Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Mechanistic Studies on the Stereoselectivity of the Serotonin 5-HT$_{1A}$ Receptor

Shuguang Yuan[+],* *Qian Peng*[+], *Krzysztof Palczewski, Horst Vogel, and Slawomir Filipek**

anie_201603766_sm_miscellaneous_information.pdf

Supporting information

1. Methods
2. Figure S1-S10

Methods

Homology model building. Homology models of 5-HT_{1A} were obtained by Modeller 9v15 using the crystal structures of the agonist bound 5-HT_{1B} receptor (pdb: 4IAQ)[1] and the antagonist bound M3 muscarinic receptor (pdb: 4U15)[2] as templates. The initial 3D sequence alignments between 5-HT_{1A} and template structures were done with Strap[3]. All highly conserved residues and motifs in each TM helix were manually adjusted for proper alignment. A total of 20,000 models were generated in Modeller with a simulated annealed protocol in a multiple template approach, and the optimal model was chosen according to its DOPE (Discrete Optimized Protein Energy) score. Over 5000 loops were generated for the ECL2 loop refinement and the lowest DOPE scored model was kept as output.

The output structural model of the 5-HT_{1A} receptor superimposes very well with the template structures of the 5-HT_{1B} and the M3 receptor (Figure S1) with RMSD of the TM backbones <1.5 Å. As $W^{6.48}$ adopts different rotamer states in the 5-HT_{1B} and M3 receptor template structures we compared the side chain conformations of the highly conserved $W^{6.48}$ within all available GPCR crystal structures (Figure S2A). In almost all cases the long axis of the indole ring orients preferentially parallel to the TM helices. The only exception is found in the M3 crystal structure, where the long axis of the indole ring of $W^{6.48}$ is oriented perpendicular to the TM helices. In our homology model of the 5-HT_{1A} receptor we have adjusted the side chain conformation of, $W^{6.48}$ to that found in most GPCR structures (Figure S2B), and we retained this particular orientation to start all our MD simulations and docking experiments.

Ligand docking. Ligand structures were obtained from the PUBCHEM[4] online database. Ligand preparation utility in the Schrödinger 2013 suite was used to optimize the geometry of each initial structure with OPLS_2005 force field.[5] Ionization states of ligands were calculated with the Epik[6] tool using Hammett and Taft methods together with ionization and tautomerization tools. The nitrogen atom in both ligands was protonated.

The 5-HT_{1A} receptor was prepared with the Protein Preparation Tool (ProPrep) in Schrödinger 2013 suite software. Asn, Gln, and His residues were checked for flips automatically in ProPrep. Hydrogen atoms were added to the 5-HT_{1A} model according to the physiological pH environment with the PROPKA[7] tool in Maestro[8] along with an optimized hydrogen bond network.

The docking procedure was performed with Glide[9, 10] (Schrödinger 2013 suite). Each ligand molecule was initially placed in the binding pocket in a pose similar to that observed in the corresponding crystal structure. Cubic boxes centered on the ligand mass center with a radius of 12 Å defined the docking binding regions before flexible ligand docking was executed. Five poses out of 20,000 per ligand were included in post-docking energy minimization. Since the structures of the ligands were quite rigid and the top scored 5 poses were quite similar to each other (RMSD<1 Å), we chose the best scored pose as the initial structure for MD simulations.

Molecular dynamics simulations. The receptor containing membrane system was built using the g_membed[11] tool in Gromacs with the receptor structure pre-aligned in the OPM (Orientations of Proteins in Membranes) database[12, 13]. Pre-equilibrated 140 POPC lipids coupled with 10,200 TIP3P water molecules in a periodic box of 72 Å x 72 Å x 100 Å were used to build the protein/membrane system. Proteins, lipids, water molecules and ions were parameterized with the CHARMM36 force field[14] parameter set, whereas the CHARMM CGenFF small molecule force field[15] was used for the ligands. All ligands were submitted to the GAUSSIAN 09 program[16] for structure optimization at the B3LYP/6-31G* level prior to force field parameter generation. All bond lengths to hydrogen atoms in each protein/membrane system were constrained with M-SHAKE[17]. Van der Waals and short-range electrostatic interactions were cut off at 10 Å. Results obtained from the MD simulations were analyzed in Gromacs[18] and VMD[19]. Each case was repeated three times starting from the same structure but various random seeds for initial velocities.

Helix bending calculations. The bending of a helix with respect to the channel axis was calculated in Bendix[20] with default settings. Bendix uses a sliding window of four residues to give local helix axes that are joined by a spline.[21] Analyses are performed on this axis to provide local angles along the length of the helix versus time. This helix distortion distribution is imaged across residues by heatmap color-coding according to the local angle magnitudes, that highlights the non-linear helix behavior.

Protein ligand interaction fingerprint calculation. The interaction fingerprint between protein and ligand was done with IChem[22]. We first extracted 100 snapshots from the final 50 ns MD simulation, and then submitted them to IChem. IChem converts protein−ligand coordinates into a fingerprint (TIFP) of 210 integers registering the corresponding molecular interaction pattern. TIFP fingerprints have been calculated for ca. 1000 protein−ligand complexes, enabling a broad comparison of relationships between interaction pattern similarities and ligand or binding site pairwise similarities. The ligands of 5-HT_{1A} in this work had only two different types of interactions: an ionic interaction and a hydrophobic contact. We defined the calculation zone within 6 Å of the ligand mass center and kept the default parameters of IChem in the calculations.

Interaction network between amino acid residues. The interaction network between amino acid residues of the receptor was obtained with RINerator[23], an application in Cytoscape[24]. RINerator allows users to select amino acids represented by nodes within a protein to determine the intrinsic interaction network and interactively display and analyze the 3D protein structures in StructureViz[25] and UCSF Chimera[26]. Such representations visualize complicated interaction patterns, not easily seen in the protein structure directly, and also their temporal changes. An average structure,

which was calculated by g_covar tool in Gromacs based on the final 50 ns period of MD simulations, was introduced for RINerator analysis. The structure with explicit hydrogen atoms served as input for the program Probe[27]. Probe identifies the contacts between amino acids in a protein by evaluating their atomic packing in small-probe contact dot surfaces. A small virtual probe (typically 0.25 Å) is rolled around the van der Waals surface of each atom and an interaction (contact dot) is detected if the probe touches another non-covalently bonded atom. The Probe program summarizes these scoring data for all parts of a protein structure and provides an output file with information for every amino acid residue. The last step of this network generation method is accomplished by the Python package RINerator that creates an undirected weighted network with multiple edges. The nodes represent the amino acid residues of the protein, whereas the links between them represent the non-covalent interactions identified by Probe. The final residues' interaction network is depicted in a circular layout style.

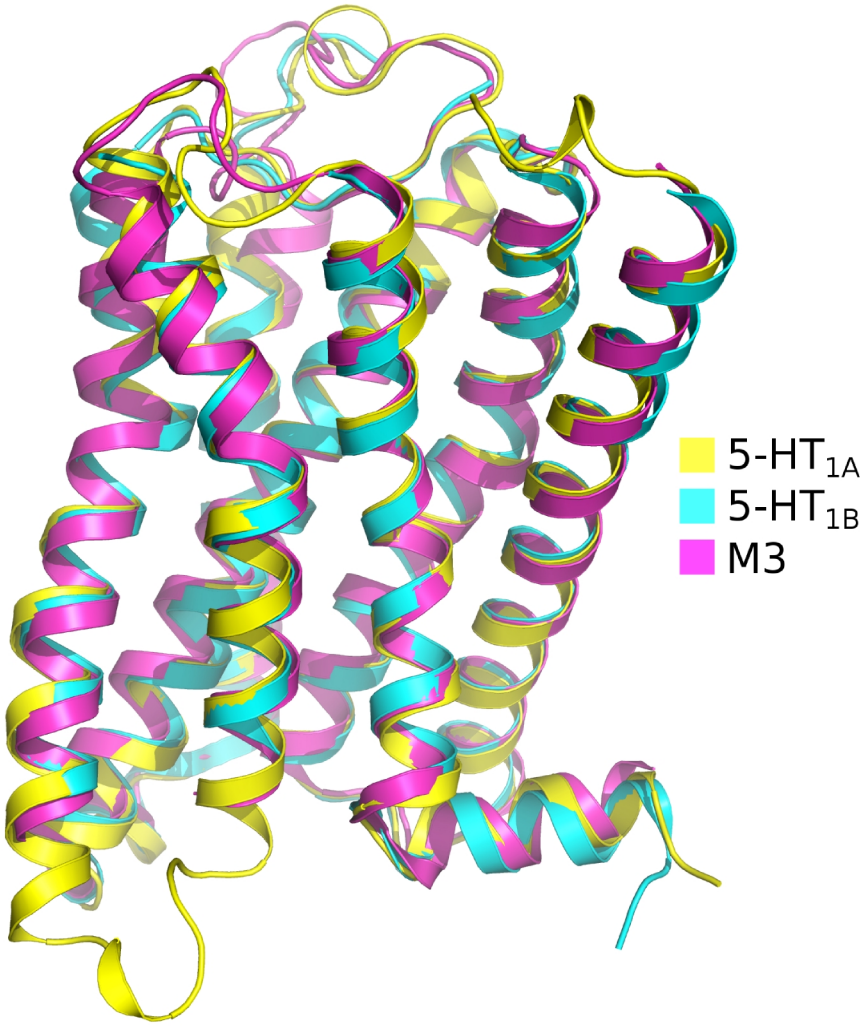

Figure S1. The superimposed structures of the 5-HT_{1A}, 5-HT_{1B} and M3 receptors represent the receptors in an inactive state.

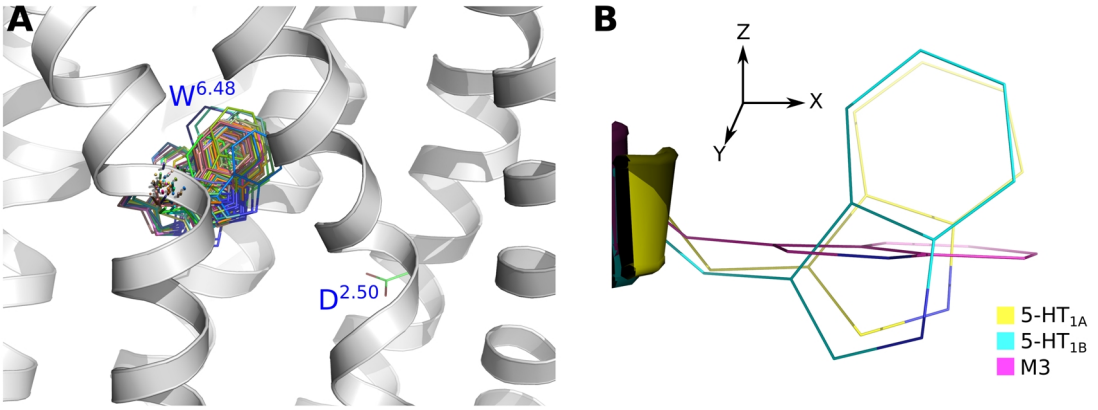

Figure S2. Rotamer conformations of the highly conserved $W^{6.48}$. (A) Superposition of all GPCR crystal structures. (B) Particular rotamer conformation of $W^{6.48}$ in the 5-HT_{1A}, the 5-HT_{1B} and the M3 receptor with respect to the TM6 helix (green).

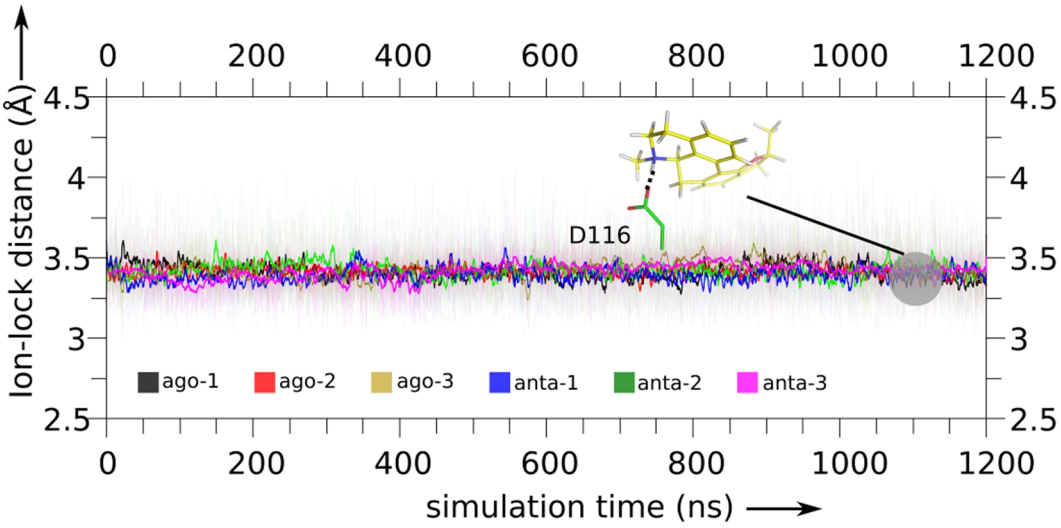

Figure S3. A stable salt bridge distance between the carboxylic group in the side chain of D116$^{3.38}$ and a protonated nitrogen from a ligand. Black and red: two trajectories for agonist-bound 5-HT$_{1A}$ receptor. Blue and green: two trajectories for antagonist-bound 5-HT$_{1A}$ receptor.

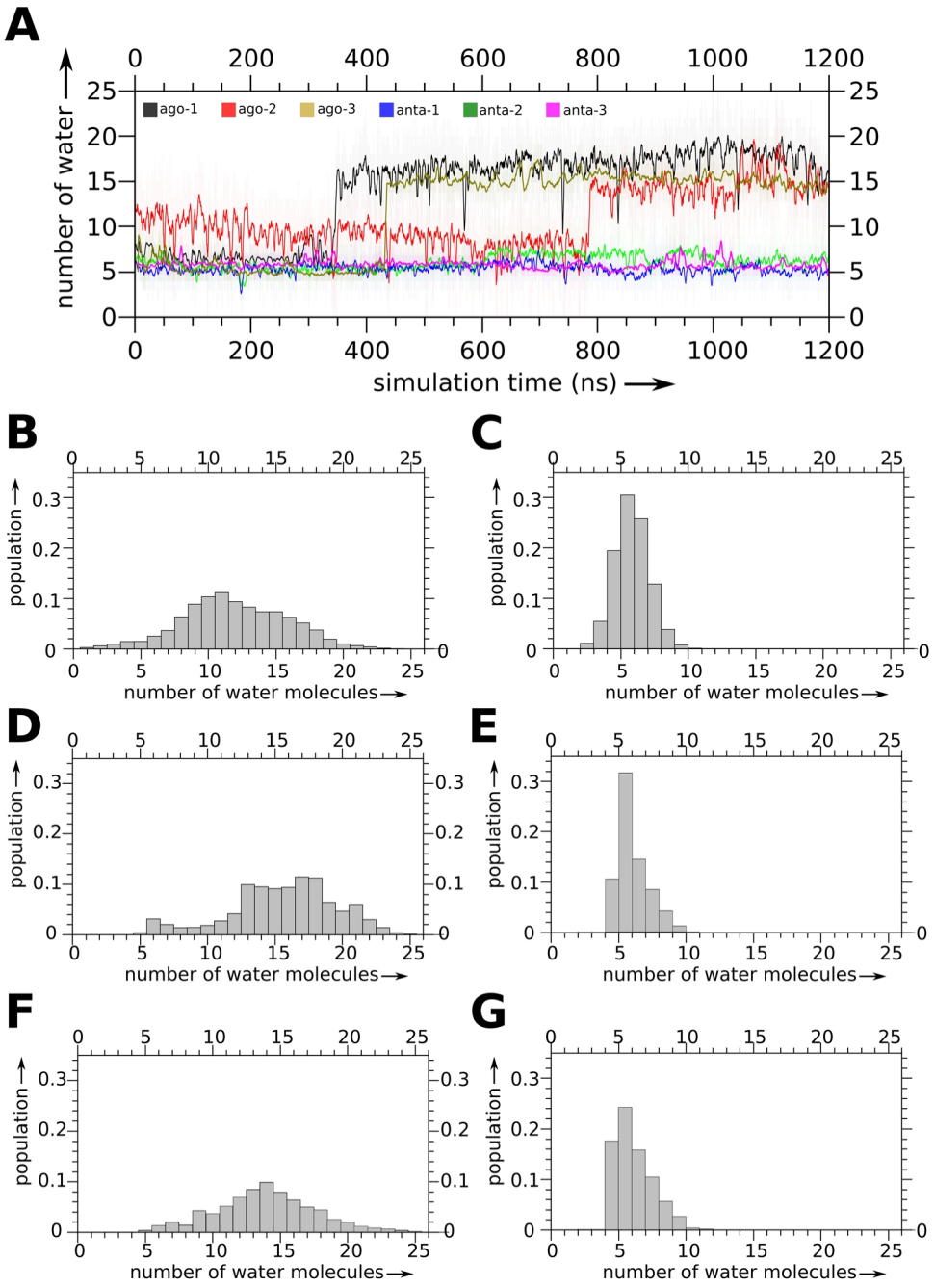

Figure S4. (A) Number of water molecules within a 4 Å distance of D82$^{2.50}$ located in the allosteric site of the receptor for agonist- (ago-1 & ago-2) and antagonist- (anta-1 & anta-2) bound 5-HT$_{1A}$ receptor. Average number of water molecules <N> and its population shown for (B, D & F) agonist- and (C, E & G) antagonist-bound complexes during final 50 ns. During the initial 380 ns ago-1 and 800 ns ago-2 simulation period, <N> was 8±2 and 6±2, respectively. From 380 ns in ago-1 and from 800 ns in ago-2, the <N> of agonist-bound 5-HT$_{1A}$ was stabilized at 15±2. For the antagonist bound cases (anta-1 & anta-2), <N> revealed very small fluctuations during the total simulation, 6±2 for each trajectory.

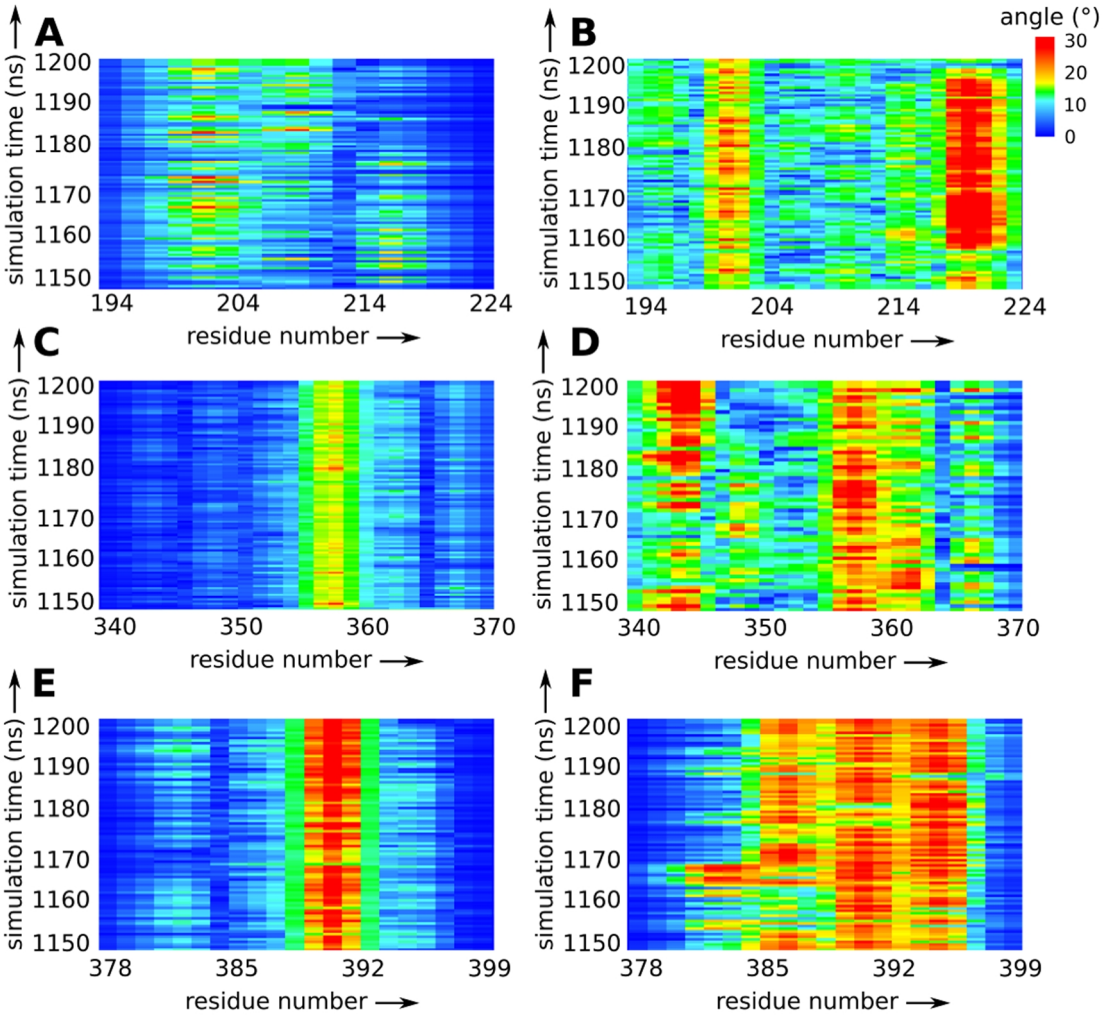

Figure S5. Bending of TMs in the 5-HT_{1A} receptor during the final 50 ns of MD simulations. (A) TM5 bending of antagonist-bound receptor. (B) TM5 bending of agonist-bound receptor. (C) TM6 bending of antagonist-bound receptor. (D) TM6 bending of agonist-bound receptor. (E) TM7 bending of antagonist-bound receptor. (F) TM7 bending of agonist-bound receptor.

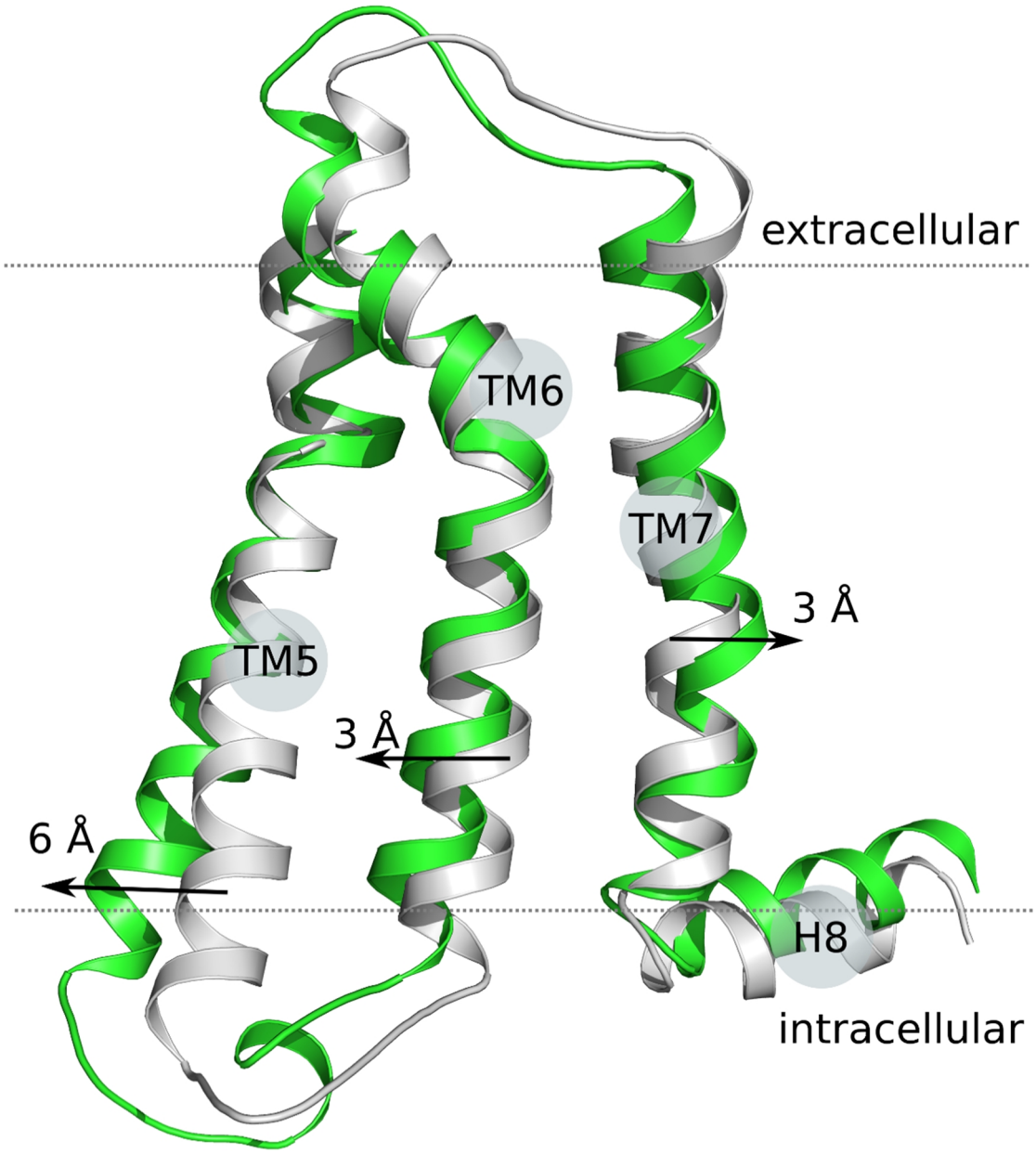

Figure S6. Movement of TM helices in the 5-HT_{1A} receptor during receptor activation by agonist binding. Gray: final snapshot from antagonist bound 5-HT_{1A} receptor in MD simulations. Green: final snapshot from agonist bound 5-HT_{1A} receptor in MD simulations.

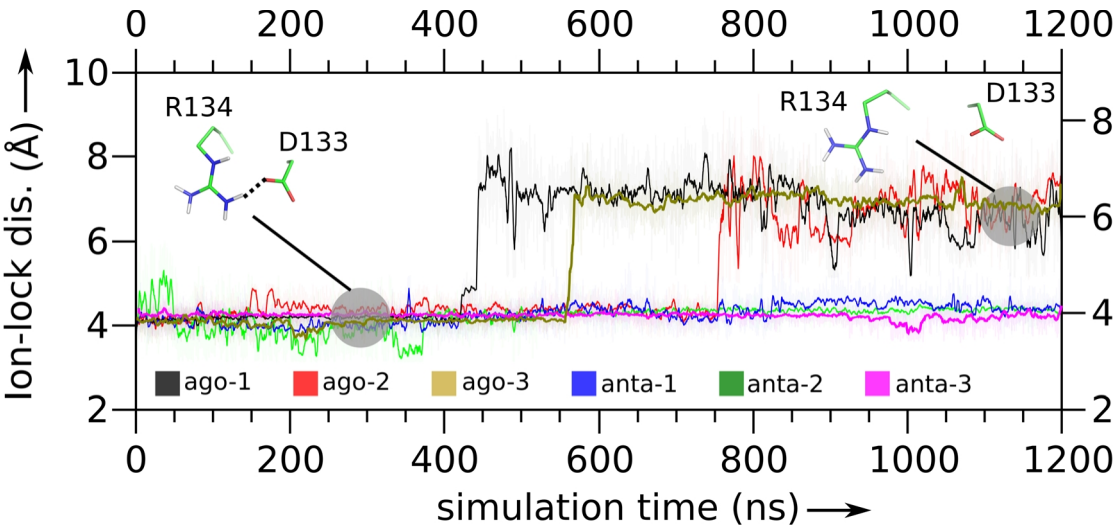

Figure S7. Temporal changes in the salt bridge distance between D133$^{3.49}$ and R134$^{3.50}$ in the DRY motif of the 5-HT$_{1A}$ receptor. Break of the salt bridge can induce rotation of the helices in the subsequent stages of activation.

References

[1] C. Wang, Y. Jiang, J. M. Ma, H. X. Wu, D. Wacker, V. Katritch, G. W. Han, W. Liu, X. P. Huang, E. Vardy, J. D. McCorvy, X. Gao, X. E. Zhou, K. Melcher, C. H. Zhang, F. Bai, H. Y. Yang, L. L. Yang, H. L. Jiang, B. L. Roth, V. Cherezov, R. C. Stevens, H. E. Xu, *Science* **2013**, *340*, 610.

[2] T. S. Thorsen, R. Matt, W. I. Weis, B. K. Kobilka, *Structure* **2014**, *22*, 1657.

[3] C. Gille, M. Fahling, B. Weyand, T. Wieland, A. Gille, *Nucleic Acids Res* **2014**, *42*, W3.

[4] Y. L. Wang, J. W. Xiao, T. O. Suzek, J. Zhang, J. Y. Wang, Z. G. Zhou, L. Y. Han, K. Karapetyan, S. Dracheva, B. A. Shoemaker, E. Bolton, A. Gindulyte, S. H. Bryant, *Nucleic Acids Research* **2012**, *40*, D400.

[5] D. Shivakumar, J. Williams, Y. J. Wu, W. Damm, J. Shelley, W. Sherman, *Journal of Chemical Theory and Computation* **2010**, *6*, 1509.

[6] J. R. Greenwood, D. Calkins, A. P. Sullivan, J. C. Shelley, *J Comput Aided Mol Des* **2010**, *24*, 591.

[7] C. R. Sondergaard, M. H. M. Olsson, M. Rostkowski, J. H. Jensen, *Journal of Chemical Theory and Computation* **2011**, *7*, 2284.

[8] Schrödinger, LLC, **2013**.

[9] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J Med Chem* **2004**, *47*, 1739.

[10] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J Med Chem* **2004**, *47*, 1750.

[11] M. G. Wolf, M. Hoefling, C. Aponte-Santamaria, H. Grubmuller, G. Groenhof, *J Comput Chem* **2010**, *31*, 2169.

[12] A. L. Lomize, I. D. Pogozheva, H. I. Mosberg, *J Chem Inf Model* **2011**, *51*, 918.

[13] A. L. Lomize, I. D. Pogozheva, H. I. Mosberg, *J Chem Inf Model* **2011**, *51*, 930.

[14] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, R. W. Pastor, *Journal of Physical Chemistry B* **2010**, *114*, 7830.

[15] K. Vanommeslaeghe, E. P. Raman, A. D. MacKerell, Jr., *J Chem Inf Model* **2012**, *52*, 3155.

[16] M. J. Frisch, et. al, Gaussian, Inc., Wallingford, CT, USA, **2009**.

[17] V. Krautler, W. F. Van Gunsteren, P. H. Hunenberger, *Journal of Computational Chemistry* **2001**, *22*, 501.

[18] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **2013**, *29*, 845.

[19] W. Humphrey, A. Dalke, K. Schulten, *Journal of Molecular Graphics & Modelling* **1996**, *14*, 33.

[20] A. C. E. Dahl, M. Chavent, M. S. P. Sansom, *Bioinformatics* **2012**, *28*, 2193.

[21] J. A. Dalton, I. Michalopoulos, D. R. Westhead, *Bioinformatics* **2003**, *19*, 1298.

[22] G. Marcou, D. Rognan, *J Chem Inf Model* **2007**, *47*, 195.

[23] J. Walstab, G. Rappold, B. Niesler, *Pharmacol Ther* **2010**, *128*, 146.

[24] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, *Genome Res* **2003**, *13*, 2498.

[25] J. H. Morris, C. C. Huang, P. C. Babbitt, T. E. Ferrin, *Bioinformatics* **2007**, *23*, 2345.

[26] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J Comput Chem* **2004**, *25*, 1605.

[27] S. Yuan, K. Palczewski, Q. Peng, M. Kolinski, H. Vogel, S. Filipek, *Angew Chem Int Ed Engl* **2015**, *54*, 7560.