogen bond are important for antagonists, we can categorize all ligands as agonists and antagonists. This criteria correctly predict the nature of all ligands studied here. Recently, an agonist in the dipropylaminotetralin series (39) has been reported with no hydroxyl group. This finding may indicate a more complicated criterion for agonists and antagonists, we can categorize all ligands as agonists and antagonists, because agonists involve tight binding between TM3 and TM5 whereas antagonists involve tight binding of the ligand between TM helices 3, 4, 5, and 6 but that the strongest binding antagonists bind to a site involving TM helices 3 and 5 whereas antagonists involve tight binding of ligand between TM helices 3 and 6.

The validation of the predictions for D2 justify carrying out similar studies to predict the structures and ligand binding to the other four human DRS. We hope that this study will provide a basis for designing agonists and antagonists selective to binding to just one of the five DR subtypes, which could be of tremendous value in treating dopamine-related diseases while minimizing side effects.

This research was initiated with support from the Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI) and was completed with support from National Institutes of Health Grants BGR01-GM625523, R29AI40567, and HD36385. The computational facilities were provided by a Shared University Research grant from IBM and Defense University Research Instrumentation Program grants from ARO and the Office of Naval Research (ONR). The facilities of the
Materials and Process Simulation Center are also supported by the Department of Energy in addition to the National Science Foundation,