doi.org/10.26434/chemrxiv.8216060.v1

Multi-Target Dopamine D3 Receptor Modulators: Actionable Knowledge for Drug Design from Molecular Dynamics and Machine Learning

Mariarosaria Ferraro, Sergio Decherchi, Alessio De Simone, Maurizio Recanatini, Andrea Cavalli, Giovanni Bottegoni

Submitted date: 01/06/2019 • Posted date: 03/06/2019

Citation information: Ferraro, Mariarosaria; Decherchi, Sergio; De Simone, Alessio; Recanatini, Maurizio; Cavalli, Andrea; Bottegoni, Giovanni (2019): Multi-Target Dopamine D3 Receptor Modulators: Actionable Knowledge for Drug Design from Molecular Dynamics and Machine Learning. ChemRxiv. Preprint.

Building on our previously reported studies on the combination of molecular dynamics and machine learning (Decherchi et al., Nature Comm 2015; Decherchi et al., JCIM 2018), we applied a combination of these techniques to identify the structural determinants causing efficacy cliffs at the D3 receptor in a small series of previously reported multi-target compounds.

Multi-Target Dopamine D3 Receptor Modulators: Actionable Knowledge for Drug Design from Molecular Dynamics and Machine Learning

Mariarosaria Ferraro,[1] Sergio Decherchi,[2] Alessio De Simone,[3] Maurizio Recanatini,[4] Andrea Cavalli,[2,4,*] Giovanni Bottegoni.[5*]

[1] Istituto di Chimica del Riconoscimento Molecolare, Consiglio Nazionale delle Ricerche (ICRM-CNR), Via Mario Bianco 9, 20131, Milan, Italy, mariarosaria.ferraro@icrm.cnr.it; [2] Computational & Chemical Biology, Italian Institute of Technology, via Morego 30, 16163 Genoa, Italy, sergio.decherchi@iit.it, andrea.cavalli@iit.it; [3] Sygnature Discovery Ltd, Bio City, Pennyfoot St, Nottingham NG1 1GR, United Kingdom, a.desimone@sygnaturediscovery.com; [4] Dept. of Pharmacy and Biotechnology, University of Bologna, via Belmeloro 6, 40126 Bologna, Italy, maurizio.recanatini@unibo.it; [5] School of Pharmacy, University of Birmingham, Sir Robert Aitken Institute for Medical Research, Edgbaston B15 2TT, United Kingdom, g.bottegoni@bham.ac.uk.

* To whom correspondence should be addressed: andrea.cavalli@iit.it; g.bottegoni@bham.ac.uk

Abstract: Local changes in the structure of G-protein coupled receptors (GPCR) binders largely affect their pharmacological profile. While the sought efficacy can be empirically obtained by introducing local modifications, the underlining structural explanation can remain elusive. Here, molecular dynamics (MD) simulations of the eticlopride-bound inactive state of the Dopamine D3 Receptor (D3DR) have been clustered using a machine learning-based approach in the attempt to rationalize the efficacy change in four congeneric modulators. Accumulating extended MD trajectories of receptor-ligand complexes, we observed how the increase in ligand flexibility progressively destabilized the crystal structure of the inactivated receptor. To prospectively validate this model, a partial agonist was rationally designed based on structural insights and computational modeling, and eventually synthesized and tested. Results turned out to be in line with the predictions. This case study suggests that the investigation of ligand flexibility in the framework of extended MD simulations can assist and inform drug design strategies, highlighting its potential role as a powerful in silico counterpart to functional assays.

Keywords: Dopamine D3 Receptor, Drug design, Machine Learning, GPCR, Molecular dynamics, Molecular recognition.

Abbreviations: D3DR: Dopamine D3 Receptor; FAAH: Fatty Acid Amide Hydrolase; D2DR: Dopamine D2 Receptor; D4DR: Dopamine D4 Receptor; cAMP: cyclic Adenosine 3′,5′-monophosphate; IL: Intracellular Loop; EL: Extracellular Loop; MTDLs: Multi-Target Directed Ligands; BPMC: Biased Probability Monte Carlo; POPC: Palmitoyl Oleoyl Phosphatidylcholine; PCA: Principal Component Analysis.

Introduction

The pharmacological properties of a drug are commonly considered to be continuous functions in chemical space: small changes in chemical structure lead to small differences in a compound's pharmacological profile. However, recent evidence has highlighted the existence of discontinuities: sometimes, small structural changes lead to large differences in one or more features. Activity cliffs are the best characterized form of discontinuity,[1] but this concept can be extended to the study of other relevant properties. GPCRs have emerged as a target class whose modulators explore rugged chemical landscapes. Small variations in GPCR binders can lead to significant changes in the efficacy, with or without affecting binding affinity.[2] This is particularly important in drug discovery because agonists, antagonists, or inverse agonists can all be therapeutically relevant, depending on the receptor and pathological framework. In some cases, the efficacy can be tuned by adopting empirical synthetic strategies. However, the underlying structural mechanisms have often remained elusive. GPCRs exist in an equilibrium ensemble of metastable conformations whose stabilization, following ligand binding, is crucial to eliciting a particular response.[3-5] The energy difference among states is often minimal, which likely explains how small structural variations in a ligand could affect the receptor's conformational ensemble. Here, we tested the possibility of using MD simulations and cluster analysis in a comparative fashion, to rationalize and predict how local structural variations affect the efficacy of four modulators which we had previously reported as multi-target binders of Fatty Acid Amide Hydrolase (FAAH), an enzyme involved in the endocannabinoid signalling pathway, and dopamine D3 receptor (D3DR).[6] Both proteins have been independently studied for nicotine addiction. Despite these targets being biochemically and structurally unrelated, we were able to conceive molecules with high affinity for both of them. Yet, accurately predicting efficacy at D3DR remained problematic. In this context, computational methods improving the design of multi-target

directed ligands (MTDLs) holds great potential toward the development of efficient drugs against tobacco dependence.

Here, we selected a training set of four compounds that, despite their high structural similarity, show increasing levels of efficacy at D3DR without any relevant change in affinity.

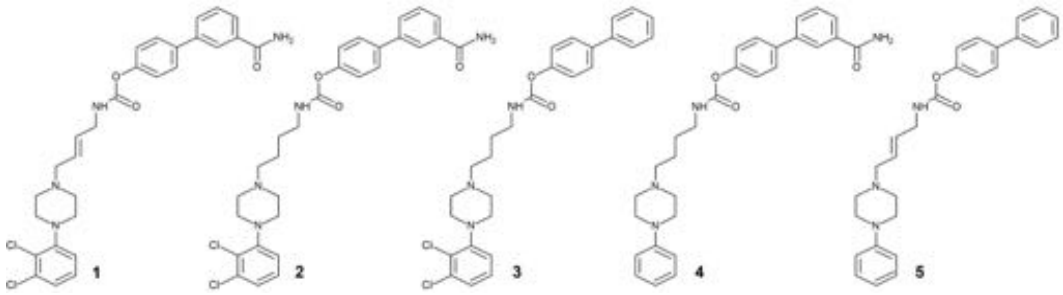


Figure 1. Chemical structure of compounds **1-5.**

Compound **1** (Figure 1) is a neutral antagonist that bears: i) a 2,3-dichloro substitution on the pendant aromatic ring of the piperazine; ii) an unsaturation in the butyl/(E)-but-2-ene linker, and iii) a carboxamide substituent in the distal ring of the biphenyl group. In the presence of a saturated linker, **2** behaves like a partial agonist, eliciting only 56% of the inhibitory response. The removal of the carboxamide in **3** does not alter the efficacy (65%) profile with respect to **2** (58%), despite a moderate decrease in activity. Compound **4** differs in features i) and ii), and is therefore an almost full agonist with 88% efficacy compared to the effects of 300 nM of dopamine on cAMP inhibition[6-8]. To derive our model and to understand which structural features affect efficacy, we attempted to interpret how the conformational behavior of each ligand within the binding site induces or stabilizes different interactions in residues H349$^{6.55}$, Y365$^{7.35}$ and S193$^{5.43}$, at the same time destabilizing the inactive crystal structure of D3DR (the superscript number indicating conserved positions according to Ballesteros and Weinstein numbering) [9]. H-bond networks involving conserved pairs of amino acids in positions 6.55, 7.35 and 5.43 have been subject of several studies on the D2-like dopamine receptors sub-family (D2, D3 and D4). In the homologous (78% sequence homology) dopamine D2 receptor (D2DR) in complex with agonists and partial agonists,[10-13] these interaction networks have been associated with low-energy patterns and functional bias. Three conserved functional serine residues on TM5, i.e., S192$^{5.42}$, S193,$^{5.43}$ and S196$^{5.46}$ are crucial in the GPCR activation pathway that involves the formation of H-bonds between ligand, water,

and receptor.[10, 14] The inward movement and anticlockwise rotation (from the extra cellular point of view) of TM5 is required to orient the serines toward the binding site. However, the serine in position 5.43 is only secondarily involved in catechol binding and has been found to establish favorable agonist-induced H-bonds with $H393^{6.55}$ in D2DR.[10, 12-13] In $H393^{6.55}A$ and $H393^{6.55}F$ mutants, a 28-fold drop in dopamine binding affinity has been observed, which correlated with reduced efficacy and confirmed the role of an aromatic residue with H-bonding capabilities in that position.[11] In the eticlopride-bound D3DR crystal structure, $Y365^{7.35}$ is fixed in a stable interaction network involving $H349^{6.55}$ and I183 on the extracellular loop 2 (EL2).[15] This interaction is absent in the D4 Dopamine Receptor (D4DR) structure complexed with nemonapride, where a valine substitutes the tyrosine in position 7.35.[16] $Y365^{7.35}V$ D3DR mutants show decreased constitutive activity of the receptor, while $V430^{7.35}Y$ cross-mutation causes the opposite effect in D4DR, highlighting that a H-bond network involving the two residues plays a functional role in regulating specific response at D3 receptor subtype. Furthermore, in microsecond-long MD simulations on an active D2DR homology model in complex with dopamine, the χ_1 torsion (C-CA-CB-CC) of $H349^{6.55}$ mainly adopted three specific dihedral angle values, namely -60°, 60°, and 140°.[12] While these angles could all be induced in the presence of the natural substrate, MD simulations with known partial agonists revealed that, in that case, the χ_1 dihedral angle of $H349^{6.55}$ was mainly stabilized at 60° or 140°.[13] These different orientations influence the interactions and the dynamics of the D2DR-G protein interface.

Therefore, the H-bond patterns established by $H349^{6.55}$ with $Y365^{7.35}$ and $S193^{5.43}$, hereafter referred to as interaction pattern 1 and interaction pattern 2, respectively, have been proposed to play a crucial role in modulating the response of D2-like sub-family of dopamine receptors in presence of molecules with different pharmacological profiles. Based on the insights gained from our studies, we designed, synthesized and tested a novel D3DR partial agonist (compound **5** in **Figure 1)** with dual FAAH/D3DR affinity and the sought efficacy profile.

Material and Methods

Ligand Docking and MD simulations

Receptor coordinates of the human D3DR crystal structure in complex with eticlopride (PDB ID: 3PBL)[15] were retrieved from the PDB and used for docking and refinement procedures implemented in the ICM software suite.[17] Hydrogen atoms were added. Polar hydrogen atoms and the positions of asparagine and glutamine side chain amidic groups were optimized and assigned the lowest energy conformation. After optimization, histidines were automatically assigned the tautomerization state that improved the hydrogen bonding pattern. Finally, the original ligand was deleted from the holo structure. Binding sites were identified with the IcmPocketFinder tool as implemented in ICM3.7.3.[18] The tolerance value was set equal to 4.6. The macro provides a mesh associated with every detected pocket. The graphical object closest to the co-crystallized ligand position was selected. All the residues with at least one side chain non-hydrogen atom within 3.5 Å of the selected mesh were considered part of the pocket. Ligands were assigned the right bond orders, stereochemistry, hydrogen atoms, and the most abundant protonation states predicted at pH 7.4. Each ligand was assigned the MMFF force field atom types and charges.[19] The docking engine used was the Biased Probability Monte Carlo (BPMC) stochastic optimizer, as implemented in ICM3.7 (Molsoft LLC, San Diego, CA – USA).[18] The ligand binding site at the receptor was represented by precalculated 0.5 Å spacing potential grid maps, representing van der Waals potentials for hydrogen and heavy atom probes, electrostatics, hydrophobicity, and hydrogen bonding. The van der Waals interactions were described with a smoother form of the 6-12 Lennard-Jones potential with the repulsive contribution capped at 4.0 kcal/mol. The electrostatic contribution was buffered, artificially increasing the distance between oppositely charged atoms to avoid their collapse when the electrostatic attractive energy prevailed on the softened van der Waals repulsion. The molecular conformation was described with internal coordinate variables. The adopted force field was a modified version of the ECEPP/3 force field with a distance-dependent dielectric constant.[20] Given the number of rotatable bonds in the ligand, an adaptive algorithm (thoroughness 3.0) calculated the basic number of BPMC steps to be carried out. The binding energy was assessed with the standard ICM empirical scoring function.[21] For each ligand, the best scoring pose was selected as representative of the ligand-bound conformation.

The D3DR in the inactive crystallized structure was properly refined and minimized after ligand placement, and chosen as a starting structure for MD simulations. Here, we focused on the ability of our compounds to destabilize the crystallized inactive state of the D3DR, triggering the onset of early local events connected to full structural transitions. In fact, large scale conformational changes remain out of reach for an antagonist-bound initial state, even with microseconds of MD trajectories. By comparing different ligands bound to the same inactive structure, we did not have to model an active state by homology. Furthermore, small differences (5-10 folds) in energy have been found between D3DR conformations with high and low affinity for agonists.[22] In the crystal structure, the long intracellular loop 3 (IL3) (R222-R318) involved in G-protein binding is not solved but substituted by T4-lysozime for stability reasons.[15] As in previously reported long MD simulations of β_2-adrenergic GPCR,[23] we did not attempt to model IL3, since IL3 does not affect ligand binding.[24-25] In this study, we adopted the same strategy for four reasons: i) ligand binding to D3DR is not significantly modified by the presence of guanyl nucleotides (G-shift), meaning that G-protein binding and activation has only a weak influence on the orthosteric binding pocket;[22] ii) D3DR expressed in *E. Coli* has shown similar ligand binding capabilities in the presence and absence of $G_{i/o}$;[25], and iii) modeling extended protein loops does not ensure the reproduction of salient features of these domains.[26]

Using the membrane embedding procedure implemented in BiKi LifeSciences (http://www.bikitech.com/),[27] we enclosed each complex in a simulation box of 8x8x10 nm containing 182 molecules of palmitoyl oleoyl phosphatidylcholine (POPC) lipids. Systems were prepared as described in the Supplementary Material. Rather than simulating multiple replicates, we preferred to carry out single MD simulations accumulating 3.05 μs on each system. This choice allowed us to increase the probability to observe rare events in the inactive structures of the receptor, enabling systematical comparisons of ligand-induced states on longer time-scales. To avoid any memory of the initial structure, we discarded the first 50 ns from each production run; the analyses were performed on 30000 snapshots extracted from each trajectory.

Cluster analysis

Conformers were pooled together based on their shared scaffold (**Figure S1**) and clustered by *k-medoids* algorithm (see description of the algorithm in the Supplementary Material).[28] To simplify the analysis and get more interpretable and robust results, three

clustering rounds were performed, merging compounds **1**, **2**, **4** in the first run, and **1**, **3**, **4** in the second run. Similarly, compounds **1**, **5**, **4** were merged in a meta-trajectory to study ligands partitioning in presence of the predicted ligand **5.**

Principal Component Analysis (PCA) in the space of dihedral angles

To get a per ligand retrospective characterization of the space sampled by compounds **1-5**, we built a dataset such that each line represented a ligand, while the columns (the features) were obtained by a processing of dihedral angles time data (30000 values x ligand). In detail, we first computed the sinuses and cosines of each angle for each ligand to correctly take into account the periodicity of the variables. Next, for each dihedral angle time series (either the sin/cos values) we computed a histogram with *nb* bins. This binning allowed us to get a discrete version of the distribution function that, at equilibrium, must be stationary. Hence, each ligand was represented by a column of ni entries where ni = 3**nb**2. The multiplier by three derives from the number of angles analyzed, while the multiplier by two derived from the sinus and cosine representations. On this matrix we performed PCA analysis.

Synthesis and Pharmacological evaluation of Compound 5

Description of synthesis and structural characterization of compound **5** are reported in the dedicated section of the Supplementary Material. The pharmacological profile of the molecule was assessed based on the ligands' inhibiting effects on cAMP accumulation *via* activation of G_i protein, as described in details in ref. [6-8].

Results

Analysis of the clustering graph in the training set (Compounds 1-4)

Compounds **1-4** were docked into the binding site of D3DR inactive structure (PDB ID: 3PBL) (see Material and Methods).[15] Similarly to the eticlopride-bound structure, all the best-scoring ligand poses reproduced the driving interaction between the basic nitrogen of the piperazine ring and the side chain of $D110^{3.32}$. As expected, given the similarity, all molecules docked consistently and in qualitative agreement with the binding mode proposed by Chien and colleagues for D3-selective derivatives.[15] In **Figure 2**, the clustering results on the **1**, **2**, **4** set are reported as a graph. Bent and extended conformations of the ligands were isolated from each trajectory but with varying frequencies. In the graph, clusters display a selective enrichment in one or more ligands (**Table S1**) with different efficacy profiles and are connected through a heterogeneously populated hub node mainly characterized by bent poses (cluster 0, **Figure 2c**). The insets 2a, 2b, 2d-2f in **Figure 2** highlight global differences in ligand scaffold orientations in the binding site (cyan licorice) relatively to the representative conformation of the hub node (gold licorice). Cluster 1 and cluster 3 were almost exclusively populated by compound **1**, with 10145 and 7624 members (**Figure 2a-b**). The substituted 2,3-dichlorophenyl ring and the trans double bond in the linker led to extended rearrangements that could be found in these clusters and that were stable in time. For example, the conformation observed in cluster 1 only appeared after 1.85 μs but was thereafter stably preserved until the end of the simulation. Lacking both the unsaturation in the linker and substituents in the pendant aromatic ring, **4** sampled the broader portion of the conformational space, showing substantial enrichment in node 7 (2005 members, bent conformation, **Figure 2f**) and almost exclusively populating cluster 5 (**Figure 2d**) and cluster 8. In the trajectory, the bent conformation associated with cluster 7 appeared after 100 ns and lasted for an additional 400 ns. Agonist-specific conformations associated with cluster 5 (9672 members) and cluster 8 (6778 members) appeared late in the simulation (around 2 μs) and were stably preserved. Compound **2** mostly interconverted between bent conformations of cluster 4 (10305 members) and cluster 6 (11773 members) (**Figure 2e**). Through MD simulations and cluster analysis carried out on the second dataset (compounds **1**, **3**, **4**) we: i) assessed the robustness of the algorithm in reproducing agonist-selective and antagonist-selective clusters obtained in the first

group; and ii) investigated the influence of the carboxamide substitution at the biphenyl group on the dynamics of **2** and **3**. The topology of the graph obtained for the set formed by **1**, **3**, and **4** (**Figure S2**) is consistent with the one reported in **Figure 2**, robustly returning a similar partition (**Table S2**), in line with the overlapping efficacies of **2** and **3** (see discussion in the Supplementary Material).

Taken together, our results suggest that different efficacy profiles could be linked to preferential stabilization of ligand-specific bent and extended conformations. Whereas compound **1** stabilized extended rearrangements, the two partial agonists **2** and **3** preferentially bound the receptor with exclusive bent conformations (clusters 4 and 6) without never transitioning into agonist-/antagonist-specific nodes. Similarly, the full agonist **4** was able to enrich exclusive clusters, but it preserved the unique feature of switching between selective bent (cluster 7) and extended conformations (clusters 5 and 8). At the receptor level, ligand-induced changes involved functional residues as $H349^{6.55}$, $Y365^{7.35}$, $S193^{5.43}$ and extracellular loops 2 and 3 (EL2, EL3) conformations. To compare our results with relevant findings on D2-like sub-family of DRs,[10-13, 15-16] we monitored the status of ligand-induced networks and local interaction patterns (1 and 2) involving the phenyl ring of the ligands, $H349^{6.55}$, $Y365^{7.35}$, and $S193^{5.43}$, in each system (**Figure 3**). Also, we collected the values of χ_1 dihedral angles in $H349^{6.55}$ and $S193^{5.43}$, which recent studies have linked to the earliest stages of the activation process. The fluctuations of these internal variables are summarized in **Figure 3** and they will be separately discussed and compared in the light of the experimental efficacies of the individual compounds.

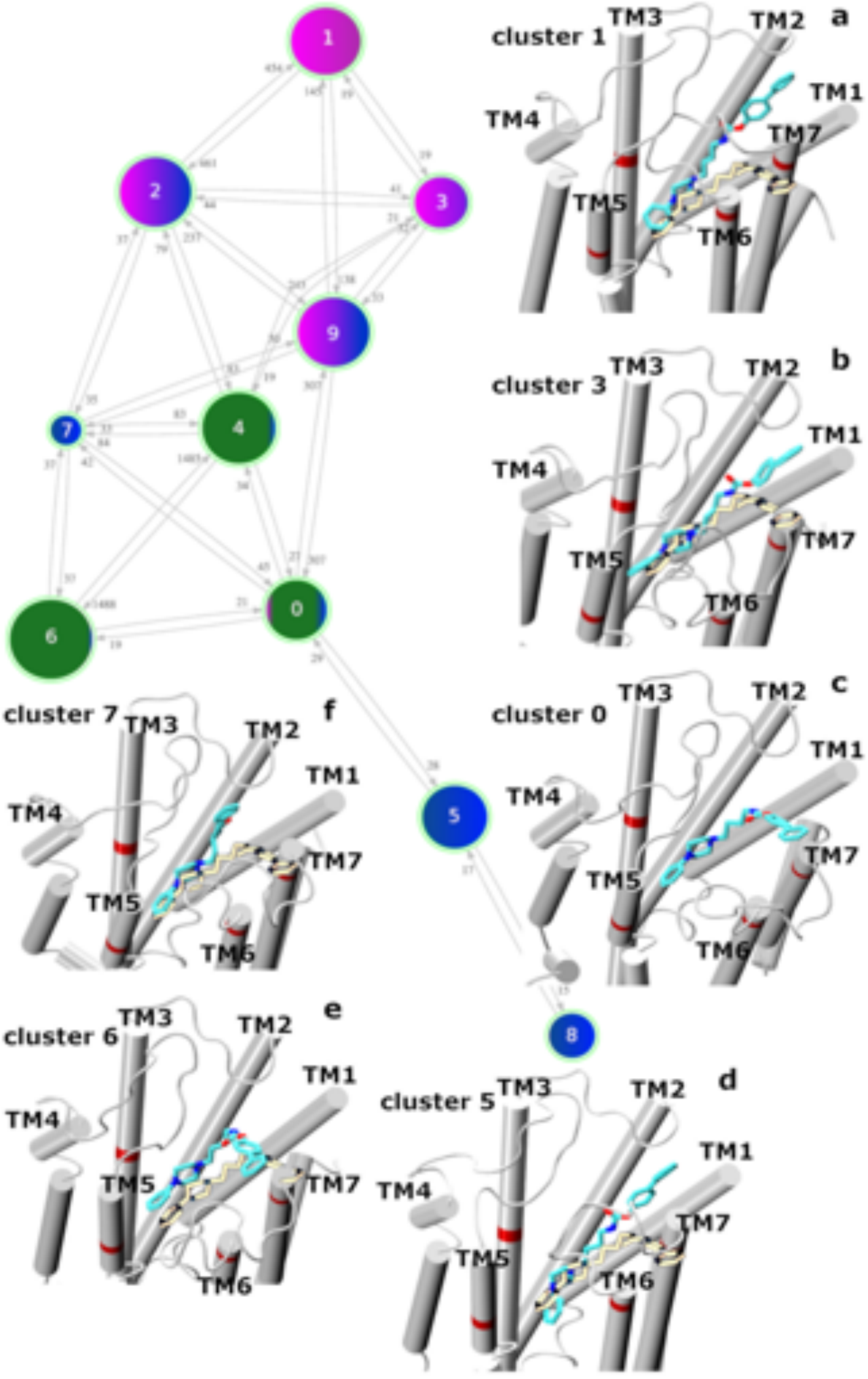

1
2
3
4
5
6
7
8
9
0
a
cluster 1
TM3
TM2
TM1
TM4
TM7
TM5
TM6
b
cluster 3
TM3
TM2
TM1
TM4
TM7
TM5
TM6
c
cluster 0
TM3
TM2
TM1
TM4
TM7
TM5
TM6
d
cluster 5
TM3
TM2
TM1
TM4
TM7
TM5
TM6
e
cluster 6
TM3
TM2
TM1
TM4
TM7
TM5
TM6
f
cluster 7
TM3
TM2
TM1
TM4
TM7
TM5
TM6

Figure 2. Cluster graph of the conformations explored by 1 (magenta), 2 (green), and 4 (blue). Each node represents a conformation. The size of each node is proportional to the cluster population. Each node is colored according to the relative cluster enrichment. Edges define transitions between clusters. In the insets (a-f), representative medoids (cyan) from each cluster are shown in complex with their corresponding D3DR conformation, and compared to "hub" medoid pose (gold) in cluster 0. Red circles on TM3, TM5, TM6, and TM7 indicate diagnostic residues D110$^{3.32}$, S193$^{5.43}$, H349$^{6.55}$, and Y365$^{7.35}$. EL2 connects TM4-TM5; EL3 connects TM6-TM7.

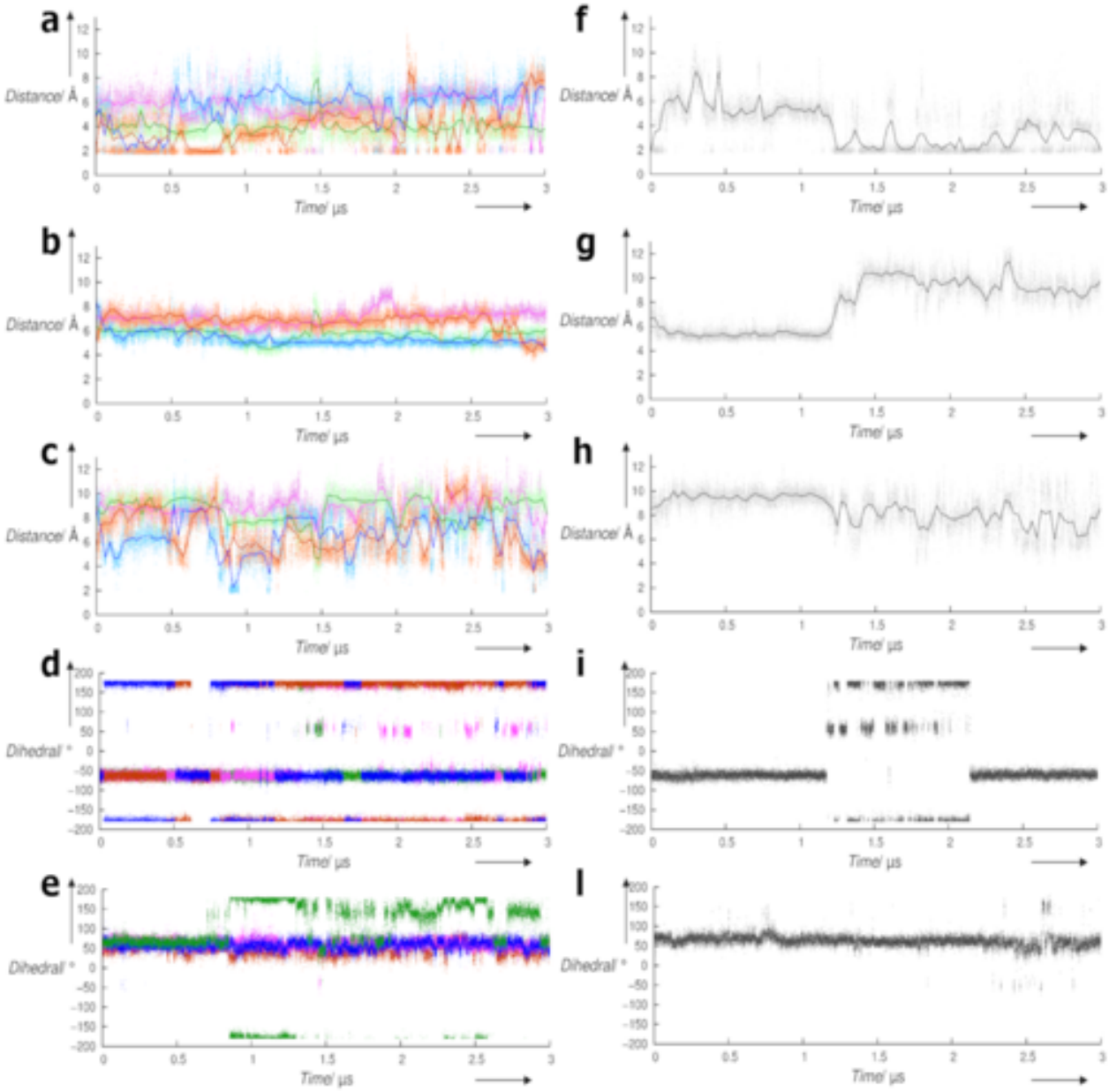


Figure 3. Diagnostic interactions monitored in D3DR-ligand systems. (**a-e**): ligands **1**-**4**; (**f-l**): predicted ligand **5**. (**a, f**) H-bond distance H349$^{6.55}$(N)-Y365$^{7.35}$(H); H-bond distances indicating the interaction pattern 1 are calculated between the hydrogen on the epsilon nitrogen atom of H349$^{6.55}$ and the phenolic

hydrogen atom of Y365$^{7.35}$; (**b, g**) aromatic interaction distance H349$^{6.55}$-phenyl ring (ligands **1-5**) calculated between rings' centers of mass. (**c, h**) H-bond distance H349$^{6.55}$(N)-S193$^{5.43}$(H); H-bond distances connoting the interaction pattern 2 are calculated between the epsilon-bound hydrogen atom of H349$^{6.55}$ and the oxygen atom of S193$^{5.43}$ side chain; (**d, i**) χ_1 dihedral angle of S193$^{5.43}$. (**e, l**) χ_1 dihedral angle of H349$^{6.55}$. Color codes for ligands **1-5** are consistent with Figures 2 and S2.

Antagonist-induced conformations

Antagonist-specific clusters 1 and 3 were both characterized by an extended conformation of the common core (**Figure 4**) and stabilized D3DR in a closed state due to concerted motions of TM6-TM7 toward TM1-TM2 (compare TM7 in **Figures 2a-c**). This shift caused EL2 to come in close contact with EL3 (**Figure 2a**). H349$^{6.55}$ and Y365$^{7.35}$ remained around 6 Å apart. In line with recent work on antagonist-bound D3DR structures, H349$^{6.55}$ and S193$^{5.43}$ were found at around 8 Å apart, indicating that the antagonist stabilized longer TM5-TM6 interface distances.[29] Therefore, neither interaction pattern 1 (**Figure 3a**, magenta line) nor interaction pattern 2 (**Figure 3c**, magenta line) were ever observed. Instead, they were replaced by stable H-bond gating bridges established by H349$^{6.55}$ and Y365$^{7.35}$ with EL2 residues, as I183 and S182 (**Figure 4** and **Figure 5a, c**). In cluster 3, the antagonist further stabilized Y365$^{7.35}$ orientation *via* T-shape π-π interactions with the biphenyl moiety (**Figure 4b**).

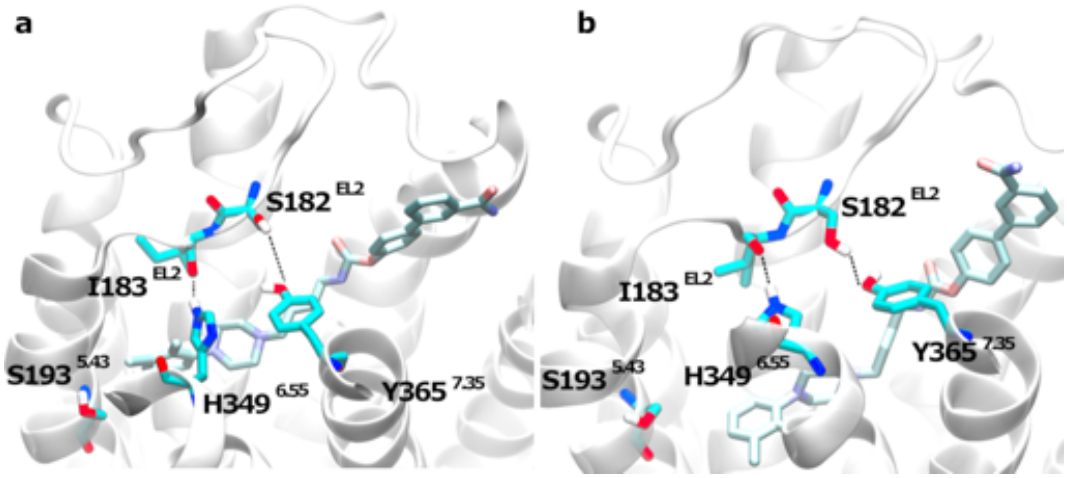


Figure 4. H349 and Y365 Interactions stabilized in the antagonist-bound system. **a**: cluster 3; **b**: cluster 1. EL3 has been removed for clarity

The antagonist limited the fluctuations of EL2, stabilizing the loop in a conformation that sealed the binding site from above (**Figure 2a**), as also observed from the lowest average number of waters surrounding the ligand along the trajectory (**Table S3** and **Figure S3**). This shielding process was enhanced by the van der Waals interactions

established by I183 side chain in the phenylpiperazine binding site. Our findings are in agreement with recently reported MD simulations and mutagenesis studies, which recognize EL2 as a crucial element in GPCR activation,[30-31] and I183 as an important residue for antagonist binding.[32]

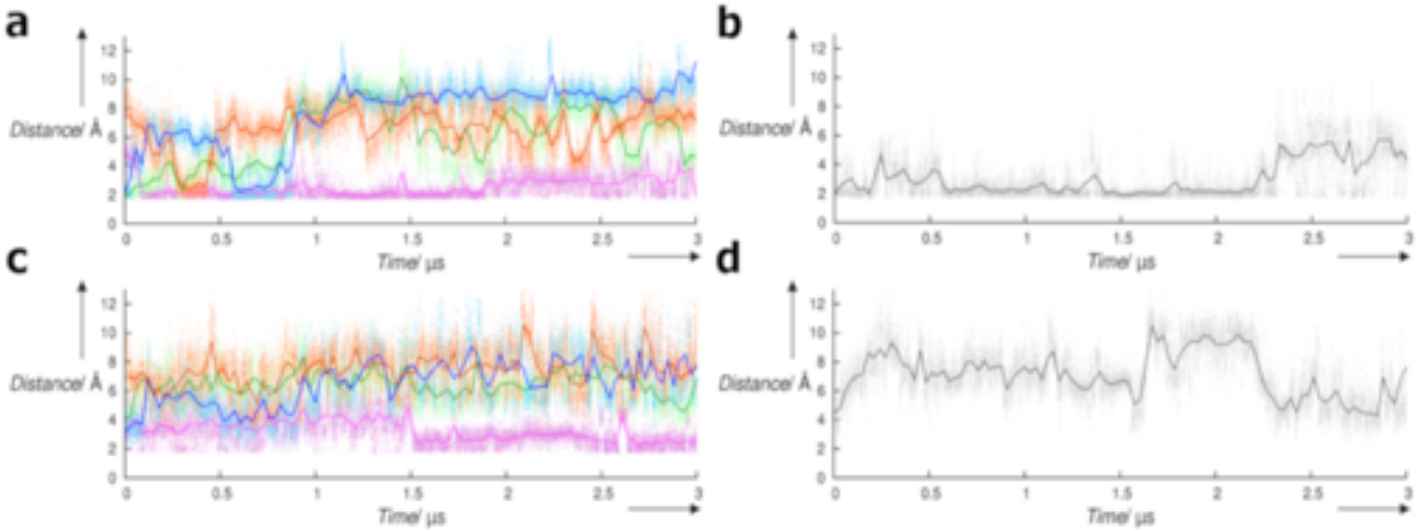


Figure 5. H-bond bridges over time for the pairs: H349$^{6.55}$(H)-I183(O) (**a-b**) and Y365$^{7.35}$(O)-S182(H) (**c-d**). Distances are shown for the four simulated ligands **1-4** (**a**, **c**) and for the predicted compound **5 (b**, **d)**. As observed from antagonist (magenta) and full agonist (blue) profiles, interaction distances between gating residues show a clearly opposite trend, with partial agonists (green, gray and orange lines) showing an intermediate behavior. Color codes for ligands **1-5** are consistent with Figures 2 and S2.

The key difference between the two antagonist clusters was in the orientation of the 2,3-dichlorophenyl ring relative to membrane plane. While cluster 3 identified a perpendicular orientation (**Figure 4a**), cluster 1 displayed a parallel one (**Figure 4b**). The 2,3-dichlorophenyl group established very weak aromatic and hydrophobic interactions with H349$^{6.55}$ (distance ≈ 6.5 Å, magenta line in **Figure 3b**). After 2 µs, the transition from cluster 3 to cluster 1 occurred and the distance between the imidazole ring in H349$^{6.55}$ and the ligand further increased, reaching an average value of 7.5 Å. This change was induced by the rotation of the C(Cl)CNC dihedral angle (see **Figure S1** and **Figures 2a, 4b**) observed in cluster 1, where the unfavorable interaction with H349$^{6.55}$ was replaced by a T-shape aromatic stacking with the side chain of F188$^{5.38}$. This phenylalanine made an inward rotation into the binding site, which was followed by the formation of π-π interactions with **1**. Contacts established with position 5.38 have been associated with β-arrestin activation in 5-HT$_{2B}$ and D2DR.[33-34] When the antagonist

stabilized in cluster 1, rotation of the phenylpiperazine core also induced F346$^{6.52}$ side chain to shift outward at the interface of TM5-6, where optimized parallel stacking between the ring of **1** and the side chains of F197$^{5.47}$, F338$^{6.44}$, and F346$^{6.52}$ was observed. In agreement with our observations, such outward rotation opens a "cryptic pocket" which was found to be crucial in explaining the efficacy of a D3DR antagonist.[32] In contrast, Michino and coworkers have recently observed that F346$^{6.52}$ rotation toward TM5-TM6 interface facilitates the inward motion of TM6 and can be considered a signature of partial agonists-driven destabilization of the inactive state of D3DR. The authors observed how the aromatic residue pointed toward the receptor core in presence of an antagonist, hindering TM6 movement.[29] Our results showed that F346$^{6.52}$ sampled both orientations, pointing toward receptor core when **1** was in cluster 3 and toward the TM5-TM6 interface when the antagonist populated cluster 1. On the same line, along the trajectory, the χ1 dihedral angle of H349$^{6.55}$ showed small fluctuations around 60° (**Figure 3e**) while S193$^{5.43}$ mainly oriented outward (-50°) (**Figure 3d**). In recent MD studies on dopamine-bound active model of D2DR, H349$^{6.55}$ χ_1 value could indeed be sampled, but with much lower frequency with respect to values conducive to a fully substrate-activated state of the receptor (-60°), whereas S193$^{5.43}$ χ_1 value stabilized around 160°.[12]

Partial agonist-induced conformations

In **2** and **3**, a partial agonist efficacy profile was associated with a different behavior, as a consequence of the increased ligand flexibility. Namely, bent conformations of **2** observed in nodes 0, 4, 6, and 7 were stabilized by T-shape stacking between the ligand 2,3-dichlorophenyl ring and F346$^{6.52}$. This residue pointed toward the receptor core, as in the receptor crystallographic structure, without rotating outward. This is particularly relevant in light of the results recently reported by Ferruz and coworkers,[32] who have shown how inward/outward conformations of F346$^{6.52}$ could have a crucial role in D3DR response to ligands of varying efficacies (see the Supplementary Material for detailed discussion of partial agonist **3**). In our simulations, partial agonist binding dynamics did not lead to antagonist-induced H-bond bridges (**Figure 5**). Conversely, we could observe EL2 displacement **(Figure 2e and Figure S2a**), enhancing water influx into the orthosteric binding pocket relatively to **1** (**Figure S3 and Table S3**). Compounds **2** and **3** facilitated this process preserving a dynamic coupling of TM6-TM7 interface. Indeed, distances between interacting atoms in H349$^{6.55}$ and Y365$^{7.35}$ were substantially shorter

in these trajectories than in the antagonist-bound receptor, dropping on average by 2 to 4 Å. The 2,3-dichlorophenyl ring of **2**, as also seen for **1,** made only weak T-shape π-π interactions with H349$^{6.55}$ (**Figure 3b**). Furthermore, in contrast to the antagonist, bent conformations of compounds **2** and **3** formed a parallel stacking interaction with the side chain of Y365$^{7.35}$, promoting the breaking of the H349$^{6.55}$-I183 and Y365$^{7.35}$-S182 H-bonds and the formation of the interaction pattern 1 (**Figure 6** and **Figure S4**). For compound **2**, the break occurred more frequently in the pair Y365$^{7.35}$-S182 (**Figure 5c**). Indeed, in cluster 4, this ligand stabilized also a partially open state of the gate, retaining H349$^{6.55}$ in proximity of both Y365$^{7.35}$ and I183 (**Figure 6a**).

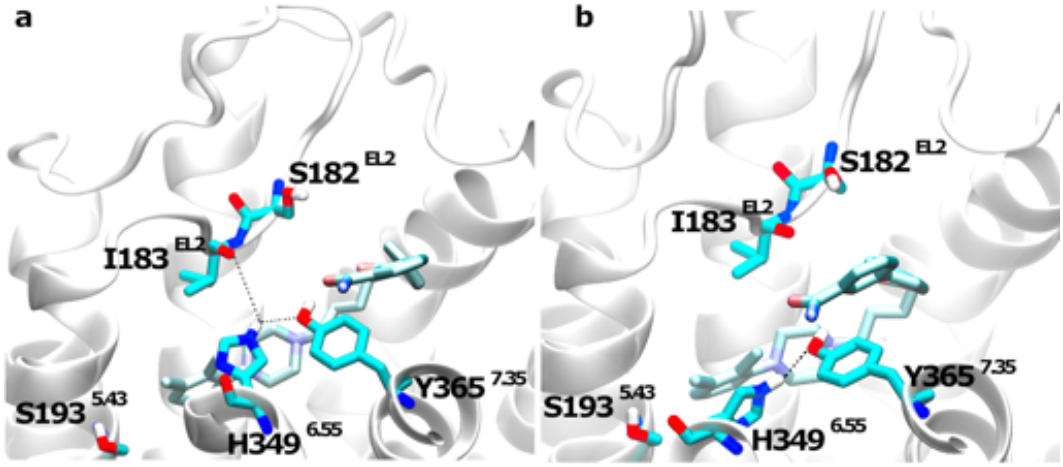


Figure 6. Interaction pattern 1 observed in D3DR by partial agonist **2**. **a**: cluster 4; **b**: cluster 6. EL3 has been removed for clarity.

Accordingly, the H349$^{6.55}$-S193$^{5.43}$ distance was on average longer for partial agonists-bound receptor, with values fluctuating in the range of 4-10 Å (**Figure 3c**). Partial agonists could block progression toward a fully active state by preventing stable TM5-TM6 interface contacts.[13] Thus, none of the simulated partial agonists was able to establish significant H-bond interactions between residues involved in interaction pattern 2. Our results agree with recent MD simulations of D3DR-ligand complexes, where bent poses have been associated with partial agonism.[29] In line with these findings, in selective clusters 4 and 6 (**Figure 2** and **Figure S2**) the distance between TM5 and TM6 increased, while tightening interactions at TM6-TM7 interface could be observed. The formation of the interaction pattern 1 in partial agonists simulations stabilized this inter-helical rearrangement. Inward rotation of the χ_1 angle of the H349$^{6.55}$

in cluster 6 did not break pattern 1 (**Figure 3a, e, 6b**) and induced substantial side-bending of TM6. The analyzed binding modes were in very good agreement with the effect induced by the partial agonist FAUC350 on the same dihedral angle of H349$^{6.55}$ in the active ternary model of D2DR, where the ligand promoted coupling of TM6-7 and formation of the H349$^{6.55}$-Y365$^{7.53}$ H-bond. Interestingly, this study reports that differential stabilization of inter-helical interaction patterns in the G-protein-bound model of D2DR is responsible for loosening intracellular coupling between the G-protein and D2DR, likely identifying structural patterns at the basis of partial agonism at the highly homologous D2DR.[13]

Agonist-induced conformations

In analogy with partial agonists, the almost full agonist **4** (hereafter simply referred to as agonist) initially induced interaction pattern 1 in the representative bent conformation of cluster 7 (**Figure 3a** and **Figure 7a**), while also inducing TM6-TM7 coupling.

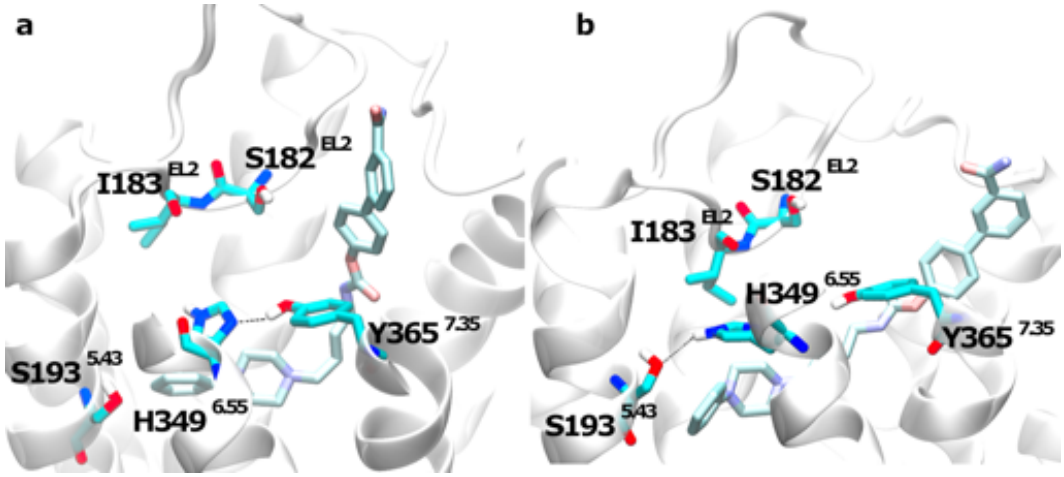


Figure 7. Interaction patterns observed in D3DR in complex with the agonist 4 a: cluster 7; b: cluster 5. EL3 has been removed for clarity.

The gating H-bond interaction between Y365$^{7.35}$ and S182 was not formed when **4** populated this cluster, again indicating that TM6-TM7 coupling promoted receptor opening (**Figure 5c**). Consistently, EL2 displacement was observed and this was, in turn, conducive to a pronounced increase in solvation (**Figure S3** and **Table S3**). These results are in line with NMR studies on rhodopsin in which H-bond network reorganization

between EL2 and TM4-6 has been coupled to EL2 displacement from the binding site during substrate-induced activation.[30-31]

In cluster 7, the pendant phenyl ring in the arylpiperazine moiety of **4** and the side chain of H349$^{6.55}$ formed an aromatic T-shape π-π interaction which was uniquely preserved by the agonist along the whole trajectory. Indeed, the distance between the two rings was stably preserved at approximately 5 Å (**Figure 3b**). The dihedral angle of S193$^{5.43}$ frequently rotated inward, around 160° (**Figure 3d**), i.e. the value observed for dopamine in the active model of D2DR.[12-13]

Our simulations revealed that interaction pattern 1 destabilized after 500 ns (**Figure 3a**) and the trajectory evolved toward the conformations that populate cluster 5 and cluster 8. These agonist-selective extended conformations caused the largest increase in the H349$^{6.55}$-I183 distance (**Figure 5a**), which further allowed waters to reach the pocket (**Table S3** and **Figure S3**). Besides the aromatic interactions with H349$^{6.55}$, the phenyl ring of **4** established tight parallel stacking with F346$^{6.52}$. While H349$^{6.55}$ side chain maintained its initial orientation ($\chi 1 \approx 60°$, **Figure 3e**), F346$^{6.52}$ side chain underwent an exclusive rotation into the binding site in the opposite direction with respect to what was observed for **1**, and pointed toward the intracellular region of the receptor. The ability to maintain stable interactions with H349$^{6.55}$ and F346$^{6.52}$ was a unique feature of **4**. The ligand-induced rotameric state of F346$^{6.52}$ was observed only in response to the shift of the phenyl ring of the agonist at the TM5-TM6 interface, where it preserved an orientation perpendicular to membrane plane (**Figure 7b**). Inward rotation of F346$^{6.52}$ observed in our simulations minimized the steric hindrance that hampers TM6 inward motion, which is a crucial event in the destabilization of the receptor inactive state. The biphenyl ring interacted with Y365$^{7.35}$ pushing TM7 toward TM5-6. Furthermore, stable interactions of the phenylpiperazine ring of **4** with H349$^{6.55}$ at TM5-TM6 interface brought the histidine to point toward TM5 with higher frequency. In this conformation, H349$^{6.55}$ and S193$^{5.43}$ were only 2 to 4 Å apart, a rearrangement never observed in ligands with lower efficacy (**Figure 3c** and **Figure 7b**). The interaction pattern 2 first appeared in the trajectory around 1 μs, remained stable for roughly 100 ns, and was later transiently re-visited, as observed in **Figure 3c** (spikes in the blue line). The agonist, probably due to the persistent memory of the initial inactive state of D3DR, was not able to induce a complete rotation of H349$^{6.55}$ and S193$^{5.43}$ χ_1 torsions toward -60° and 160°, respectively, which are the values that fully optimize the H349$^{6.55}$-S193$^{5.43}$ interaction in dopamine-bound

D2DR.[12] However, even without stably preserving interaction pattern 2, in our simulations of D3DR in complex with **4**, a series of agonist-specific changes in the interaction networks, which did not occur in antagonist- and partial agonist-bound complexes, could be observed. In particular, compound **4** selectively enriched bent and extended clusters, in which it was the only ligand able to preserve tight interaction with the H349$^{6.55}$. In doing so, the agonist stabilized the interaction pattern 1 for 500 ns and also induced transient formation of the interaction pattern 2 for a maximum time of 100 ns, complementing the shift of H349$^{6.55}$ at TM5-TM6 interface. Such changes promoted abrupt opening of the extracellular portion of receptor and temporarily increased the contacts between TM5 and TM6. Maintaining aromatic interactions with F346$^{6.52}$ in a uniquely inward-rotated state contributed to this scenario, reducing the hindrance at the interface.

Design, MD analysis and biological evaluation of the predicted compound 5

Taken together, our results provided the structural basis for understanding the varying efficacy of the **1**-**4** series. The 2,3-dichloro substitution and the butyl/(E)-but-2-ene linker were both needed to obtain a full antagonist, i.e. a molecule that was unable to promote any of the agonist-stabilized interaction patterns and that preserved D3DR in a closed configuration. Conversely, the removal of chloride atoms and the introduction of a flexible butyl linker led to an almost full agonist activity for **4**. Compound **4** was able to uniquely establish long-lasting contacts with functional histidine in position 6.55, orienting this residue to establish both of the investigated interaction patterns, albeit to different extents. Interestingly, the partial agonist properties of semi-flexible **2** and **3** (bearing the 2,3-dichloro substitutions, but a saturated butyl linker) were explained according to their ability to stably induce only one of the intermediate interaction schemes, thus hampering but not completely blocking the cascade of events that concurred in the perturbation of the inactive receptor state. These results are in good agreement with the structure-efficacy relationship previously reported for other series of compounds.[13]

To gain confidence on this model, we designed a new compound introducing on the shared scaffold local modifications that, based on our understanding of the structure-efficacy relationship, were likely conducive to a partial agonist profile. Two main aspects were considered in the design of the new ligand. First, our MD studies tried to provide a rationale for previous studies on a series of structurally related phenylpiperazine

derivatives, which showed that the simple replacement of a butyl linker with a butyl/(E)-but-2-ene in two identical ligands could transform an agonist into a partial agonist, and a partial agonist into a full antagonist.[35] We reasoned that the introduction of a butyl/(E)-but-2-ene linker in **4** could reduce ligand flexibility, and in doing so, it could prevent the stabilization of both patterns. Second, we reduced the molecular weight of the new compound by removing the carboxamide substituent in the *meta* position of the biphenyl group. Our SAR studies on *O*-aryl carbamate derivatives revealed that this substituent influences the affinity for the receptor but not the efficacy of the ligand.[6,8] Insights gained from our simulations revealed that removing this moiety allowed **3** to sample a wider set of conformations, which, however, resulted in the induction of the interaction pattern 1, that is, the hallmark of partial agonists' profiles (see the Supplementary Material and Figure S2). The designed compound **5** is shown in **Figure 1**.

The conformations obtained from 3.05 μs of MD simulations were then subjected to the previous analysis, merging the trajectories of compounds **1**, **5**, and **4** to assess how the new ligand conformations partitioned in the presence of our reference agonist and antagonist. The obtained clusters are reported in **Figure 8**. As in previous cases, the graph highlighted a hub cluster, cluster 0, which was almost equally populated by the three ligands (**Table S4**). In analogy with **3**, thanks to the increased flexibility in the analyzed dataset (**1**, **5**, **4**), the hub node was characterized by extended and not bent conformations (**Figure 8a**). The featuring binding modes of **5** grouped in cluster 4 and cluster 6 (**Figure 8b-c**), with 9688 and 8429 conformations, respectively. An interesting difference between the partitions was the relative abundance of **5** in the agonist-selective cluster 7 (**Figure 8**). This was three times greater than **4** (6463 vs. 2173 members) (**Table S4).** In the previous analysis, this cluster contained only 569 conformations of **3** and 126 of **2**. Our model identified this medoid as a crucial intermediate for establishing interaction pattern 1 in the agonist simulation. Interaction scheme 1 was never formed in cluster 4 (**Figure 9a**). It only appeared after 1.2 μs when **5** transitioned first to cluster 7 (refer to medoid in **Figure 7a**) and then to cluster 6 (**Figure 9b**). Indeed, this pattern was stably preserved until the end of the simulation (**Figure 3f**), with H349$^{6.55}$-Y365$^{7.35}$ distance fluctuating in the range 2-4 Å. While the histidine kept a stable χ_1 angle of 60° (**Figure 3l**), following the formation of the interaction pattern 1 and population of the agonist-like cluster 7, we observed inward rotation of S193$^{5.43}$ in TM5 (from 60° to 180°) (**Figure 3i**). In line with results on **2** and **3**, the interaction pattern 2 was never formed in presence of **5** (**Figure 3h**), as the distance between H349$^{6.55}$ and S193$^{5.43}$ was stably over 4 Å for the

whole trajectory. In clusters 4, 6 and 7, the receptor was found in an open state with TM6-7 being dynamically coupled through stabilization of the interaction pattern 1 (see also TM6-TM7 in **Figure 8b-c**). The average number of solvent molecules around the ligand was comparable to **2-3** and in between the values calculated for **1** and **4** (**Table S3** and **Figure S3**).

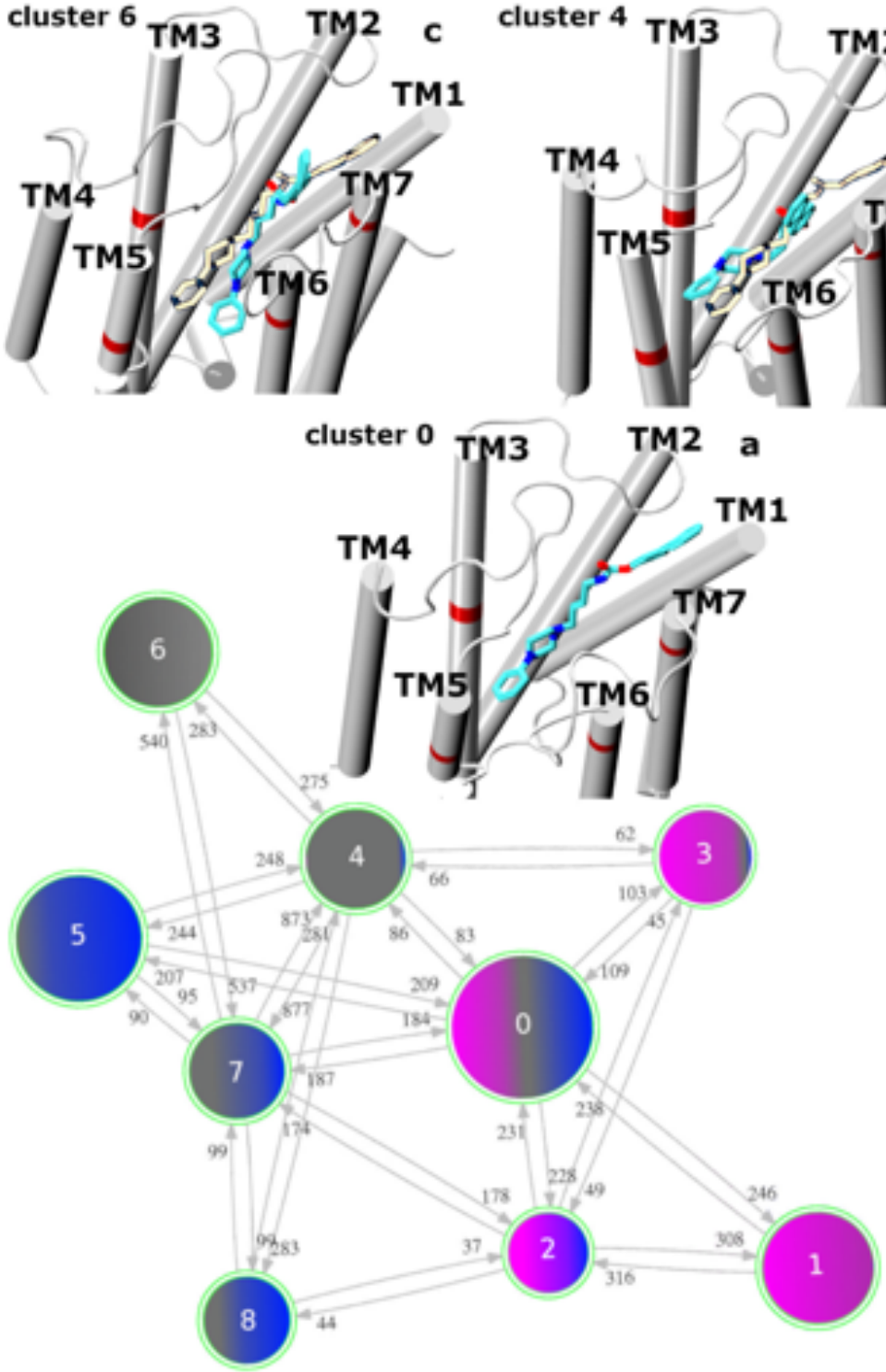


Figure 8. Cluster graph of the conformations explored by compounds **1** (magenta), **5** (gray), and **4** (blue). See Figure 2 for graph description. In the insets (**a-c**), representative medoids (cyan) of each cluster are shown in complex with their corresponding D3DR conformation and compared to the most populated "hub" medoid pose (gold) in cluster 0.

Indeed, bending of compound **5** in cluster 4 caused the breaking of the H-bond between $Y365^{7.35}$ and S182. Later in the trajectory, the ligand stabilized in the bent orientation of cluster 6, where both gating interactions were broken (**Figure 5b, d**) and interaction pattern 1 was stably induced (**Figure 3f and Figure 9**)

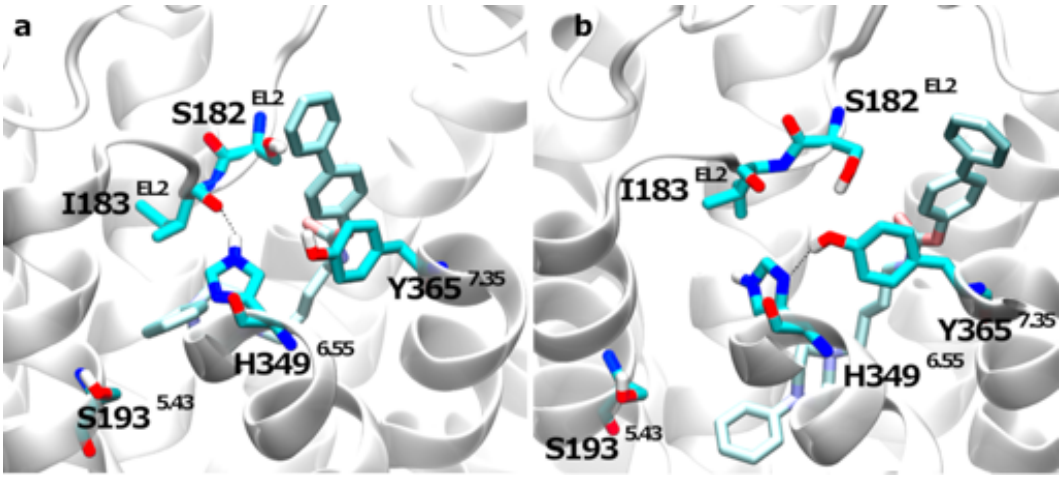


Figure 9. Interaction patterns induced by compound **5**. **a**: cluster 4; **b**: cluster 6. EL3 has been removed for clarity.

In analogy with **3**, in cluster 6, the lack of the carboxamide substituent in the distal phenyl ring of the biphenyl group induced the ligand to drift deeper into the pocket (Figure 8c and Figure 9b), compromising the aromatic interaction between the ligand phenyl ring and $H349^{6.55}$ (**Figure 3g**). Ligand **5** interacted with $F338^{6.44}$, $F197^{5.47}$, and $F346^{6.52}$ at the interface of TM5-TM6, inducing outward rotation of $F346^{6.52}$ side chain and causing an increase in distance between these two helixes. To release the steric clash between $F346^{5.47}$ and $F338^{6.44}$, TM5 slightly rotated clockwise, inducing EL2 to partially extend over the binding site (see relative orientations of TM5 and EL2 in cluster 4 to 6 in **Figure 8b**, **c**). These results are in agreement with the recently proposed mechanism of D3DR activation by a phenylpiperazine series of partial agonists.[29] In contrast, recent MD studies on D3DR antagonists from Ferruz and colleagues have associated the outward rotation of $F346^{6.52}$ side chain to the formation of the cryptic pocket responsible for antagonist-like responses.[31] Our observations suggest that the interaction of partial

agonists, like **3** and **5** with this cryptic site could actually be responsible for antagonist-like properties, resulting in hampered activation of D3DR and partial agonism profiles. However, MD simulations of partial agonist **2** showed that this compound stabilized the rotameric state observed for F346$^{6.52}$ in the eticlopride-bound D3DR,[15] suggesting that both conformations are likely to reduce the efficacy depending on the preferred bent and extended binding mode. In other words, we propose that an antagonist would likely elicit no response if an outward rotation of the F346$^{6.52}$ occurs in D3DR. When a rigid ligand binds the receptor in an extended mode, it blocks solvent access, preventing the formation of any interaction pattern. In contrast, in presence of semi-flexible partial agonists, which stabilize a more open conformation of the binding site, an outward rotation of the active site phenylalanine could still concur to destabilization of the inactive state. Interestingly, when our flexible agonist increased receptor solvation, assuming bent and extended conformations, F346$^{6.52}$ rotated in an unexpected direction, orienting its side chain toward the center of the helical bundle. Tight binding of H349$^{6.55}$ was also found to be uniquely preserved in the full agonist simulation, where both interaction patterns were visited. In contrast, this residue seems not crucial for antagonism. In partial agonists simulations, weak interactions with this residue were responsible for the induction of just one of the functional interaction schemes involved in D3DR activation.

Importantly, the partition for compounds **1** and **4** was very robust. This is because, despite changing the initial set of conformations, medoid positions along the trajectory did not change. Indeed, we found that clusters 1, 3, 5, 7, and 8 were consistent with those already identified in the two previous sets of compounds. According to our model, compound **5** behaved as a partial agonist, destabilizing the inactive state, promoting receptor opening *via* TM6-TM7 coupling and binding site solvation. In line with other partial agonists, it induced only one of the two interaction patterns associated with agonist-like properties. To validate our hypothesis, **5** was synthesized and its experimental efficacy was tested *in vitro* for its ability to inhibit accumulation of cAMP.[6-8]

Compound **5** revealed partial agonist properties when compared to the effects elicited by 300 nM of dopamine (D3DR efficacy: 60%). Moreover, dropping the terminal carboxamide led to a D3DR 23 nM EC_{50}. As expected, this is weaker than that of **2** and **3**. However, it is a promising value in light of the reduced molecular weight.

Dihedral PCA

To further characterize our results, we also performed a retrospective compound-based PCA analysis of torsional angles (**Figure 10**). Namely, we attempted to identify key torsional angles in the analyzed series of compounds. The dataset was obtained by extracting the values of three dihedral angles from **1-4** (**Figure S1**) over the entire 12 μs (120000 data points) of aggregate production runs. Next, data were binned in a reduced number of representative histograms. We performed a dimensionality reduction through dihedral principal component analysis (PCA) on this space and projected on the three main components to help visualize our ligands. Compound **5** was then embedded in the space identified by the four known compounds. The two distances obtained by averaging those from **1** and **4** were $1.43*10^4$ and $1.46*10^4$ for compound **2** and compound **5**, respectively, making them equidistant from the two reference points. These results were robust to changes in the initial number of bins.

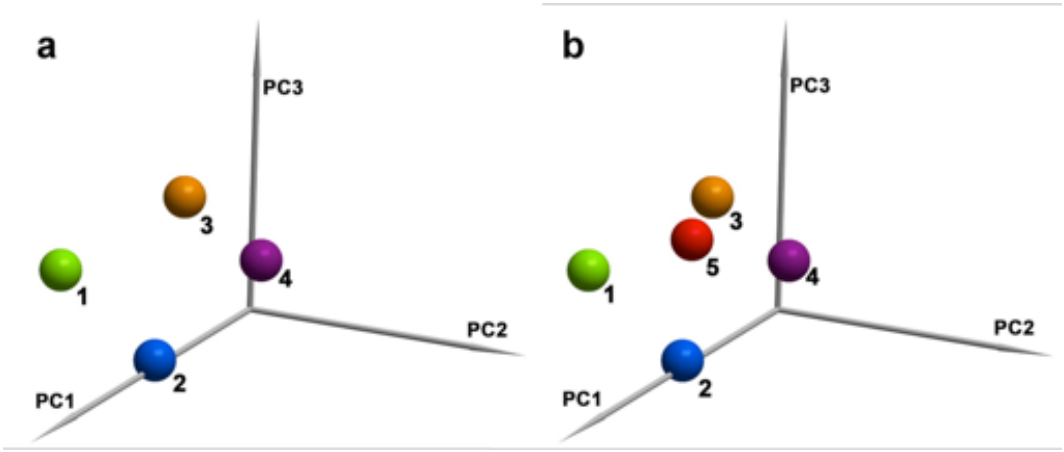


Figure 10. The first three components of the vectorial space defined by the four ligands' dihedral angles are shown after PCA. **a**) Projections of compounds **1-4**. **b**) Projection of the predicted compound **5** into the same space. Colored spheres identify the position of the five ligands.

Compound **5** actually stabilized in a specific orientation, which, as for **2** and **3**, was somewhere in between the agonist and antagonist ones. This suggests that **5** could possess the same functionality of **2** and **3** at D3DR. Overall, although limited to a dataset of only four compounds, we found this (retrospectively applied) analysis useful in obtaining a quick and concise understanding of the simulations, recapitulating insights obtained with more complex analytical frameworks. The highly symmetric behavior of **2**,**3** and **5** when compared to **1** and **4** suggests that this could be an effective vector space where the efficacy profile of new compounds belonging to this series could be prospectively

characterized. This also make us optimistic that a similar framework could be duplicated for other series and, possibly, other targets.

Conclusions

In summary, we used MD simulations to rationalize the experimentally observed efficacies of *O*-aryl carbamate derivatives. Our comparative analyses on the destabilizing effect of our ligands on the D3DR inactivated structure, we got evidence that structurally similar molecules can engage in subtly different interaction patterns and that these are, in turn, conducive to different efficacy profiles. The conformational changes reported from our simulations could be only related to the destabilization of the D3DR reference state, albeit known to be connected to initial steps in the concerted process of receptor activation. We found that the extent of these conformational changes was helpful in discriminating between ligand efficacies, and could therefore be of great help in designing a new ligand with a tailored pharmacological profile. Compound **5** was designed based on SER data to further test the consistency of the simulative outcome and was eventually synthesized and tested. As expected, this ligand behaved as a partial agonist. In due time, and in light of the ever-increasing computational power available to the scientific community, this work could pave the way to a more systematic application of MD as the in silico counterpart of functional assays, much as docking and free-energy methods can be regarded as the in silico counterpart of binding assays.

Acknowledgments

We sincerely thank Dr. Matteo Masetti for constructive scientific discussions.
This research did not receive any specific grant from funding agencies in the public, commercial or not-profit sectors.

References

[1] J. Bajorath, *Expert Opin. Drug Discov.* **2017**, *12* (9), 879-883.
[2] P. I. Dosa, E. A. Amin, *J. Med. Chem.* **2015**, *59*, 810-840.
[3] M. J. M. Niesen, S. Bhattacharya, N. Vaidehi, *JACS* **2011**, *133*, 13197-13204;
[4] P. Ghanouni, Z. Gryczynski, J. J. Steenhuis, T. W. Lee, D. L. Farrens, J. R. Lakowicz, B. K. Kobilka, *J. Biol. Chem.* **2001**, *276*, 24433-24436;
[5] R. Seifert, K. Wenzel-Seifert, U. Gether, B. K. Kobilka, *J. Pharmacol. Exp. Ther.* **2001**, *297*, 1218-1226.

[6] A. De Simone, D. Russo, G. F. Ruda, A. Micoli, M. Ferraro, R. M. C. Di Martino, G. Ottonello, M. Summa, A. Armirotti, T. Bandiera, A. Cavalli, G. Bottegoni, *J. Med. Chem.* **2017,** *60* (6), 2287-2304
[7] A. De Simone, G. F. Ruda, C. Albani, G. Tarozzo, T. Bandiera, D. Piomelli, A. Cavalli, G.Bottegoni, *Chem. Commun.* **2014**, *50*, 4904–4907.
[8] Micoli, A. De Simone, D. Russo, G. Ottonello, G. Colombano, G. F. Ruda, T. Bandiera, A. Cavalli, G. Bottegoni, *Med. Chem. Comm* **2016**, *7*, 537-541
[9] J. A. Ballesteros, H. Weinstein, in *Methods in Neurosciences, Vol. 25* (Ed.: C. S. Stuart), Academic Press, **1995**, pp. 366-428
[10] J. C. Fowler, S. Bhattacharya, J. D. Urban, N. Vaidehi, R. B. Mailman, *Mol. Pharmacol.* **2012**, *81*, 820-831;
[11] N. Tschammer, S. Bollinger, T. Kenakin, P. Gmeiner, *Mol Pharmacol* **2011**, *79*, 575-585
[12] R. C. Kling, H. Lanig, T. Clark, P. Gmeiner, *PLoS One* **2013**, *8*, e67244;
[13] R. C. Kling, N. Tschammer, H. Lanig, T. Clark, P. Gmeiner, *PLoS One* **2014**, *9*, e100069;
[14] M. Congreve, C. J. Langmead, J. S. Mason, F. H. Marshall, *J. Med. Chem.* **2011**, *54*, 4283-4311.
[15] E. Y. Chien, W. Liu, Q. Zhao, V. Katritch, G. W. Han, M. A. Hanson, L. Shi, A. H. Newman, J. A. Javitch, V. Cherezov, R. C. Stevens, *Science* **2010**, *330*, 1091-1095.
[16] S. Wang, D. Wacker, A. Levit, T. Che, R. M. Betz, J. D. McCorvy, A. J. Venkatakrishnan, X. Huang, R. O. Dror, B. K. Shoichet, B. L. Roth1, *Science*. **2017**, *358*, 381–386.
[17] R. Abagyan, M. Totrov, D. Kuznetsov, *J. Comput. Chem.* **1994**, *15*, 488-506.
[18] M. Totrov, R. Abagyan, *Proteins* **1997**, *Suppl 1*, 215-220.
[19] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490-519.
[20] G. Nemethy, K. D. Gibson, K. A. Palmer, C. N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, H. A. Scheraga, *J. Phys. Chem.* **1992**, *96*, 6472-6484.
[21] M. Totrov, R. Abagyan in *Drug-receptor thermodynamics: introduction and applications* (Ed: R. B. Raffa) *John Wiley & Sons, Ltd.* **2001**, 603-624.
[22] J. A. Ahlgren-Beckendorf, B. Levant, *J. Recept. Signal Transduct.* **2004**, *24*, 117-130.
[23] R. O. Dror, D. H. Arlow, P. Maragakis, T. J. Mildorf, A. C. Pan, H. Xu, D. W. Borhani, D. E. Shaw, *Proceedings of the National Academy of Sciences* **2011**, *108*, 18684-18689
[24] R. C. Rubenstein, S. K. Wong, E. M. Ross, *J. Biol. Chem.* **1987**, *262*, 16655-16662.
[25] J. F. Vanhauwe, K. Josson, W. H. Luyten, A. J. Driessen, J. E. Leysen, *J. Pharmacol. Exp. Ther.* **2000**, *295*, 274-283.
[26] F. Eisenmenger, P. Argos, R. Abagyan, *J. Mol. Biol.* **1993**, *231*, 849-860.
[27] S. Decherchi, G. Bottegoni, A. Spitaleri, W. Rocchia, A. Cavalli, *J. Chem. Inf. Model.* **2018**, *58*, 219-224.
[28] S. Decherchi, A. Berteotti, G. Bottegoni, W. Rocchia, A. Cavalli, *Nat. Commun.* **2015**, *6*, 6155;
[29] M. Michino, C. A. Boateng, P. Donthamsetti, H. Yano, O. M. Bakare, A. Bonifazi, M. P. Ellenberger, T. M. Keck, V. Kumar, C. Zhu, R. Verma, J. R. Deschamps, J. A. Javitch, A. H. Newman, L. Shi, *J. Med. Chem.* **2017**, *60*, 580-593.

[30] M. Wheatley, D. Wootten, M. T. Conner, J. Simms, R. Kendrick, R. T. Logan, D. R. Poyner, J. Barwell, *British Journal of Pharmacology* **2012**, *165*, 1688-1703.
[31] S. Ahuja, V. Hornak, E. C. Yan, N. Syrett, J. A. Goncalves, A. Hirshfeld, M. Ziliox, T. P. Sakmar, M. Sheves, P. J. Reeves, *Nat. Struct. Mol. Biol.* **2009**, *16*, 168-175.
[32] N. Ferruz, S. Doerr, M. A. Vanase-Frawley, Y. Zou, X. Chen, E. S. Marr, R. T. Nelson, B. L. Kormos, T. T. Wager, X. Hou, A. Villalobos, S. Sciabola, G. De Fabritiis, *Sci. Rep.* **2018**, *8*, 897.
[33] D. Wacker, C. Wang, V. Katritch, G. W. Han, X. P. Huang, E. Vardy, J. D. McCorvy, Y. Jiang, M. Chu, F. Y. Siu, W. Liu, H. E. Xu, V. Cherezov, B. L. Roth, R. C. Stevens, *Science* **2013**, *340*, 615-619;
[34] R. B. Free, L. S. Chun, A. E. Moritz, B. N. Miller, T. B. Doyle, J. L. Conroy, A. Padron, J. A. Meade, J. Xiao, X. Hu, A. E. Dulcey, Y. Han, L. Duan, S. Titus, M. Bryant-Genevier, E. Barnaeva, M. Ferrer, J. A. Javitch, T. Beuming, L. Shi, N. T. Southall, J. J. Marugan, D. R. Sibley, *Mol. Pharmacol.* **2014**, *86*, 96-105.
[35] M. Taylor, P. Grundt, S. A. Griffin, A. H. Newman, R. R. Luedtke, *Synapse* **2010**, *64*, 251-266.

Multi-Target Dopamine D3 Receptor Modulators: Actionable Knowledge for Drug Design from Molecular Dynamics and Machine Learning

Mariarosaria Ferraro, Sergio Decherchi, Alessio De Simone, Maurizio Recanatini, Andrea Cavalli,* Giovanni Bottegoni.*

Table of Contents

Computational Methods

Molecular Dynamics Setup

The membrane protein complex was then solvated with an average of 12300 water molecules (TIP3P model). Force fields available in AMBER 14 were used to parameterize protein,[1] lipids, and ligands, corresponding to ff14SB,[2] lipid14,[3] and GAFF,[4] respectively. Point charges for ligands were derived from the electrostatic potential calculated after geometry optimization at the Hartree-Fock level of theory with a 6-31G* basis set, following the RESP procedure as implemented in Antechamber. Simulations were performed on GPU-equipped workstations with Gromacs 4.6.7 MD engine.[5] In detail, the MD protocol encompassed three steps: minimization, equilibration, and production. Each system was minimized for 5000 steps and then thermalized to 300 K in different phases. Temperature was raised to 300 K in 300 ps within the NVT ensemble, in three consecutive increments of 100 K lasting 100 ps each. Then, volume and density were allowed to equilibrate in NPT ensemble at 300 K and target pressure of 1 bar for 200 ps. Lipids, ligands, and water molecules were equilibrated first, applying position constraints only to protein backbone (1000 $kJmol^{-1}nm^{-2}$) in NVT steps. During the NPT equilibration, protein structure constraints were removed to allow relaxation at 300 K. Production runs were performed in NPT conditions with semi-isotropic pressure control, using Parrinello-Rahman barostat; temperature was kept at 300 K using v-rescale thermostat. A cut-off of 11 Å was used to switch off van der Waals interactions, while Particle Mesh Ewald was used to calculate electrostatics of the system, with a spacing of 1.6 Å. Finally, a 2 fs time-step was used to accumulate 3.05 μs of simulated time for each of the five systems, for a total of 15.25 μs. Dumping time was set equal to 100 ps. Our analysis covered the last 3 μs of collected statistics for each of the investigated complexes.

Cluster Analysis

We used a variant of *k-means* algorithm, namely *k-medoids*,[6] as implemented in the BiKi LifeSciences suite.[7] Generally, *k-means* generates an artificial mean structure, identified by coordinates that have minimal sum of squared deviations from a cluster center. The algorithm minimizes a distance-based cost function which is the sum of squared errors (SSE) as defined in Eq. (1):

$$SSE = \sum_{c=1}^{k} \sum_{i=1}^{n} |x_i^c - m_c|^2 \qquad (1)$$

In Eq. (1), k is the number of clusters, n is the total number of conformations, x is the i*th* element of the c*th* cluster and m is the mean of the c*th* cluster. In each step, centroids are randomly chosen, closest objects are grouped around them, and SSE is calculated. Then, new arbitrary k medoids are chosen, clustering is performed again, and the new SSE is compared with the previous one in an iterative fashion until the difference between the previous and the present cycle cannot be further reduced, and medoid positions do not change anymore. While efficient in terms of computational time, *k-means* clustering is sensitive to outliers and to the initialization procedure for the random search of medoids. *K-medoids* is similar to *k-means*, as it is a medoid-centered algorithm, but instead of taking means as the centroid of the cluster, *k-medoids* assigns to centroids a physical meaning, identifying them as real objects in the data set. The new medoids are the most centrally located objects of each cluster. This modification introduces the possibility of returning to minimize the real sum of dissimilarities (distances) between the objects x in a cluster and their medoid m, which is a real representative conformation of the cluster. In other words, in Eq. (1), the difference is an absolute distance from a reference point and not a distance from the mean. Moreover, it overcomes some limitations of the classical *k-medoids* algorithm, which, as *k-means*-like algorithms, randomly select the initial medoids. This procedure affects computational efficiency and makes the results dependent on the choice of k. This *k-medoids* version provides a method to select the initial *k* medoids. The distances d_{ij} are first calculated between every pair of i and j objects, and a distance matrix is created once. Then, a variable for each j object, v_j, is calculated as in Eq. (2):

$$v_j = \sum_{i=1}^{n} \frac{d_{ij}}{\sum_{l=1}^{n} d_{il}} \qquad j = 1, \dots, n \tag{2}$$

The values obtained for each j object are sorted in ascending order and the first k medoids with minimal values of *v* are considered as initial cluster centers. This makes the algorithm deterministic because the initial k medoids are always those that minimize the total distance to all other objects i.e. the most central ones. Also, medoids are updated, finding a new one for each cluster that minimizes the total distance to any other objects in the cluster. The main advantages of this procedure are the ability to work with medoids that can be associated with sampled conformations, and to use an objective function based on absolute distances to refine the quality of the clustering. Moreover, the algorithm is robust to outliers because the most centered conformations are selected as the initial

medoids. In our systems, trajectories were concatenated based only on the common scaffold shared by the five ligands (**Figure S1**). The choice of the representative medoids was performed based on RMSD-based threshold between medoids in a given partition. The number of clusters was considered meaningful of sampling diversity if the difference between the medoids was more than 3Å.

Experimental Methods

Synthesis and Characterization of compound 5

General methods and abbreviations

Abbreviations used in the description of the examples that follow are:

Acetonitrile (MeCN); ammonium chloride (NH4Cl); BnBr (benzyl bromide); carbonyldiimidazole (CDI); cesium carbonate (Cs2CO3); cyclohexane (Cy); chloroform (CHCl3); deuterated chloroform (CDCl3 or Chloroform-d); deuterated dimethylsulfoxyde (DMSO-d6); dichloromethane (DCM); dimethylsulfoxyde (DMSO); N,N-diisopropylethylamine (DIPEA); dimethylacetamide (DMA); 4-(dimethylamino)-pyridine (DMAP); ethylene glycol monomethyl ether (EGME); ethanol (EtOH); electrospray (ES); ethyl acetate (EtOAc); hydrochloric acid (HCl); mass spectrometry (MS); microwave (MW); sulfuric acid (H2SO4); iodomethane (MeI); N,N-dimethylformamide (DMF); lithium hydroxide (LiOH); magnesium sulphate (MgSO4); methanol (MeOH); nuclear magnetic resonance (NMR); room temperature (RT); sodium bicarbonate (NaHCO3); tetrabutylammonium iodide (TBAI); triethylsilane (TES); tetrahydrofurane (THF); thin layer chromatography (TLC); triethylamine (Et3N); trifluoroacetic acid (TFA).

Automated column chromatography purifications were performed using a Teledyne ISCO apparatus (CombiFlashTM Rf) with pre-packed silica gel columns of different sizes (from 4 g to 120 g). Mixtures of increasing polarity of cyclohexane and ethyl acetate or dichloromethane and methanol were used as eluents. Preparative TLCs were performed using Macherey-Nagel pre-coated 0.05 mm TLC plates (SIL G-50 UV254). Hydrogenation reactions were performed using H-CubeTM continuous hydrogenation equipment (SS-reaction line version), with disposable catalyst cartridges (CatCartTM) preloaded with the required heterogeneous catalyst. Microwave heating was performed using ExplorerTM-48 positions instrument (CEM). NMR experiments were run on a Bruker Avance III 400 system (400.13 MHz for 1H, and 100.62 MHz for 13C), equipped with a BBI probe and Z-gradients. Spectra were acquired at 300 K, using deuterated dimethylsulfoxyde (DMSO-d6) or deuterated chloroform (CDCl3) as solvents. Chemical shifts for 1H and 13C spectra were recorded

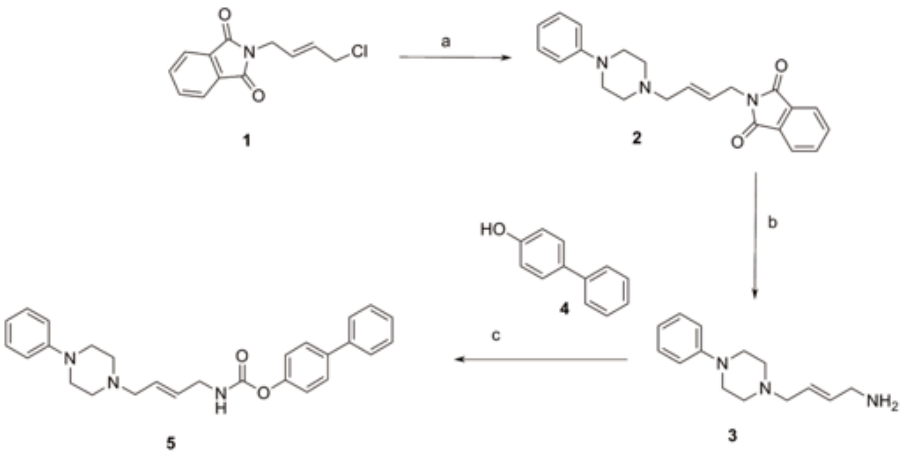
in parts per million using the residual non-deuterated solvent as the internal standard (for CDCl3: 7.26 ppm, 1H and 77.16 ppm, 13C; for DMSO-d6: 2.50 ppm, 1H; 39.52 ppm, 13C).
UPLC/MS analyses were run on a Waters ACQUITY UPLC/MS system consisting of an SQD (Single Quadrupole Detector) Mass Spectrometer equipped with an Electrospray Ionization interface and a Photodiode Array Detector. PDA range was 210-400 nm. Electrospray ionization in positive and negative mode was applied. UPLC mobile phases were: (A) 10mM NH4OAc in H2O, pH 5; (B) 10mM NH4Oac in MeCN/H2O (95:5) pH 5. Analyses were performed either with method A, B, or C, as below reported.

Method A (generic):
Gradient: 5 to 95% B over 3 min. Flow rate 0.5 mL/min. Temp. 40 °C
Pre column: Vanguard BEH C18 (1.7µm 2.1x5mm). Column: BEH C18 (1.7µm 2.1x50mm)

Method B (polar):
Gradient: 0 to 50% B over 3 min. Flow rate 0.5 mL/min. Temp. 40 °C
Pre column: VanGuard HSS T3 C18 (1.7µm 2.1x5 mm). Column HSS T3 (1.8µm 2.1x50mm)

Method C (apolar):
Gradient: 50 to 100% B over 3 min. Flow rate 0.5 mL/min. Temp. 40 °C
Pre column: Vanguard BEH C18 (1.7µm 2.1x5mm). Column: BEH C18 (1.7µm 2.1x50mm)

Synthetic Scheme and General Procedures for the synthesis of compound 5



Scheme S1. a) 1-phenylpiperazine, K_2CO_3, MeCN, reflux, 6 hours. Yield: 90%. b) $NH_2NH_2 \cdot H_2O$, MeOH, then 2N HCl, reflux, 3 hours. Yield: 91%. c) p-nitrophenylchloroformate, DIPEA, DMA:DCM 1:1, rt, 48 hours. Yield: 5%

Step A

A mixture of aryl piperazine (1 eq.), N-(halidealkyl)phthalimide (1.1 eq.) and base (K_2CO_3 or triethylamine, 3 eq.) in acetonitrile was heated to reflux for 6 h. The hot suspension was filtered and the residue washed with acetone several times. The filtrates were concentrated under reduced pressure to give the phthalimide intermediates.

Step B

The phthalimide derivative (1 eq.) and hydrazine hydrate (1.2 eq.) in methanol were heated to reflux for 2 h. To the hot solution was added 2N HCl, and reflux was continued for 1 h more. After cooling to ambient temperature, the mixture was filtered, the residue washed with methanol, and the filtrate evaporated to dryness. This residue was suspended in water and neutralized with 2N NaOH. Extraction with EtOAc afforded an oily product, which was pure enough for the next step.

Step C

The amine derivative (1.0 eq.) was treated with p-nitrophenylchloroformate (1.1 eq.) and DIPEA (1.1 eq.) in a 1:1 mixture of DMA:DCM. The reaction mixture was stirred at ambient temperature for 30 min. To the resulting p-nitrophenyl carbamate solution were added the alcohol derivative (1.25 eq.) and DIPEA (1.1 eq., 2.2 total) and the resultant mixture was stirred at room temperature for 48 h. The desired carbamate was isolated by removal of the undesired p-nitrophenol byproduct and DMA by washing several times with brine and water, then collection and concentration of the organic phase, and purification by flash chromatography (Eluent: 5% MeOH in DCM).

Synthesis of Intermediate 2

The title compound was obtained by applying the general procedure for Step A, using (E)-2-(4-chlorobut-2-enyl)isoindoline-1,3-dione 1 (470 mg, 2.00 mmol), 1-phenylpiperazine (0.28 mL, 1.85 mmol), potassium carbonate (639 mg, 4.62 mmol), and MeCN (7 mL). White solid, 600 mg (90%). UPLC-MS (method A): Rt 2.73 min, m/z 362 $[M+H]^+$. ^1H NMR (400 MHz, Chloroform-d) δ 7.88 (dd, J = 5.4, 3.1 Hz, 2H), 7.79 – 7.70 (m, 2H), 7.31 – 7.25 (m, 2H), 6.99 – 6.90 (m, 2H), 6.87 (td, J =

7.3, 1.1 Hz, 1H), 5.80 – 5.75 (m, 2H), 4.36 – 4.31 (m, 2H), 3.24 – 3.19 (m, 4H), 3.06 (d, J = 4.5 Hz, 2H), 2.60 (dd, J = 6.1, 4.0 Hz, 4H).

Synthesis of Intermediate 3

The title compound was synthesized by applying the general procedure for Step B, using 2-[€-4-[4-phenyl)piperazin-1-yl]but-2-enyl]isoindoline-1,3 dione 2 (600 mg, 1.66 mmol), hydrazine hydrate (0.1 mL, 1.99 mmol) in MeOH (5 mL), then HCl (2 mL). Slightly yellow solid, 350 mg (91%). UPLC-MS (method A): Rt 1.57 min, m/z 232 $[M+H]^+$. ^{1}H NMR (400 MHz, DMSO-*d6*) δ 7.21 (tt, J = 7.2, 1.9 Hz, 2H), 6.95 – 6.89 (m, 2H), 6.80 – 6.74 (m, 1H), 5.70 (dt, J = 15.5, 5.5 Hz, 1H), 5.56 (dt, J = 14.9, 6.6 Hz, 1H), 3.18 (s, 2H), 3.15 (dd, J = 5.4, 1.4 Hz, 2H), 3.12 (m, 4H), 2.95 (dd, J = 6.5, 1.3 Hz, 2H), 2.54 – 2.47 (m, 4H).

Synthesis of the final compound 5

The title compound was synthesized according to the general procedure for Step C, starting from p-nitrophenylchloroformate (192 mg, 0.95 mmol), DIPEA (0.33 mL, 1.90 mmol), (E)-4-(4-phenylpiperazin-1-yl)but-2-en-1-amine 3 (200 mg, 0.86 mmol), and 4-phenylphenol 4 (184.0 mg, 1.08 mmol) in a 1:1 mixture of DMA:DCM (4 mL). White solid 20 mg (5%). UPLC-MS (method A): Rt 1.86 min; m/z 428 $[M-H]^+$. ^{1}H NMR (400 MHz, DMSO-*d6*) δ 8.00 (t, J = 5.8 Hz, 1H), 7.71 – 7.62 (m, 4H), 7.51 – 7.44 (m, 2H), 7.40 – 7.33 (m, 1H), 7.26 – 7.18 (m, 4H), 6.94 (d, J = 8.2 Hz, 2H), 6.78 (t, J = 7.2 Hz, 1H), 5.69 (d, J = 3.3 Hz, 2H), 3.73 (t, J = 4.0 Hz, 2H), 3.14 (m, 4H), 3.06 – 2.95 (m, 2H), 2.52 (m, 4H). ^{13}C NMR (101 MHz, DMSO-*d6*) δ 154.71, 151.51, 151.13, 139.97, 137.41, 130.23, 129.41, 129.35, 128.41, 128.00, 127.79, 127.07, 122.64, 119.20, 115.79, 99.97, 59.91, 52.96, 52.51, 48.64, 42.38.

Supplementary Figures

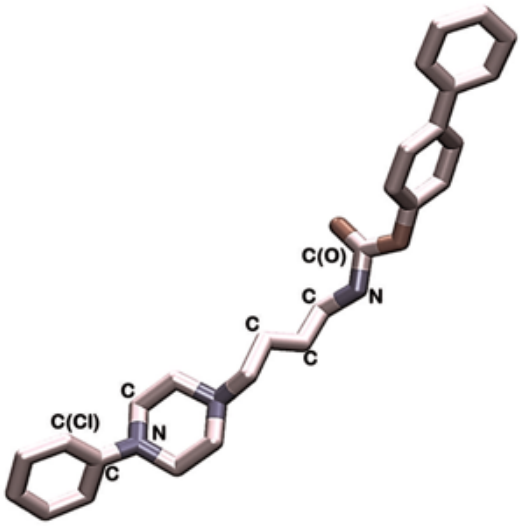


Figure S1. Common scaffold selected for cluster analysis. Key ligand dihedral angles C(Cl)CNC, CCCN, and CCNC(O) are labeled explicitly. Atoms in brackets help to identify the chloride and oxygen atoms bound to the carbon atoms defining the dihedral angles.

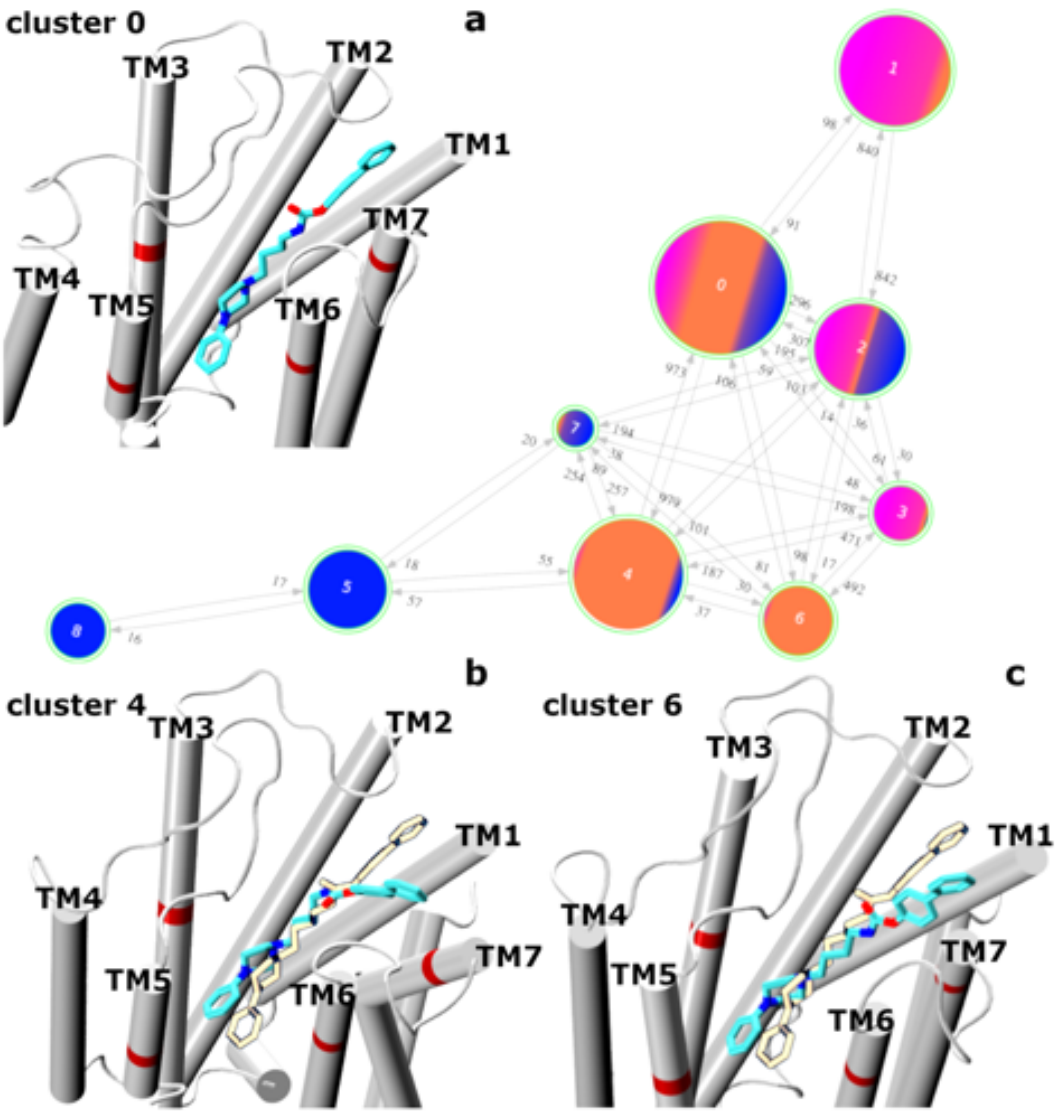


Figure S2. Cluster graph of the conformations visited by compounds **1** (magenta), **3** (orange), and **4** (blue). The nodes are colored by gradients representing the relative enrichment in conformations of the three molecules. In the insets (**a-c**), representative medoids (cyan) of each cluster are shown in complex with their corresponding D3DR conformation and compared to the most populated hub medoid pose (gold) in cluster 0 (**a**). Red circles on TM3, TM5, TM6, and TM7 indicate diagnostic residues D110$^{3.32}$, S193$^{5.43}$, H349$^{6.55}$, and Y365$^{7.35}$, respectively.

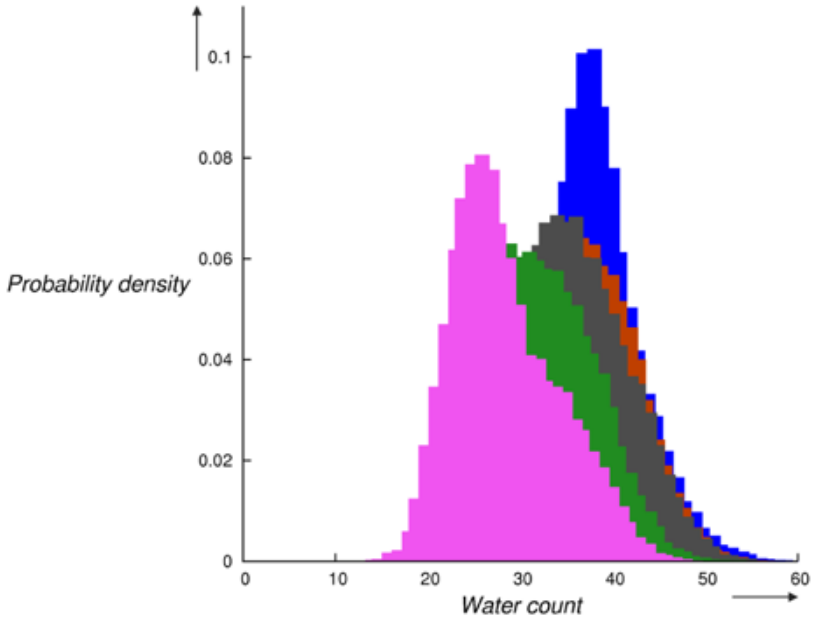


Figure S3. Histograms of the probability distribution for water molecules within 5 Å of compounds **1** (magenta), **2** (green), **3** (orange), **4** (blue), and the predicted ligand **5** (dark grey). See Table S1 for means and standard deviations.

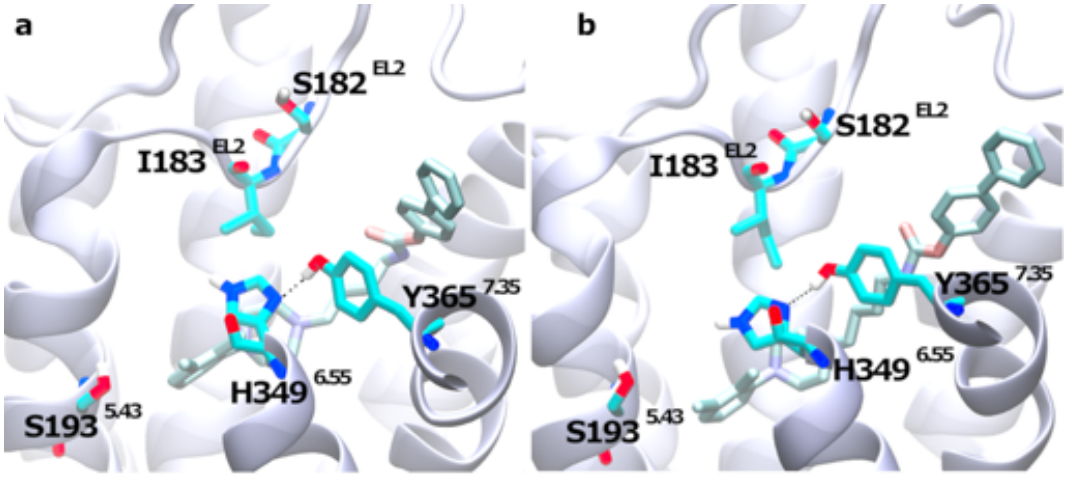


Figure S4. Interaction pattern 1 observed in D3DR by partial agonist **3**. **a**: cluster 4; **b**: cluster 6. EL3 has been removed for clarity.

Supplementary Tables

Table S1. Relative enrichment in representative clusters from analysis on compounds 1, 2, 4

Ligand	1	2	4
Cluster 0	475	7755	776
Cluster 1	10145	0	195
Cluster 2	5437	31	5136
Cluster 3	7624	0	313
Cluster 4	73	10305	434
Cluster 5	0	0	9672
Cluster 6	2	11773	320
Cluster 7	7	126	2005
Cluster 8	0	0	6778
Cluster 9	6244	4387	14

Table S2 Relative enrichment in representative clusters from analysis on compounds 1, 3, 4

Ligand	1	3	4
Cluster 0	4382	7910	4014
Cluster 1	10182	3270	253
Cluster 2	5410	876	4987
Cluster 3	5887	666	239
Cluster 4	2237	9948	1139
Cluster 5	0	1	9620
Cluster 6	1785	6771	21
Cluster 7	121	569	2962
Cluster 8	0	0	6780

Table S3. Number of water molecules within 5 Å of the ligands. Average values over 30000 frames are reported.

Ligand	1	2	3	4	5
Avg. water count	28.6 ± 5.7	32 ± 5.8	36.7 ± 5.6	38.7 ± 4.6	36.1 ± 5.6

Table S4 Relative enrichment in representative clusters from analysis on compounds 1, 5, 4

Ligand	1	5	4
Cluster 0	6545	1495	4670
Cluster 1	10006	26	196
Cluster 2	5578	398	1306
Cluster 3	7652	364	300
Cluster 4	0	9688	173
Cluster 5	0	1682	9679
Cluster 6	2	8429	594
Cluster 7	9	6463	2173
Cluster 8	3	1201	6757

Detailed Analysis on Compound 3

The clustering graph

The clustering graph in **Figure S2** shows nine medoids representative of the second dataset. Despite the presence of selectively enriched clusters (see **Table S2**), the three ligands shared a more similar portion of the sampled space: nodes appeared more homogeneously populated with respect to the first dataset of compounds and a greater number of transitions was observed among clusters. The clustering algorithm was robust enough to center the new clusters on the same representative medoids for compounds **1** and **4**. Therefore, clusters 1, 2, 3, 5, 7 and 8 were reproduced by the analysis on the new dataset and labeled accordingly. In contrast to the first set, two hub conformations were found in the second group, cluster 0 and cluster 2. These hub medoids displayed an extended conformation rather than a bent one. Moreover, cluster 2, which was only visited by **1** and **4** in the first group (**Figure 2** in the main text), here contained 876 poses from **3**, suggesting that the removal of the terminal carboxamide group increased the compound's flexibility in the binding site, thereby allowing the sampling of extended conformations not accessible to **2**. However, in analogy with partial agonist **2**, compound **3** still preferentially clustered into bent conformations. In **Figure S2**, two nodes were interested by an enrichment of partial agonist-selective conformations, consistently called cluster 4 (9948 members) and cluster 6 (6771 members). The medoid of the most populated cluster 4 adopted a bent conformation similar to that observed in **2**. In contrast, cluster 6 was found in a more extended binding mode, where the biphenyl group pointed toward TM7. From the comparison between the two clustering graphs, we can conclude that partial agonist **3**, like **2**, preferred to partition in selective clusters that did not overlap with those preferred by agonist **4** and the antagonist **1**. In the **1-3-4** set of compounds, the overall description of the system as provided by the clustering graph differed only marginally from that of the **1-2-4** set. However, the lack of a hydrogen bond donor/acceptor in the distal ring of the biphenyl group increased the mobility of the ligand along the 3 μs of sampled trajectory. This allowed it to distribute more homogeneously in clusters containing bent and extended poses. Still, the deletion of the carboxamide was not enough to increase the compound's flexibility to agonist-like levels and push **3** toward agonist-selective nodes 5 and 8. Collectively, cluster analysis revealed that ligands with different efficacies sampled specific conformations, which were clustered in exclusive or almost exclusive nodes. Our results are in line with simulative studies on another series of phenylpiperazine derivatives bound at the D3DR binding site, in which adopting specific bent and extended conformations has been associated with partial agonism and antagonism, respectively.[8]

Partial Agonist-Induced Conformations

In agreement with results on compound **2**, the partial agonist **3** showed substantial preference for the bent conformations of cluster 4, albeit it also enriched the more extended conformation represented by cluster 6. Both binding modes prevented the formation of H-bond gating interactions between EL2 and TM6-TM7 (**Figure 5a,c**), induced TM6-TM7 coupling, opened the receptor pocket and induced the interaction pattern 1 **(Figure S4** and **Figure 3a**) with subtle yet interesting differences as compared to **2**. The remarkable tilt induced by **3** in TM7 (cluster 4) drastically increased the number of waters in the site achieving average values close to the full agonist (**Figure S3** and **Table S3**). Furthermore, this partial agonist could also recover agonist-like conformations. For example, it established aromatic interactions between the pendant phenyl ring of the arylpiperazine group and H349$^{6.55}$ during the last 400 ns of simulation (**Figure 3b** in the main text). On the other hand, it could also assume extended conformations with the dichlorophenyl ring parallel to membrane plane, meaning that **3** could also significantly populate antagonist-associated clusters 1 and 3 (see **Table S2** and **Figure S2**). Due to the higher flexibility of **3** in respect of **1**, these antagonist-selective nodes were only transiently populated and receptor closure *via* formation of H-bond gating interactions between H349$^{6.55}$-I183 and Y365$^{7.35}$-S182 could not be triggered (**Figure 5a, c**).

The lack of the carboxamide substituent in the distal phenyl ring of the biphenyl group induced the ligand to drift deeper into the pocket for most of the simulated time, thus, losing the aromatic contact between the ligand dichlorophenyl ring of **3** and H349$^{6.55}$ (**Figure 3b**). In contrast to **2**, in specific clusters 4 and 6, **3** led F346$^{6.52}$ side chain to shift toward TM5-TM6 interface. In another series of recently studied piperazine derivatives in complex with D3DR,[8] a similar rearrangment has been associated to TM6 inward motion and partial D3DR activation.

Interaction pattern 2 (**Figure 3c**) was never established. This accounts for the limited efficacy of this ligand.

References

[1] D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, Walker, J. Wang, R. M. Wolf, X. Wu, P. A. Kollman, AMBER14, University of San Francisco, California. 2014.

[2] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* 2015, *11*, 3696-3713.

[3] C. J. Dickson, B. D. Madej, Å. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould, R. C. Walker, *J. Chem. Theory Comput.* 2014, *10*, 865-879.

[4] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* 2004, *25*, 1157-1174.

[5] B. Hess, C. Kutzner, D. Van Der Spoel, E. Lindahl, *J. Chem. Theory Comput.* 2008, *4*, 435-447.

[6] H. S. Park, C. H. Jun, *Expert Syst. Appl.* 2009, *36*, 3336-3341.

[7] S. Decherchi, G. Bottegoni, A. Spitaleri, W. Rocchia, A. Cavalli, *J. Chem. Inf. Model.* **2018**, *58*, 219-224.

[8] M. Michino, C. A. Boateng, P. Donthamsetti, H. Yano, O. M. Bakare, A. Bonifazi, M. P. Ellenberger, T. M. Keck, V. Kumar, C. Zhu, R. Verma, J. R. Deschamps, J. A. Javitch, A. H. Newman, L. Shi, *J. Med. Chem.* 2017, *60*, 580-593.