JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

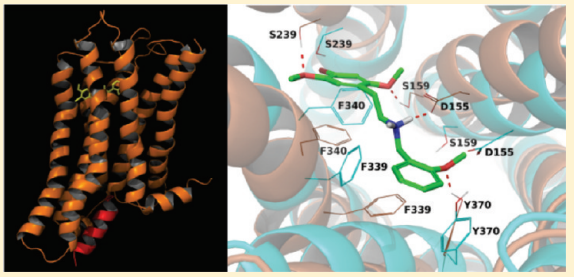
G Protein- and Agonist-Bound Serotonin 5-HT$_{2A}$ Receptor Model Activated by Steered Molecular Dynamics Simulations

Vignir Ísberg, Thomas Balle, Tommy Sander, Flemming Steen Jørgensen, and David E. Gloriam*

Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark

Supporting Information

ABSTRACT: A 5-HT$_{2A}$ receptor model was constructed by homology modeling based on the β_2-adrenergic receptor and the G protein-bound opsin crystal structures. The 5-HT$_{2A}$ receptor model was transferred into an active conformation by an agonist ligand and a $G_{\alpha q}$ peptide in four subsequent steered molecular dynamics (MD) simulations. The driving force for the transformation was the addition of several known intermolecular and receptor interhelical hydrogen bonds enforcing the necessary helical and rotameric movements. Subsquent MD simulations without constraints confirmed the stability of the activated receptor model as well as revealed new information about stabilizing residues and bonds. The active 5-HT$_{2A}$ receptor model was further validated by retrospective ligand screening of more than 9400 compounds, whereof 182 were known ligands. The results show that the model can be used in drug discovery for virtual screening and structure-based ligand design as well as in GPCR activation studies.



INTRODUCTION

The more than 800 G protein-coupled receptors (GPCRs)[1] constitute one of the largest protein superfamilies in the human genome[2] and are involved in a wealth of physiological processes. GPCRs mediate extracellular signals into intracellular responses and are activated by diverse endogenous ligands such as monoamines, peptides, lipids, and metabolites or external stimuli, such as light, smell, and taste.[3] Consequently, GPCRs have a tremendous therapeutic and commercial potential and 30−45% of drugs currently on the market target GPCRs.[4] Even so, their therapeutic value is largely unexploited as the drugs in the clinic only target around 30 different receptors, representing ca. 4% of the repertoire.[5]

The receptors for serotonin, or 5-hydroxytryptamine (5-HT), belong to the largest GPCR family, Family A or the *Rhodopsin* family. The serotonergic neurotransmitter systems, consisting of 12 GPCRs and one ligand-gated ion channel, is believed to be one of the first neurotransmitter systems to evolve and predates the muscarinic, dopaminergic, and adrenergic receptor systems.[6] Serotonergic receptors are engaged in a large number of peripheral (development, cardiovascular, gastrointestinal and endocrine function) and central (cognition, memory, aggression, appetite, sex, sleep, and mode regulation) processes.[6] Clinically, the serotonergic system has been linked to several mental and mood disorders (e.g., depression, anxiety, social phobia, schizophrenia, obsessive-compulsive disorder, and panic disorder) and also migraine, hypertension, pulmonary hypertension, eating disorders, vomiting, and irritable bowel syndrome.[7] The 5-HT$_{2A}$ receptor is one the most studied serotonergic receptors, and its activation stimulates the secretion of several hormones such as ACTH, corticosterone, oxytocin, renin, and prolactin.[6] 5-HT$_{2A}$ stimulation, for example by lysergic acid diethylamide (LSD), induces powerful psychedelic effects. 5-HT$_{2A}$ inhibition by clinical drugs has antipsychotic (e.g., clozapine and olanzapine) and antidepressive (e.g., mirtazapine and mianserin) effects.[7]

5-HT$_{2A}$ agonists are classified into three main groups: tryptamines, ergolines, and phenethylamines.[8] The tryptamines, which include serotonin/5-HT, are derivates of the amino acid tryptophan. Naturally occurring tryptamines have long been used because of their psychoactive effects.[8] The ergolines[8,9] are structurally similar to tryptamines, but they are less flexible. The best known ergoline is LSD, which discovery in 1943 marked the beginning of research on hallucinogens. For the phenethylamine class, the model compound is mescaline, which is found in the peyote cactus *Lophophora williamsii*. Mescaline is one of the oldest psychedelics and has been used for religious traditions by native American Indians for thousands of years.[10] A large number of mescaline analogs have been synthesized and the structure−activity relationships are well understood.[11] Very few selective 5-HT receptor ligands are known, and as subtype selectivity is an important clinical factor, there is a need for development of more and better subtype-selective 5-HT agonists and antagonists.

GPCRs share a common structure consisting of seven transmembrane α-helices connected by loop regions, an extracellular

Received: October 13, 2010
Published: January 24, 2011

Table 1. Overview of the Distance Constraints Used in the Steered MD Simulations[b]

		activation of model				observation of interactions	
simulation purpose	constrained bonds	1	2	3	4	5a	5b
prevention of unfolding at sites of sequence truncation	α-helical backbone bonds in 5-HT_{2A} and $G_{\alpha q}$ peptide	X	X	X	X	X	X
ligand binding	receptor and ligand	X	X	X	X	X	
G protein binding	receptor and $G_{\alpha q}$ peptide	X	X	X	X		X
interhelical stabilization	interhelical-connecting residues	X	X	X			
prevention of local helical distortions at constraints	all α-helical backbone bonds	X	X[a]				

[a] In simulation 2, the harmonic potential of the α-helical bonds was weaker than in simulation 1. [b]All distance constraints, except two aromatic ligand binding residues, represent hydrogen bonds.

amino terminus, and an intracellular carboxy terminus. The greatest sequence conservation is seen within the transmembrane helices (TMHs), which are believed to form a common activation and signal transduction machinery.[12] The idea of a common activation machinery is supported by the high structural similarity of the available GPCR crystal structures; those for (rhod)opsin,[13] the β_1-[14] and β_2[15]-adrenoreceptors, the adenosine A_{2A} receptor,[16] the chemokine CXCR4 receptor,[17] and the dopamine D_3 receptor[18] in complex with antagonists or inverse agonists. Receptor activation is thought to involve conformational changes in the form of both helical movements, primarily a tilt of helix 6 as formulated in the "global toggle switch" model,[19] as well as rotameric alteration of key side-chains, referred to as "activation micro-switches".[20] The only active receptor structure published so far is of opsin.[13a] In opsin the TMH bundle is widened, by an outward movement of helices 6 and 5, on the intracellular side to accommodate for binding of the G protein. However, opsin is not inhibited, not activated, by its natural ligand and thus might not be representative for the upper, ligand-binding part of activated receptors. Consensus is that the helical bundle tightens to accommodate agonist binding, for example aminergic ligands interact with TMHs 3, 5, and 6, but are too short to reach them simultaneously in the inactive receptor conformation.

This study presents a model of the human 5-HT_{2A} receptor, including an agonist and a $G_{\alpha q}$ peptide, that has been activated *in silico* using advanced molecular dynamics (MD) simulations steered to fulfill known structural features of GPCR activation. Subsequent unconstrained simulations revealed new interactions also associated with the active conformation. The model has been successfully validated by a retrospective ligand screening seeded with 182 known 5-HT_{2A} ligands and will constitute an important tool for the development of novel 5-HT agonists, especially of the phenylethylamine class.

METHODS

5-HT_{2A} Homology Modeling. The high-resolution crystal structure (PDB: 2RH1)[15] of the closely related, but inactive, β_2-adrenoceptor (β_2AR) was used as the main template. Prior to 5-HT_{2A} homology modeling, 4 modifications were made to the β_2AR structure to incorporate template features for active GPCR conformations only available in the G protein-bound opsin crystal structure (PDB: 3DQB).[13a] (1) The lower parts of TMHs 5 and 6 were manually tilted outward. (2) The missing structure of the β_2AR intracellular loop 3 (ICL3) was replaced by the opsin version (sequence substitution of LQKIDKSEGRFHVQNLSQVEQDGRTGHGLRRSSKFCLK for AQQQESATTQKA). (3) The $G_{\alpha i}$ peptide backbone was inserted and mutated to $G_{\alpha s}$. (4) R131$^{3.50}$ was set at its G protein-interacting rotamer. In addition, several crystal structures of $G_{\alpha s}$ proteins (PDB: 1TL7, 1U0H, and 1CJK)[17,18] were overlaid and used to elongate the $G_{\alpha s}$ peptide until the lower end of ICL3. In addition, the rotameric state of W286$^{6.48}$ (the "rotamer toggle switch") was switched to its presumed active conformation (chi1-chi2 angles of 170 and 130, respectively) closer to TMH5. This modified β_2AR structure was used as a template for homology modeling of 5-HT_{2A} with Modeler (Mod9v6).[21] The sequence alignment is available in Supporting Information, as SI1. The $G_{\alpha s}$ C-terminal peptide was manually mutated to the $G_{\alpha q}$ subtype and selected side-chains configured to interact with the 5-HT_{2A}.

5-HT_{2A} Model *In Silico* Activation Using Steered Molecular Dynamics Simulations. *Simulation Settings.* Molecular dynamics simulations were conducted using the Schrödinger implementation of the Desmond software package[22] (version 2.4) and the OPLS-AA 2001 force field. Using the Desmond system builder, a 10 Å buffered orthorhombic system with periodic boundary conditions was built with a DPPC lipid membrane and TIP4P explicit water solvent. The overall charge was neutralized by chloride ions. Simulations were performed in the NPT ensemble. The temperature, 325 K, and pressure, 1.01325 bar, were kept constant by coupling the system to a Berendsen thermostat and barostat. An integration step of 2.0 was used, Coulombic interactions were calculated using a cutoff radius of 9.0 Å, and long-range electrostatic interactions were calculated using the smooth particle mesh Ewald method. Before each MD simulation, a modified version (SI2) of the default Desmond membrane protein relaxation protocol was applied.

MD Simulations and Distance Constraints. We simulated receptor activation by agonist and G protein in 4 subsequent MD simulations of 5 ns (ns) each and 20 ns in total. Hydrogen bonds of interhelical, ligand, and G protein interactions were collated from mutagenesis and X-ray crystallographic data. These were included in the simulations as pairwise (donor−acceptor) atom distance harmonic constraints (default between hydrogen and acceptor: 1.8 Å; G protein and helix 8: 2.2 Å; protonated nitrogen in ligand and − D155$^{3.32}$ gamma carbon: 3.0 Å). These constraints drive the helical bending and tilting and side-chain rotameric changes until the hydrogen bonding distances are satisfied i.e. until a presumed active conformation is reached. To prevent local distortions, the α-helical backbone hydrogen bonds were constrained during the first 2 simulations. After the third simulation, the stabilizing interhelical constraints (adapted from the opsin structure) were removed to assess the stability of known and observe the formation of new interactions. After having completed

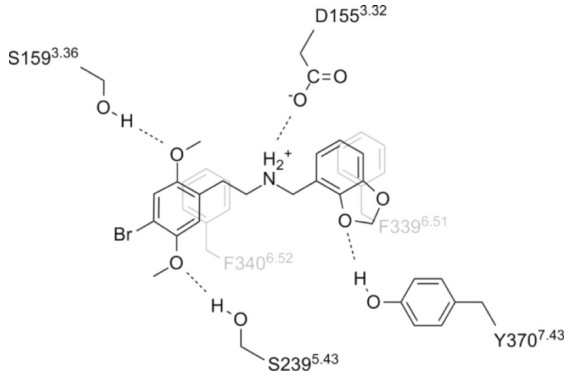


Figure 1. Binding mode of 25Br-NBMD in 5-HT$_{2A}$, as described for phenylethylamines in the literature. Constrained hydrogen bonds are displayed as dashed lines. The 2-MeO[23] and 5-MeO[24] groups accept hydrogens from S159$^{3.36}$ and S239$^{5.43}$, respectively. The amine donates one or two hydrogens to D155$^{3.32}$.[25] The two aromatic rings of the ligand interact with F339$^{6.51}$[26] and F340$^{6.52}$,[27] respectively (shaded). During the third simulation, the zeta carbon of F339$^{6.51}$ and the hydroxyl hydrogen of Y370$^{7.43}$ were constrained at a distance of 3.6 Å. Furthermore, the N-benzyl oxygen is important for affinity and was here proposed to hydrogen bond to Y370$^{7.43}$.

the model (simulation 4), two additional parallel observation simulations were performed without ligand (5b) or G protein (5a) constraints. All constraints are summarized in Table 1, described in the text below, and were added as to the Desmond structure file as "pseudo-bonds" using a python script (SI3).

Ligand - Receptor Bond Constraints. Agonist binding was simulated by docking N-(2,3-methylenedioxybenzyl)-4-bromo-2,5-dimethoxyphenethylamine (25Br-NBMD) into the 5-HT$_{2A}$ homology model and applying MD distance constraints. 25Br-NBMD was chosen because it can form all hydrogen bonds as reported from mutagenesis studies, as shown in Figure 1, and also includes an N-benzyl oxygen moiety, which has been shown to increase affinity.

G Protein - Receptor Bond Constraints. The G protein-bound opsin structure[13a] includes two interprotein hydrogen bonds between the backbone oxygen atoms of C347 and K345 to R135$^{3.50}$ and Q312$^{8.49}$, respectively. The first bond was defined in a distance constraint. The second bond could not be directly translated into our 5-HT$_{2A}$ model as the opsin Q312$^{8.49}$ is not conserved but is a threonine, which is too short to reach the backbone of the G protein. However, preliminary MD simulations of this region of the receptor suggested a reversed hydrogen bond between the same residues, i.e. the G protein lysine acting as donor, and this interaction was constrained.

Interhelical Receptor Bond Constraints. Investigation of the opsin structure reveals several interhelical hydrogen bonds that are not present in any inactive GPCR structure (Table 2). These hydrogen bonds are believed to stabilize the active conformation and were consequently constrained. In addition to the hydrogen bonds, Y380$^{7.53}$ is stabilized by a hydrophobic network consisting of multiple residues. An approach chosen was to anchor the γ-carbon of the tyrosine to its own backbone with two distance constraints, to the backbone nitrogen (3.1 Å) and oxygen (3.9 Å), respectively.

Receptor and G Protein Helical Backbone Bond Constraints to Prevent Local Distortions. Local distortions of the α-helical backbone were prevented by applying gradually decreasing constraints on the α-helical backbone hydrogen bonds during the first two simulations. In this case the purpose was to maintain, rather than change, distances, and thus the distances of all such hydrogen bonds were set at their length in the original (pre-MD) homology model. To allow for tilting and bending of helices, no constraints were defined in the immediate vicinity of proline kinks, which are known to constitute "hinge" areas.[28]

Receptor and G Protein Helical Backbone Bond Constraints To Prevent Unfolding at Truncated Sites. Regions where the structural templates were incomplete are protected by helical backbone hydrogen bond constraints through all MD simulations. As the 5-HT$_{2A}$ model is missing both the N- and C-termini, the ends of helices 1 and 8 were stabilized (residues 70−78 and 387−395, respectively). Furthermore, as ICL3 had been truncated to the corresponding length in the opsin template, the intracellular ends of helices 5 and 6 were also constrained (residues 260−268 and 312−320, respectively). Finally, since only a small part of the G_α subunit was modeled, the peptide (residues 345−356) was also stabilized by constraints.

Model Validation by Retrospective Ligand Screening. Ligand docking was carried out using Glide.[29] Using the output structure of MD simulation step 4 with waters, lipids, and ions removed; a receptor grid was created around the ligand present in the binding site. The ligand structures were built in Maestro[30] and comprised 182 known 5-HT$_{2A}$ ligands; 139 agonists (112 phenethylamines, and 27 agonists from other structural classes), 39 antagonists, and 4 inverse agonists collected from the literature. In addition, the first 10,000 compounds were extracted from the ZINC database[31] preclustered (Tanimoto score 60%) "drug-like" subset. Duplicate structures were removed, and the remaining 9257 compounds were added to the screening database. Standard precision docking was performed using input partial charges, and the options to add Epik[32] state penalties to docking score and apply strain correction terms were selected. The van der Waals radii scaling factor was set to 0.8 with a partial charge cutoff of 0.15. Ten poses per ligand were saved. Induced fit docking was carried out using the induced fit docking workflow in Maestro (implementing Glide and Prime[33]) with default settings. The output structure of MD simulation step 4 with waters, lipids, and ions removed was used for preparation of the receptor grid, which was defined around the ligand. The ligand set consists of the 182 known 5-HT$_{2A}$ ligands used in the standard precision docking.

Molecular Depiction. Two-dimensional molecular structures were drawn in ChemBioDraw Ultra.[34] Three-dimensional structures were depicted in PyMOL.[35] Throughout this article, receptor residues are referred to by their one-letter code followed by their full sequence number. TMH residues also have a Ballesteros-Weinstein (BW) index[36] in superscript.

RESULTS

Extracellular Face and Ligand Binding Site. Figure 2 shows the active 5-HT$_{2A}$ model (SI4) superimposed to the β_2AR template (Figure 2A) and initial (before MD) 5-HT$_{2A}$ homology model (Figure 2B). Most of the helical movements occurred during the first MD simulation, whereas the subsequent simulations displayed smaller changes. This is because constraints, which are harmonic potential forces, are stronger at longer (ligand−receptor interaction) distances. Figure 2A shows that, in accordance with the "global toggle switch model",[19] the largest extracellular movement is an inward tilt of TMH6. Further investigation also shows that TMH5 has moved slightly toward

Table 2. Interhelical Bonds in MD Simulations[b]

	prevalence in simulations	AA 1	AA 2	original length (Å)	model length (Å)	comment
KNOWN	all	N92$^{1.50}$	D120$^{2.50}$	1.7	2.0	only seen in opsin
	all	S115$^{2.45}$	W200$^{4.50}$	2.0	1.9	seen in all crystallized GPCRs, -inactive and active
	1−3	R173$^{3.50}$	Y254$^{5.58}$	1.6	broken	3.50 participates in the "ionic-lock"
	all	Q262$^{5.66}$	E318$^{6.30}$	2.3	1.8	6.30 participates in the "ionic-lock"
	all	Y370$^{7.53a}$	-	3.1 and 3.9	3.3 and 4.3	keeps TMH7 from TMH6
NEW	all	S115$^{2.45}$	H165$^{3.42}$	3.7	2.2	
	complete	L166$^{3.43}$, S170$^{3.47}$, R173$^{3.50}$, Y254$^{5.58}$, V328$^{6.40}$, N376$^{7.49}$, Y380$^{7.53}$		-	-	water-mediated H-bond network
	from 4 in 5b	R173$^{3.50}$, Y254$^{5.58}$, L325$^{6.37}$, Y380$^{7.53}$		-	-	H-bond network
	in 5b	D120$^{2.50}$	S373$^{7.46}$	3.3	2.0	

[a] Y370$^{7.53}$ does not form a hydrogen bond but is stabilized by a hydrophobic network (see Methods). [b] The "Original length" was defined for bonds as follows (bonds: source); the R173$^{3.50}$-Y254$^{5.58}$ bond: manually, other constrained bonds: from the G protein-bound opsin structure and new bonds: from the pre-MD 5-HT$_{2A}$ homology model.

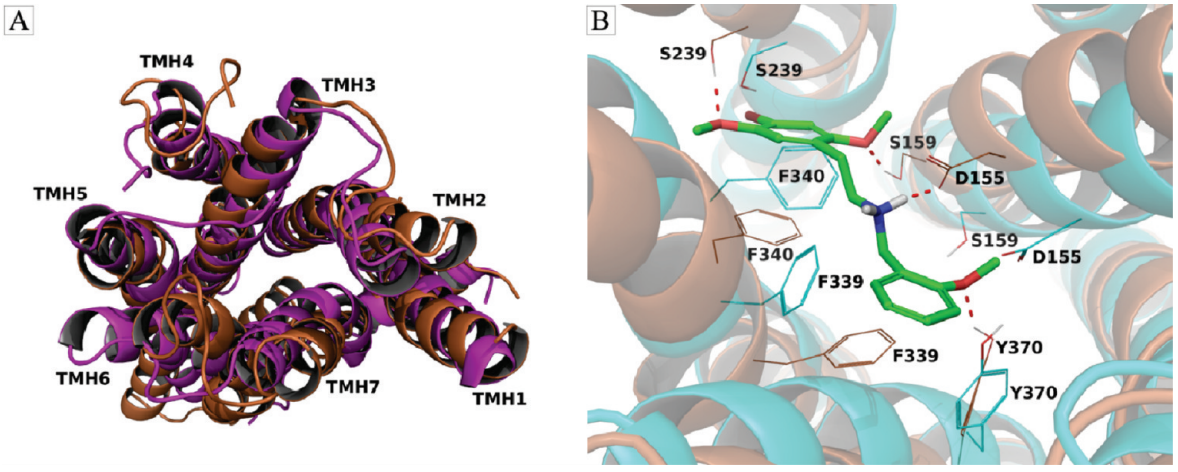


Figure 2. (**A**) Extracellular view of a superimposition of the active 5-HT$_{2A}$ model (brown) and β_2AR structure (magenta). The most important helical movements involve a tightening of the ligand binding site between TMH3, TMH5, and TMH6. (**B**) A cartoon representation of a superimposition of the active 5-HT$_{2A}$ model (brown) and the initial (pre-MD) 5-HT$_{2A}$ homology model structure (cyan). Hydrogen bonds to the ligand, 25Br-NBMeO, (green carbons) are shown as red dashed lines.

TMH6, and TMH3 is closer to TMH5. This meets the proposed requirement of a tightening of the triad TMHs 3, 5, and 6 to accommodate binding of aminergic ligands. Looking at other helices, the extracellular part of TMH7 is shifted slightly (2.4 Å of top TMH7 alpha carbons) sideways toward TMH1, a movement that seems necessary to make room for the relocation of TMH6. Finally, TMH1 and TMH2 have also moved to the side, possibly triggered by the sideways shift of TMH7.

As shown in Figure 2B, the binding residues have come closer to the ligand during the MD simulations. The differences are primarily caused by side-chain rotations, but also helical tilting and rotations are observed. The largest effect is seen for D155$^{3.32}$ and F340$^{6.52}$, but all interactions in the binding mode described in Figure 1 are satisfied. The largest movements occurred in simulations 1 (first application of constraints) and 3 (release of helical backbone constraints). The rmsd values (Å) of the ligand binding heteroatoms to the initial homology model were (simulation 1, 2, 3, 4); D155$^{3.32}$ (2.3, 1.8, 3.0, 3.6), S159$^{3.36}$ (2.5, 2.1, 2.9, 3.3), S239$^{5.43}$ (1.7, 1.0, 1.7, 1.7), and Y370$^{7.43}$ (2.2, 2.4, 1.9, 2.5). The rmsd values (Å) of the alpha carbons in ligand binding residues to the initial homology model were (simulation 1, 2, 3, 4); D155$^{3.32}$ (2.4, 2.5, 3.9, 4.0), S159$^{3.36}$ (2.9, 2.3, 3.6, 3.6), S239$^{5.43}$ (1.7, 1.8, 1.9, 2.1), and Y370$^{7.43}$ (2.1, 2.3, 1.1, 3.2).

As shown in Figure 1, all hydrogen bonds except that for the N-benzyl moiety have been supported experimentally. Preliminary ligand docking and MD identified two possible hydrogen bond donors for the N-benzyl moiety: N343$^{6.55}$ and Y370$^{7.43}$. However, mutation, N343A$^{6.55}$, does not affect the binding affinity of N-benzyl ligands, and tyrosine is a stronger hydrogen bond donor than asparagine. Consequently, we identified the tyrosine as the most likely hydrogen bond donor to the 2′-oxygen of the 25Br-NBMD N-benzyl moiety, and the corresponding distance constraint was applied in the MD simulations. Although, TMH7 has moved away from the ligand binding site during activation, Y370$^{7.43}$ has maintained a good position for hydrogen bonding interaction with the N-benzyl moiety of the ligand.

Intracellular Face and G Protein Binding Site. Figure 3A shows that TMH6 has moved outward to make room for the G

dx.doi.org/10.1021/ci100402f |J. Chem. Inf. Model. 2011, 51, 315–325

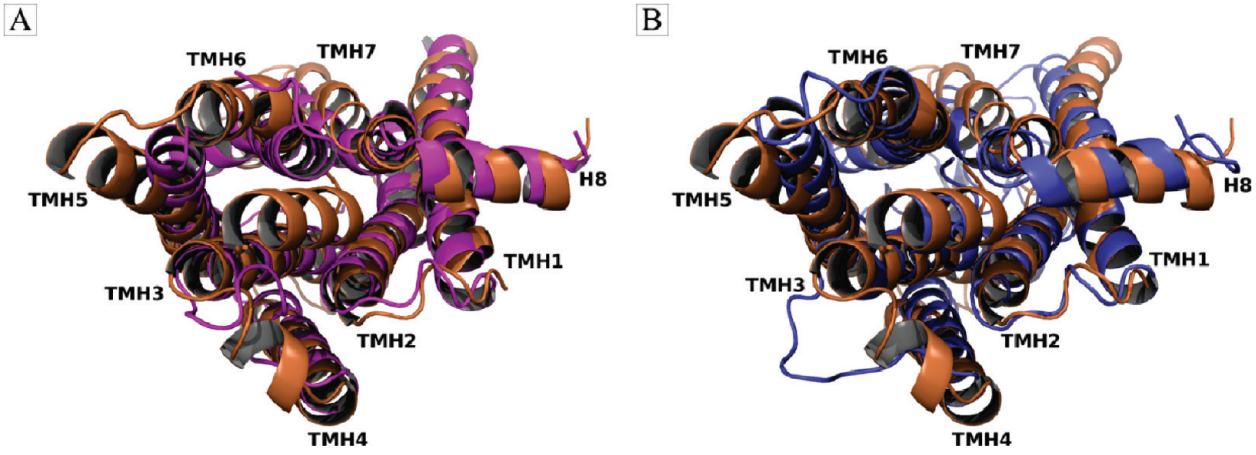


Figure 3. A view from the intracellular side of the active 5-HT$_{2A}$ model (brown) superimposed on the crystal structures of (**A**) β_2AR (magenta) and (**B**) G protein-bound opsin (blue).

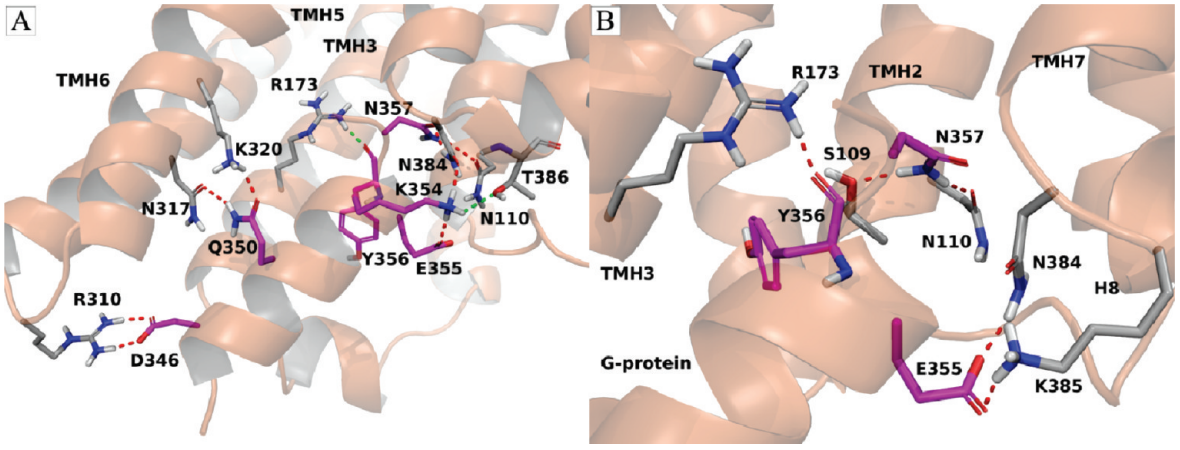


Figure 4. Hydrogen bonding networks between the 5-HT$_{2A}$ receptor (gray carbons) and the G protein (magenta carbons) as seen in (**A**) the finished model (after MD simulation 4) and (**B**) an extended simulation (5a) where the G protein constraints on known bonds have been released. Hydrogen bonds shown in green were previously known, whereas hydrogen bonds in red color formed during the MD simulations.

protein, in line with the "global toggle-switch model"[19] as well as biophysical measurements[37] and the opsin structure.[13] Furthermore, TMH5 has moved sideways to give room to TMH6. Figure 3B shows that the movements of TMHs 5 and 6 are larger in the 5-HT$_{2A}$ model than in the opsin structure. Similarly, an agonist-bound nanobody-stabilized crystal structure of β_2AR displays an 11 Å outward movement of cytoplasmic TMH6.[38] It is plausible that variations in interprotein interactions and conformations reflect differences in the sequences of receptors and G protein subtypes.

The G protein peptide backbone moved moderately (rmsd of 3.055) and preserved its orientation during the MD simulations. However, interestingly in addition to the two constrained, eight new hydrogen bonds (red color in Figure 4A) were formed during the simulations. The C-terminus, which is in a near identical position as before MD, displays a six-residue interaction network around K354. K354 was constrained only to T386$^{8.49}$ on helix 8, but post-MD it interacts with three residues; two from the receptor, T386$^{8.49}$ and N384$^{8.47}$, and one from the G protein, E355. Furthermore, this network includes hydrogen bonds to two asparagines on the G protein, N357, and TMH2, N110$^{2.40}$. In the middle of the G protein peptide, Q350 has formed hydrogen bonds with K320$^{6.32}$ and N317$^{6.29}$ on TMH6. Finally, in the lower end, the strongest interaction, a salt bridge, is observed between D356 and R310 on ICL3.

The G protein constraints were released (simulation 5a) to investigate the stability of existing and formation of new hydrogen bonds. As shown in Figure 4 B, the C-terminal network has changed, although the C347-R173$^{3.50}$ (known from opsin) and N357−N110$^{2.40}$ (formed in simulation 4) bonds are intact. The K354-Q312$^{8.49}$ (constrained initially) hydrogen bond has broken, and instead new ones are formed between E355 and the helix 8 residues N384$^{8.47}$ and K385$^{8.48}$. As the E355-N384$^{8.47}$ and E355-K385$^{8.48}$ bonds formed without constraints, they are considered as serious alternatives to the original K354-Q312$^{8.49}$ modeled prior MD. These bonds were considered in the modeling prior to MD, but the interacting residues were too distant before the simulations.

Activating Rotamer "Micro-switches". The "rotamer toggle switch", **W336**$^{6.48}$, (Figure 5A) is believed to swing toward TMH5 upon activation,[19] although this was not observed in opsin.[13a] W336$^{6.48}$ moved during the MD simulations to a position close to TMH5 where it can interact with F243$^{5.47}$, which has been suggested to stabilize the position of in the active state based on mutation data.[19] Figure 5A is also in agreement with experiments showing that long substituents in the 4-position

dx.doi.org/10.1021/ci100402f |J. Chem. Inf. Model. 2011, 51, 315–325

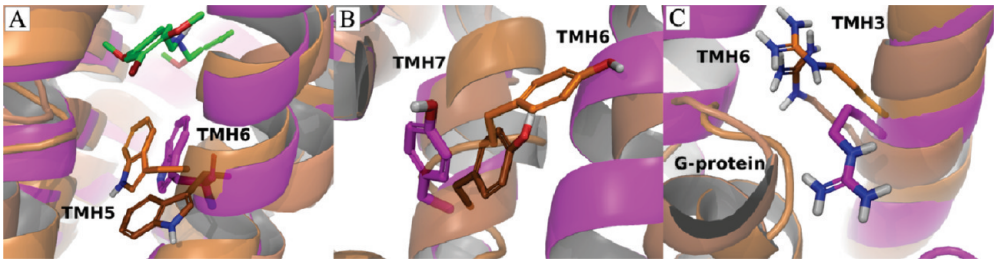


Figure 5. A comparison of the three activating microswitches: (**A**) W336$^{6.48}$, (**B**) Y380$^{7.53}$, and (**C**) R173$^{3.50}$ in β_2AR (magenta carbons), G protein-bound opsin (orange carbons), and the active 5-HT$_{2A}$ model (brown carbons). In **A**, the ligand, 25Br-NBMeO, is shown with green carbons.

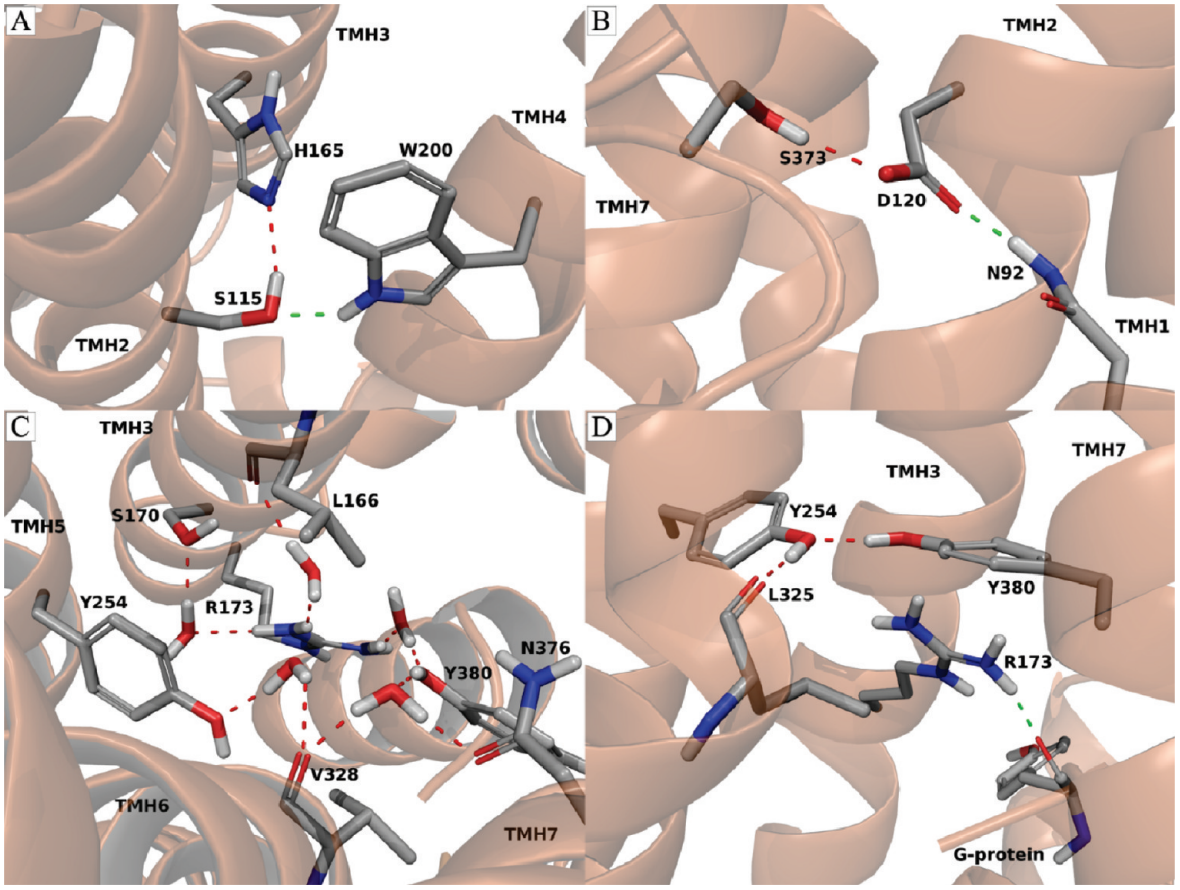


Figure 6. Interhelical hydrogen bonds formed spontaneously during simulations and constrained initially are displayed with red and green dashed lines, respectively. (**A**) A three-helix hydrogen bonding network between TMH2, TMH3, and TMH4. (**B**) A three-helix hydrogen bonding network between TMHs 1−2 and 7. (**C**) A water-mediated hydrogen bonding network between TMH3, TMH5, and TMH7. (**D**) A direct hydrogen bonding network between TMH5 and TMH7 and possibly a cation-PI interaction to TMH3.

(between TMH5 and TMH6) of phenethylamine agonists can transform them into antagonists,[39] because extra bulk would interfere with the movement of W336$^{6.48}$ to the active rotamer.

Y380$^{7.53}$ (Figure 5B) swings toward TMH6 upon activation and thereby stabilizes this helix in its outward shifted position making room for the G protein. In the 5-HT$_{2A}$ model, as in opsin, Y380$^{7.53}$ has moved closer to TMH6 and here participates in two different hydrogen bond networks (see below). **R173$^{3.50}$** (Figure 5C) is part of the ionic lock in rhodopsin but interacts with D130$^{3.49}$ in β_2AR. In opsin, and in the 5-HT$_{2A}$ model, this residue interacts with Y223$^{5.58}$ and the C-terminal of the G protein. The bond to Y254$^{5.58}$ broke upon release of its constraint. This is likely to be an effect of the slight movement of the intracellular end of TMH5 away from TMH3, thought to make additional room for the G protein. However, no conclusions should be drawn from this observation as both the ICL3 and the G protein structures are incomplete.

Interhelical Bonds. In order to study the stability of existing and the formation of new interhelical bonds, these were not constrained in the last simulation (simulation 4) in the production of the 5-HT$_{2A}$ model. Table 2 lists the interhelical contacts and their prevalence in the simulations. One bond, R173$^{3.50}$-Y254$^{5.58}$, broke (see above), whereas four were preserved, indicating that, overall, the 5-HT$_{2A}$ simulations are compliant with the natural dynamics of an active GPCR. Previously known interhelical bonds were within hydrogen bonding distances

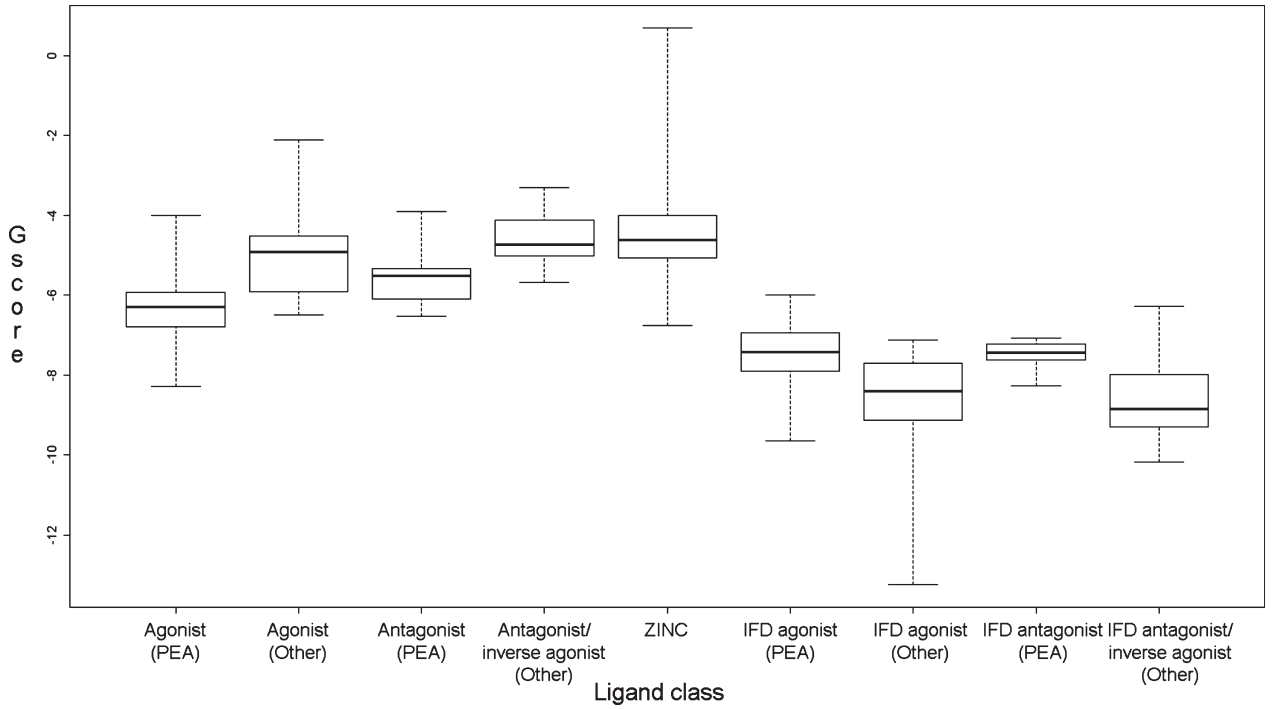


Figure 7. Box plot generated with R[40] of the distribution of Glide[29] G scores (lower value is better) of the different classes of compounds in the docking set. The four boxes to the right represent scores from induced-fit docking. The bands inside boxes are median values, and the tops and bottoms are the upper and lower quartiles, respectively. The horizontal bars at the ends of the dotted lines represent minimum and maximum of the scores.

already prior to the MD simulations and have therefore served to maintain rather than modify the relative positions of helices. In contrast, the new hydrogen bonds could not have been formed without the *in silico* activation.

H165$^{3.42}$ is responsible for a three-helix connection between TMH2, TMH3, and TMH4 (cf. Figure 6A) consisting of one known (to W200$^{4.50}$) and one new (to S115$^{2.45}$) hydrogen bond, which was formed in the first and preserved throughout the MD simulation. S115$^{2.45}$ and H165$^{3.42}$ are closer in the G protein-bound opsin structure than in that of rhodopsin (2.0 and 2.8 Å, respectively), but the angle between them does not allow for hydrogen bonding. Figure 6B displays a second three-helix hydrogen bonding network formed between TMH1, TMH2, and TMH7 in simulation 5b. N92$^{1.50}$ was known to interact with D120$^{2.50}$ and, after release of ligand constraints found to also bind to S373$^{7.46}$. The new hydrogen bond is supposed to stabilize TMH7, which was shifted toward TMH1 and TMH2, to give room for the movement of THM6 closer to the ligand binding pocket. During the fourth MD simulation a water molecule-mediated hydrogen bond network spontaneously formed between the residues; L166$^{3.43}$, S170$^{3.47}$, R173$^{3.50}$, Y254$^{5.58}$, Y328$^{6.40}$, N376$^{7.49}$, and Y380$^{7.53}$ was (cf. Figure 6C). Figure 6D depicts an alternative hydrogen bond network observed after the release of the ligand constraints. Here three of the residues, R173$^{3.50}$, Y254$^{5.58}$, and Y380$^{7.53}$, together with L325$^{6.37}$ interact directly i.e. without bridging waters.

Validation by Retrospective Ligand Screening. In order to validate the 5-HT$_{2A}$ model we screened a set of 182 known ligands and 9257 diverse compounds from the ZINC database.[31] The majority of the best scoring hits were phenylethylamine agonists, which constituted the top 29 of the ranked hits. The complete table of screening hits is found in SI5. Figure 7 shows box plots of the docking scores of each class of compounds in the test set. The best scoring class is, with large margin, the phenylethylamine agonists, and the phenethylamine antagonists score significantly higher than the other antagonists. The docking scores are approximately the same among the other molecule classes i.e. agonists, other antagonists/inverse agonists and ZINC compounds, although the variance increases in this order. As binding data are not provided by ZINC, it cannot be determined whether these represent known 5-HT$_{2A}$ ligands. To test whether rearrangements in the binding site would allow a better accommodation of nonphenyltehylamine agonists, we also performed induced-fit docking (the four rightmost boxes in Figure 7). Here, the docking scores improved for all four sets, and, interestingly, nonphenylethylamine agonists here score better than phenylethylamine agonists.

DISCUSSION

Active GPCR structures are highly needed for activation studies and ligand design, but modeling of active-state GPCRs has been, and still is, very challenging due to the limited number of related experimental structures. However, the recent structure of opsin complexed with a $G_{\alpha q}$ C-terminal peptide[13a] has provided the structure of the intracellular face of an active GPCR and revealed part of the G protein binding site and binding mode. Preliminary, agonist-bound adrenergic receptor structures have recently been announced[38,41] but not yet published. However, these display very small changes compared to the inactive structures, and it is uncertain whether they represent truly active receptor conformations. Mutagenesis of a number of ligand families has shown that the binding sites of many agonists have to tighten upon activation to fully accommodate their binding modes.[42] In our modeling of an active 5-HT$_{2A}$ conformation, we activated the receptor *in silico* by means of constrained ligand−receptor interactions. The resulting tightening of the ligand

binding site and opening of the G protein binding site are in agreement with mutational mapping of ligand binding,[42] biophysical studies of helical movements,[37] the "global toggle-switch model",[19] and the recent presentations of agonist-bound GPCR structures.[38,41] Furthermore, the system proved stable when releasing ligand and G protein constraints, thereby indicating that our 5-HT_{2A} model represents a plausible active GPCR conformation.

As shown in Figure 1, the binding mode of the phenylethylamine scaffold is well characterized with exception of the structural rationale behind the contribution to affinity for some N-benzyl-substituted analogs. A (F339L$^{6.51}$) mutation has shown that the phenylalanine, F339$^{6.51}$, may provide aromatic stabilization of N-benzyls analogs. However, binding affinities show that N-benzyl derivates that have a hydrogen bond acceptor (e.g., 25I-NBMD) have much higher affinities than those that do not (e.g., 25I-NB),[26] indicating that an uncharacterized hydrogen bonding interaction is missing. An analysis of the binding site combined with data from mutation studies identified a tyrosine in TMH7, Y370$^{7.43}$, as the most likely hydrogen bonding partner for the N-benzyl moiety. This residue has been implicated to have a bigger role in ligand-mediated activation, also for other GPCRs. In the inactive β_1-[14] and β_2-[15]adrenergic receptor structures, Y370$^{7.43}$ interacts with D155$^{3.32}$, something which seems possible in all aminergic receptors as they all have an aspartate in position 3.32 and a tyrosine or tryptophan in position 7.43, including 5-HT_{2A}. Furthermore, 22 additional Family A receptors (in total 20% of the members) share the combination of an aspartate or glutamate in position 3.32 and a tyrosine or tryptophan in position 7.43.

As this bond stabilizes the inactive conformation, the agonist-mediated breakage may be a key to receptor activation. It has been shown that the Y370A$^{7.43}$ mutation results in loss of affinity and efficacy in 5-HT_{2A}, but this effect is general for agonists and is thus likely to result from structural changes in the receptor.[43] We propose that a more conservative mutation, such as Y370F$^{7.43}$, would be more informative, especially in the testing of the hypothesis of Y370$^{7.43}$ as hydrogen bonding partner of 25I-NBMD and related compounds. The Y370F$^{7.43}$ mutation has already proven milder compared to Y370A$^{7.43}$ in the homologous muscarinic M_1 acetylcholine[44] and 5-HT_4 receptors.[45]

The MD simulations generated new information about the multitude of residues and noncovalent contacts that may play a role in stabilization of active GPCR conformations. In particular, two residues, H165$^{3.42}$ and S373$^{7.46}$, are interesting (Figure 6A-B). The first residue, H165$^{3.42}$, already had a known contact to W200$^{4.50}$ and rapidly formed a new hydrogen bond in the first simulation to S115$^{2.45}$. S115$^{2.45}$ and H165$^{3.42}$ are closer in the G protein-bound opsin structure than in the rhodopsin structure (2.8 Å), but the angle between them does not allow for hydrogen bonding. Sequence analysis shows that position 3.42 is occupied by a hydrogen bond acceptor in 48% of Family A GPCRs and 97% of the aminergic receptors. The second residue, S373$^{7.46}$, formed an interaction, after release of ligand constraints, to N92$^{1.50}$, which was already bonded to D120$^{2.50}$. Sequence analysis shows that position 7.46 is occupied by a hydrogen bond donor in 73% of Family A GPCRs and 100% of the aminergic receptors. These two residues, H165$^{3.42}$ and S373$^{7.46}$, and their three-helix networks are located slightly below and above the middle, respectively, of the receptor. The lower area is located two helical turns below the ligand binding site, an area which mediates signal transduction to the G protein. S373$^{7.46}$ is located only one helical turn above the conserved proline residue, constituting a hinge in TMH7. Thus, the S373$^{7.46}$ residue may be involved in the stabilization of TMH7, which moved to allow THM6 to close in on the ligand binding pocket.

The MD simulations also produced additional information about known rotameric microswitches, and in particular Y380$^{7.53}$, which displayed great versatility in various modes of stabilizing interactions. Comparing the inactive β_2AR[15] and active opsin structures,[13a] Y380$^{7.53}$ moved toward TMH6 and so stabilizes this helix in its outward shifted position making room for the G protein. In the 5-HT_{2A} model, Y380$^{7.53}$ has moved even closer to TMH6 as compared to the opsin structure. Moreover, this residue participates in two different hydrogen bond networks. In the active 5-HT_{2A} model the network is water-mediated involving four water molecules (Figure 6C), but as the constraints on the ligand were released (simulation 5b) a largely overlapping set of receptor residues formed a hydrogen bonding network directly between the protein residues (Figure 6D). These two alternative networks contribute to the stability of the active receptor states by keeping R173$^{3.50}$ close to the G protein, blocking the inward movement of TMH6 and locking the relative positions of TMH3, TMH5, and TMH7.

The G protein peptide backbone moved only slightly during the MD simulations. However, as many as eight new hydrogen bonds were formed during the four simulations leading to the active 5-HT_{2A} model. Furthermore, several residues displayed alternative variants of hydrogen bonds when releasing the G protein constraints. Most striking, the hydrogen bond with helix 8 changed direction and residues as well as increased from one to two hydrogen bonds, when replacing K354-Q312$^{8.49}$ with E355-N384$^{8.47}$ and E355-K385$^{8.48}$. As the new bonds were formed without constraints, they constitute likely alternatives to the original K354-Q312$^{8.49}$ bond modeled prior to the MD simulations. These bonds were considered in the homology model but could only get close enough after the rearrangements of helices observed during the MD simulations. Given that we have identified at least two possibilities for interactions between the G protein and helix 8, and similar interactions are found in the opsin crystal structure, it is likely that helix 8 is an important anchor point for the G protein in the active state of the 5-HT_{2A} receptor. Furthermore, the G protein displayed interactions with TMH6, which together with ICL3, are also considered as generic anchoring points. However, in this region, G protein interactions will vary among receptors, to achieve the subtype specificity and reflect the diversity of loop lengths and sequences.

In the retrospective ligand screening of more than 9400 compounds, the best 29 hits were phenylethylamines (SI5), showing that the model is suitable for virtual screening and structure-based design of 5-HT_{2A} phenylethylamine agonists. Furthermore, the screening results showed similar scores for other agonists, antagonists/inverse agonists, and ZINC compounds. Ergolines, and many tryptamines, are larger than phenylethylamines and therefore do not fit into the binding site. These compounds are also structurally different and have different binding modes. This raises the question if a single GPCR model can be used to accurately dock structurally different ligands. As the agonist classes bind in approximately the same site and, although varying somewhat in size, are all small molecules, it is likely that rearrangements in side-chain configurations without further MD simulations would be adequate for the model to accommodate nonphenylethylamine agonists. We tested this by performing induced-fit docking of the 5-HT_{2A} ligands. In this docking, the scores improved for all three sets of compounds and in

particular for nonphenylethylamine agonists, which scored best. This would imply that backbone movements are not needed, avoiding a full *in silico* activation process for each ligand class, although our steered MD approach could easily be adapted to other agonists, antagonists, or inverse agonists as well as other receptors.

Another question of high interest for ligand design is which features can be identified in the binding site that may provide interactions specific for the 5-HT_{2A} receptor. Initially we expected that unique features for 5-HT_{2A} would be found primarily in the N-benzyl-binding region. However, comparison of the amino acid residues in the 5-$HT_{2A\text{-}C}$ receptors showed that this is not the case. Only one residue has been identified in 5-HT_{2A} that has the potential to provide specific interactions. This is a serine residue on TMH5, $S159^{5.46}$. Human 5-HT_{2B} and 5-HT_{2C} have an alanine in this position. This difference could be used in design of selective ligands, for example by forming hydrogen bonds to 5-HT_{2A}. Furthermore, analysis of the mouse and rat 5-HT_{2A} sequences shows that these also have an alanine in this position. This suggests that ligands designed to interact with the serine have to be assayed against the human receptor.

To summarize, the MD simulations generated new information on residues and nonbonded contacts associated with receptor activation. The model may be used for further activation studies or design of 5-HT_{2A} agonists, in which case we suggest additional longer MD simulations and analysis of interesting rotamer states, statistically. However, our intention is to use the 5-HT_{2A} model for agonist design, in the first case phenylethylamines, for which the model has been validated by retrospective virtual ligand screening, and subsequently for other agonist classes after optimization of the binding site. Finally, this study is a thorough example of the steered MD approach for *in silico* GPCR activation, which too, may inspire similar analyzes.

■ ASSOCIATED CONTENT

Supporting Information. **SI1**: Sequence alignment of β_2AR and 5-HT_{2A}, in pir format, used as input to Modeler for building of the 5-HT_{2A} homology model. **SI2**: A modified version of the default Desmond membrane protein relaxation protocol. **SI3**: The python script used to add distance constraints as "pseudo-bonds" to the Desmond structure file. **SI4**: PDB structure file of the model produced in this study of active 5-HT_{2A}. **SI5**: A table of all hits from the ligand screening of the 5-HT_{2A} model. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +46703148390. Fax: +4535 33 60 40. E-mail: davidgloriam@gmail.com.

■ ACKNOWLEDGMENT

The authors would like to thank Professor David Nichols for valuable input. D.E.G. acknowledges financial support from the Alfred Benzon Foundation. T.B. was supported by grants from the Lundbeck Foundation.

■ ABBREVIATIONS

25Br-NBMD, N-(2,3-methylenedioxybenzyl)-4-bromo-2,5-dimethoxyphenethylamine; 25I-NB, N-benzyl-4-iodo-2,5-dimethoxyphenethylamine; 25I-NBMD, N-(2,3-methylenedioxybenzyl)-4-iodo-2,5-dimethoxyphenethylamine; 5-HT, 5-hydroxytryptamine (serotonin); 5-HT2A, 5-hydroxytryptamine receptor 2A; ACTH, adrenocorticotropic hormone; β_2AR, β_2 adrenergic receptor; DPPC, dipalmitoylphosphatidylcholine; fs, femtosecond; $G\alpha_i$, α-subunit of a heterotrimeric G_i protein; $G\alpha_q$, α-subunit of a heterotrimeric G_q protein; $G\alpha_s$, α-subunit of a heterotrimeric G_s protein; GPCR, G protein-coupled receptor; ICL, intracellular loop; K, kelvin; LSD, lysergic acid diethylamide; MeO, methoxy; MD, molecular dynamics; Ns, nanosecond; OPLS-AA, optimized potentials for liquid simulations (all atom); PDB, protein data bank (www.pdb.org); ps, picosecond; RMSD, root mean square deviation; TMH, transmembrane helix; Å, ångström

■ REFERENCES

(1) Fredriksson, R.; Lagerström, M. C.; Lundin, L.-G.; Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003**, *63* (6), 1256–1272.

(2) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y. H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigo, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y. H.; Coyne, M.; Dahlke, C.; Mays, A.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.;

Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. The sequence of the human genome. *Science* **2001**, *291* (5507), 1304–1351.

(3) Bockaert, J.; Pin, J. P. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J.* **1999**, *18* (7), 1723–1729.

(4) (a) Drews, J. Drug Discovery: A Historical Perspective. *Science* **2000**, *287* (5460), 1960–1964. (b) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1* (9), 727–730.

(5) Schlyer, S.; Horuk, R. I want a new drug: G-protein-coupled receptors in drug development. *Drug Discovery Today* **2006**, *11* (11−12), 481–493.

(6) Nichols, D. E.; Nichols, C. D. Serotonin receptors. *Chem. Rev.* **2008**, *108* (5), 1614–1641.

(7) Hoyer, D.; Hannon, J. P.; Martin, G. R. Molecular, pharmacological and functional diversity of 5-HT receptors. *Pharmacol., Biochem. Behavior* **2002**, *71* (4), 533–554.

(8) Nichols, D. E. Hallucinogens. *Pharmacol. Ther.* **2004**, *101* (2), 131–181.

(9) Ray, T. S. Psychedelics and the human receptorome. *PloS One* **2010**, *5* (2), e9019–e9019.

(10) Shulgin, A.; Shulgin, A. Shuras voice. In *PiHKAL: A chemical love story*, 1st ed.; Transform Press: Berkeley, CA, 1991; pp 15−18.

(11) Blaazer, A. R.; Smid, P.; Kruse, C. G. Structure-activity relationships of phenylalkylamines as agonist ligands for 5-HT(2A) receptors. *ChemMedChem* **2008**, *3* (9), 1299–1309.

(12) Kobilka, B. K. G protein coupled receptor structure and activation. *Biochim. Biophys. Acta* **2007**, *1768* (4), 794–807.

(13) (a) Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krauss, N.; Choe, H.-W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008**, *455* (7212), 497–502. (b) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le, T. I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289* (5480), 739–745.

(14) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454* (7203), 486–491.

(15) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318* (5854), 1258–1265.

(16) Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322* (5905), 1211–1217.

(17) Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science* **2010**, *330* (6007), 1066–1071.

(18) Chien, E. Y. T.; Liu, W.; Zhao, Q.; Katritch, V.; Hanson, G.; Shi, M. A.; Newman, L.; Javitch, A. H.; Cherezov, J. A.; Stevens, V. R. C., Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. *Science* **2010**, *330* (6007), 1091–1095.

(19) Schwartz, T. W.; Frimurer, T. M.; Holst, B.; Rosenkilde, M. M.; Elling, C. E. Molecular mechanism of 7TM receptor activation--a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* **2006**, *46*, 481–519.

(20) Nygaard, R.; Frimurer, T. M.; Holst, B.; Rosenkilde, M. M.; Schwartz, T. W. Ligand binding and micro-switches in 7TM receptor structures. *Trends Pharmacol. Sci.* **2009**, *30* (5), 249–259.

(21) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.-Y.; Pieper, U.; Sali, A., Comparative protein structure modeling using MODELLER. *Current protocols in protein science/editorial board*; John E. Coligan ... [et al.] **2007**; Chapter 2, Unit 2.9-Unit 2.9.

(22) Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossváry, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. In *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters*; Tampa, FL, 2006; Tampa, Florida.

(23) Braden, M. R. *Towards a Biophysical Basis of Hallucinogen Action*; 2007.

(24) Braden, M. R.; Nichols, D. E. Assessment of the roles of serines 5.43(239) and 5.46(242) for binding and potency of agonist ligands at the human serotonin 5-HT2A receptor. *Mol. Pharmacol.* **2007**, *72* (5), 1200–1209.

(25) Chambers, J. J.; Nichols, D. E. A homology-based model of the human 5-HT2A receptor derived from an in silico activated G-protein coupled receptor. *J. Comput.-Aided Mol. Des.* **2002**, *16* (7), 511–520.

(26) Braden, M. R.; Parrish, J. C.; Naylor, J. C.; Nichols, D. E. Molecular interaction of serotonin 5-HT2A receptor residues Phe339-(6.51) and Phe340(6.52) with superpotent N-benzyl phenethylamine agonists. *Mol. Pharmacol.* **2006**, *70* (6), 1956–1964.

(27) Choudhary, M. S.; Craigo, S.; Roth, B. L. A single point mutation (Phe340-->Leu340) of a conserved phenylalanine abolishes 4-[125I]iodo-(2,5-dimethoxy)phenylisopropylamine and [3H]mesulergine but not [3H]ketanserin binding to 5-hydroxytryptamine2 receptors. *Mol. Pharmacol.* **1993**, *43* (5), 755–761.

(28) Cordes, F. S.; Bright, J. N.; Sansom, M. S. Proline-induced distortions of transmembrane helices. *J. Mol. Biol.* **2002**, *323* (5), 951–960.

(29) *Glide, 5.5*; Schrödinger LCC: New York, 2009.

(30) *Maestro, 9.0*; Schrödinger LCC: New York, 2009.

(31) Irwin, J. J.; Shoichet, B. K. ZINC--a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177–182.

(32) (a) *Epik, 2.0*; Schrödinger LCC: New York, 2009.(b) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21* (12), 681–691.

(33) *Prime, 2.1*; Schrödinger LCC: New York, 2009.

(34) *ChemBioDraw Ultra, 12*; Cambridgesoft: Cambridge, 2010.

(35) *The PyMOL Molecular Graphics System, 1.2r1*; Schrödinger LCC: New York, 2009.

(36) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366–428.

(37) (a) Gether, U.; Lin, S.; Kobilka, B. K. Fluorescent labeling of purified beta 2 adrenergic receptor. Evidence for ligand-specific conformational changes. *J. Biol. Chem.* **1995**, *270* (47), 28268–28275. (b) Devanathan, S.; Yao, Z.; Salamon, Z.; Kobilka, B.; Tollin, G. Plasmon-waveguide resonance studies of ligand binding to the human beta 2-adrenergic receptor. *Biochemistry* **2004**, *43* (11), 3280–3288.

(38) Rasmussen, S. G.; Choi, H. J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; Devree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunhara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **2011**, *469* (7329), 175–180.

(39) Seggel, M. R.; Yousif, M. Y.; Lyon, R. A.; Titeler, M.; Roth, B. L.; Suba, E. A.; Glennon, R. A. A structure-affinity study of the binding of 4-substituted analogues of 1-(2,5-dimethoxyphenyl)-2-aminopropane at 5-HT2 serotonin receptors. *J. Med. Chem.* **1990**, *33* (3), 1032–1036.

(40) Team, R. D. C. *R: A Language and Environment for Statistical Computing, 2.11.1*; R Foundation for Statistical Computing: Vienna, 2010.

(41) (a) Marshall, F. H. In *A new approach to drug discovery utilising stabilised GPCRs*, G Protein-Coupled Receptors in Drug Discovery, Berlin, Germany, 17 March; Informa Life Sciences: Berlin, Germany, 2010. (b) Mason, J. In *Enabling Structure Based Compound Design of GPCRs Through Biophysical Characterization*, GPCR Congress 2010, Montreux, Switzerland, 10 June; Montreux, Switzerland, 2010.

(42) Kristiansen, K. Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function 2. *Pharmacol Ther.* **2004**, *103* (1), 21–80.

(43) Roth, B. L.; Shoham, M.; Choudhary, M. S.; Khan, N. Identification of conserved aromatic residues essential for agonist binding and second messenger production at 5-hydroxytryptamine2A receptors. *Mol. Pharmacol.* **1997**, *52* (2), 259–266.

(44) Lu, Z. L.; Saldanha, J. W.; Hulme, E. C. Transmembrane domains 4 and 7 of the M(1) muscarinic acetylcholine receptor are critical for ligand binding and the receptor activation switch. *J. Biol. Chem.* **2001**, *276* (36), 34098–34104.

(45) Rivail, L.; Giner, M.; Gastineau, M.; Berthouze, M.; Soulier, J.-L.; Fischmeister, R.; Lezoualc'h, F.; Maigret, B.; Sicsic, S.; Berque-Bestel, I. New insights into the human 5-HT4 receptor binding site: exploration of a hydrophobic pocket. *Br. J. Pharmacol.* **2004**, *143* (3), 361–370.