JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

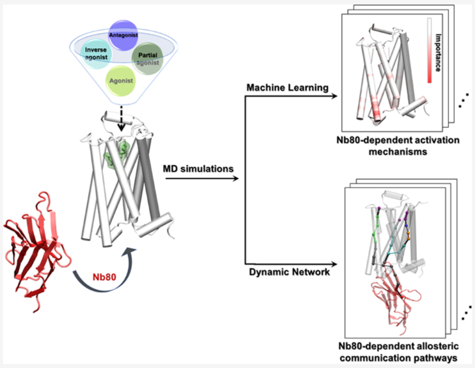
pubs.acs.org/jcim Article

Allosteric Effect of Nanobody Binding on Ligand-Specific Active States of the β2 Adrenergic Receptor

Yue Chen, Oliver Fleetwood, Sergio Pérez-Conesa, and Lucie Delemotte*

Cite This: *J. Chem. Inf. Model.* 2021, 61, 6024−6037 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



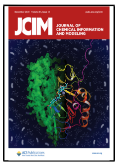
ABSTRACT: Nanobody binding stabilizes G-protein-coupled receptors (GPCR) in a fully active state and modulates their affinity for bound ligands. However, the atomic-level basis for this allosteric regulation remains elusive. Here, we investigate the conformational changes induced by the binding of a nanobody (Nb80) on the active-like β2 adrenergic receptor (β2AR) via enhanced sampling molecular dynamics simulations. Dimensionality reduction analysis shows that Nb80 stabilizes structural features of the β2AR with an ∼14 Å outward movement of transmembrane helix 6 and a close proximity of transmembrane (TM) helices 5 and 7, and favors the fully active-like conformation of the receptor, independent of ligand binding, in contrast to the conditions under which no intracellular binding partner is bound, in which case the receptor is only stabilized in an intermediate-active state. This activation is supported by the residues located at hotspots located on TMs 5, 6, and 7, as shown by supervised machine learning methods. Besides, ligand-specific subtle differences in the conformations assumed by intracellular loop 2 and extracellular loop 2 are captured from the trajectories of various ligand-bound receptors in the presence of Nb80. Dynamic network analysis further reveals that Nb80 binding triggers tighter and stronger local communication networks between the Nb80 and the ligand-binding sites, primarily involving residues around ICL2 and the intracellular end of TM3, TM5, TM6, as well as ECL2, ECL3, and the extracellular ends of TM6 and TM7. In particular, we identify unique allosteric signal transmission mechanisms between the Nb80-binding site and the extracellular domains in conformations modulated by a full agonist, BI167107, and a G-protein-biased partial agonist, salmeterol, involving mainly TM1 and TM2, and TM5, respectively. Altogether, our results provide insights into the effect of intracellular binding partners on the GPCR activation mechanism, which should be taken into account in structure-based drug discovery.

■ INTRODUCTION

The G-protein-coupled receptor (GPCR) superfamily is the largest and most distinct group of membrane receptors in eukaryotes, comprising over 800 diverse human cell-surface receptors.[1] They mediate signaling of a variety of extracellular stimuli, including photons, odorants, hormones, peptides, and proteins, and regulate many physiological processes. Not surprisingly, GPCRs are important targets for the binding of drugs, which account for ∼34% of all US Food and Drug Administration (FDA)-approved medicines, highlighting the indispensable role of GPCRs in health and disease.[2,3] All GPCRs share a common seven-transmembrane (7TM) domain helices architecture. They show differences in their extracellular domains, where extracellular ligands bind, and in their intracellular domains, where signaling transducers, such as G-proteins and β-arrestins, bind.[4,5] Upon activation by extracellular ligands, the receptor undergoes certain conformational changes and engages intracellular transducers, which modulates different downstream signaling pathways.

GPCR activation is an allosteric process, involving transducing a signal initiated by various external stimuli into cellular response and downstream regulation of various aspects of human physiology. Therefore, understanding the mechanism underlying the allosteric signaling of GPCRs is of importance for drug discovery and pharmacology research. Several highly conserved residues on the pathway connecting the ligand-binding and the G-protein-binding pockets have been identified by experimental and computational studies.[6,7] These residues are organized in microscopic clusters, often referred to as microswitches. Their dynamics play an important role in GPCR activation. For example, the outward movement of transmembrane helix 6 (TM6), located at the cytoplasmic domain of the receptor, is a hallmark of GPCR activation. Some evolutionarily conserved sequence motifs are also identified as microswitches, such as $N^{7.49}P^{7.50}xxY^{7.53}$ (superscripts referring to Ballesteros−Weinstein numbering[8]), $D^{3.49}R^{3.50}Y^{3.51}$, $P^{5.50}I^{3.40}F^{6.44}$, $C^{6.47}W^{6.48}xP^{6.50}$, and the sodium-binding pocket at $D^{2.50}$, distributed in the TM domains.[6,7] In

Received: July 12, 2021
Published: November 15, 2021



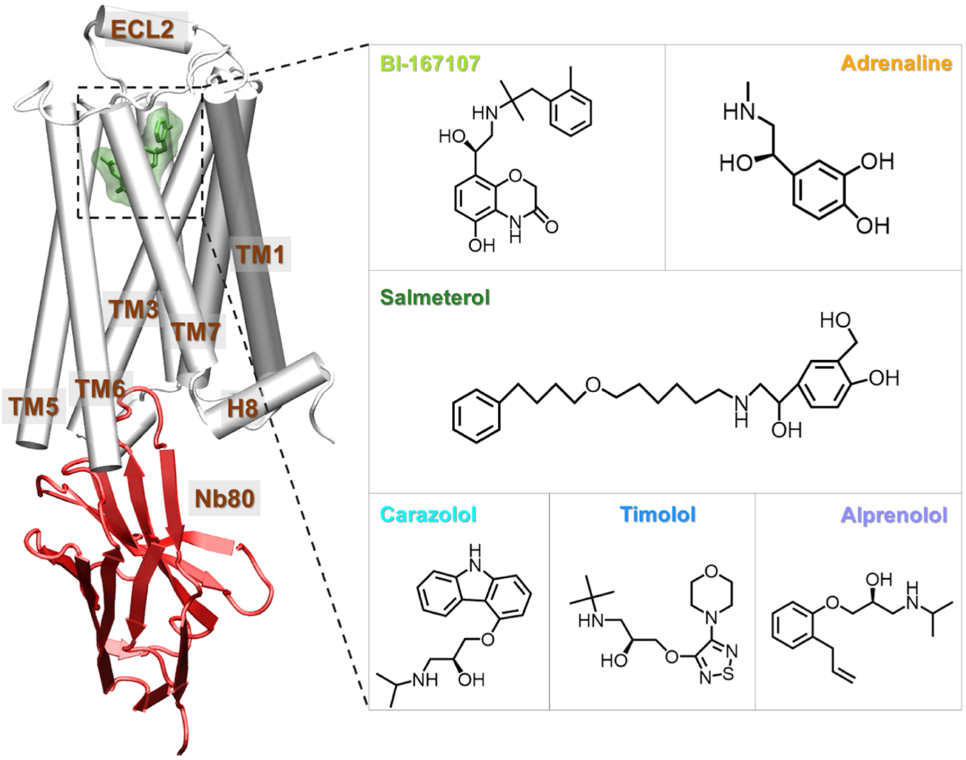


Figure 1. Structure of the β2AR-Nb80: A molecular dynamics snapshot of BI167107-bound β2AR with Nb80 in the active-like state (simulation starting structure: PDB 3P0G[26]) and ligands examined in this study: agonists BI167107 and adrenaline; biased partial agonist salmeterol; inverse agonist carazolol, and antagonists timolol and alprenolol. The receptor is represented as white cartoon, Nb80 as red ribbons, the bound ligand as green sticks, and transparent surface.

addition, great efforts have also been made to characterize the mechanism underlying the G-protein activation,[9,10] biased signaling,[11,12] and allosteric modulation[13,14] via diverse approaches.[15−18] Several studies, using, in particular, various spectroscopic techniques, have pointed out that a simple two-state model involving a single inactive and a single active state is an oversimplification and that the activation mechanism instead involves multiple inactive, intermediate, and active receptor states.[19,20] Many studies have focused on the ligand-dependent conformational changes implicated in the enhancement in the binding affinity of intracellular transducers.[21−24] However, the structural basis underlying transducer-induced allosteric communications and how they are related to ligand efficacy is not fully understood.

Previously, Fleetwood et al.[25] focused on analyzing conformational ensembles of the β2 adrenergic receptor (β2AR) modeled in the absence of intracellular binding partner and revealed that ligands with varying efficacy profiles could stabilize different intermediate-active states of the receptor using enhanced sampling molecular dynamics (MD) simulations coupled with data-driven methods. Here, we build on this work and investigate the structural changes induced by G-protein-mimicking Nanobody80 (Nb80)[26] to the β2AR bound with six ligands with different efficacies (Figure 1). This nanobody has been designed, and used, to stabilize the receptor approaching a fully active state: the region of Nb80 that comes in interaction with the receptor indeed mimics the interactions made with a nucleotide-free G-protein. We first sampled the different active-like ensembles of unliganded and ligand-bound β2AR in the absence and presence of Nb80 using MD simulations. Using dimensionality reduction analysis, we found that Nb80 binding stabilizes a highly active-like state with a 12−14 Å outward movement of TM6 independent of ligand binding. More specifically, BI167107 (full agonist)[26] and salmeterol (G-protein-biased partial agonist)[27] generate different subtle conformational distributions, compared to the other ligands. In addition to the intracellular end TM6 microswitch, specific residues in TM3, TM5, and TM7 are identified as important features for distinguishing Nb80-bound and N80-free states. In the presence of Nb80, ligand-specific conformational differences mainly show up in the ECL and ICL domains. Furthermore, dynamic network analysis reveals that communication across the receptor is greatly strengthened when binding to Nb80. Interestingly, BI167107- and salmeterol-specific optimal signal transmission pathways from the Nb80-binding site to the ligand-binding site primarily involve TM1 and TM2, and TM5, respectively. Taken together, our findings provide a structural basis for the enhancement of ligand affinity and ligand-specific effects on the receptor activation, controlled by intracellular binding partners, a phenomenon that should be taken into account during structure-based drug discovery of, for example, biased agonists for GPCRs.

■ METHODS

Molecular Simulations System Setup. We based all β2AR simulation systems on the fully active state structure 3P0G,[26] which is bound to BI167107 and Nb80. The nanobody-bound systems had the same system configuration and followed the same equilibration protocol as the previously published nanobody-free simulations.[28,29] All simulations were initiated with CHARMM-GUI[30] and used the CHARMM36m force field.[31] To account for missing residues and mutations present in the experimental structure, we reversed the N187E mutation and capped chain termini with acetyl and methylamide. The ligands included in this study are all resolved in the different β2AR PDB structures 2RH1,[32] 3NYA,[33] 3D4S,[34]6MXT,[35] and 4LDO.[36] Due to their localization in hydrophobic environments, E122$^{3.41}$ and the ligands' amine groups were protonated, while H172$^{4.64}$ and H178$^{4.70}$ were protonated at their epsilon positions. With a complete model of the β2AR and the Nb80, the protein complex was embedded in a homogeneous POPC[37] lipid bilayer, and surrounded by a solution consisting of TIP3P water molecules[38] with a 0.15 M concentration of sodium and chloride ions. For the nanobody-bound receptor, we inserted 190 membrane molecules and 120 water molecules per lipid. In the smaller nanobody-free systems, we used 180 membrane molecules and 79 water molecules per lipid. We performed the MD simulations with GROMACS 2018.6.[39]

The systems were energetically minimized with the steepest descent algorithm, then particle velocities were generated at a 310.15 K temperature. Next, the systems were equilibrated following a six-step protocol under decreasing positional constraints, the first three simulations with 1 fs timestep for a total of 125 ps of simulation time, and the last three steps with a 2 fs timestep for a total of 500 ps simulation length, following CHARMM-GUI's default protocol. We used the LINCS algorithm to constrain hydrogen bonds. To control temperature and pressure, we used a Nosé−Hoover thermostat with a 1 ps time constant, and a semi-isotropic Parrinello−Rahman barostat with a time constant of 5 ps, and a compressibility of 4.5×10^{-5}/bar for production runs, whereas the initial equilibration used a Berendsen thermostat and barostat. Long-range effects were handled with the fast smooth particle-mesh Ewald (SPME) electrostatics and a Verlet list for neighbor searching. Input files and simulation trajectories are available publicly online (https://osf.io/b5rav/).[28,29]

Single-State Sampling Simulations. The conformational ensembles were sampled using kinetically trapped active-like state sampling, or single state sampling, a recently published enhanced sampling technique.[29] In this relatively simple framework, 24 simulation replicas, each of 7.5 ns length, were launched from the starting structure (Table 1). Their center point, c, was computed in a high-dimensional space spanned by a set of collective variables (CVs), which have previously been shown to well characterize the β2AR's activation mechanism.[29] For every replica, i, we computed its distance to the center, x_i, and the average replica distance to the center, d. A weight, $w_i(x) = \mathrm{e}^{-\frac{|x_i - c|^2}{d}}$, was assigned to every replica. For the next iteration of the method, the replicas were extended, with the number of copies proportional to $\frac{w_i}{\sum_j w_j}$, keeping the total replica count at 24. By performing these steps iteratively, the replicas eventually diffused around a well-equilibrated state. In other words, for every ligand−receptor complex, we obtained an ensemble of structures sampled from the closest kinetically stable state accessible from the active starting structure. In line with convergence analysis performed on the original dataset of nanobody-free simulations,[29] we monitored the distance between the center points of subsequent iterations. Convergence was obtained when the distance between center points was smaller than the standard error of the replicas' distance to the centers. Finally, the trajectories of the last iteration were further analyzed, as described in the next section.

Table 1. Total Simulation Time Per System

system	ligand	number of iterations	simulation time (μs)
apo-β2AR		8	1.44
BI167107-β2AR	BI167107	8	1.44
adrenaline-β2AR	adrenaline	8	1.44
salmeterol-β2AR	salmeterol	8	1.44
carazolol-β2AR	carazolol	8	1.44
timolol-β2AR	timolol	8	1.44
alprenolol-β2AR	alprenolol	8	1.44
apo-β2AR-Nb80		8	1.44
BI167107-β2AR-Nb80	BI167107	8	1.44
adrenaline-β2AR-Nb80	adrenaline	8	1.44
salmeterol-β2AR-Nb80	salmeterol	8	1.44
carazolol-β2AR-Nb80	carazolol	8	1.44
timolol-β2AR-Nb80	timolol	8	1.44
alprenolol-β2AR-Nb80	alprenolol	8	1.44

Dimensionality Reduction Methods. We derived the active-like conformational ensemble from the last iteration of the swarms. To project the ensemble onto a lower-dimensional manifold, we used two different dimensionality reduction methods, principal component analysis (PCA)[40] and multidimensional scaling (MDS).[41]

PCA is one of the most used techniques for dimensionality reduction. It projects the data on principal components in the linear regime by computing the eigenvectors of the data covariance matrix. The first principal component represents the dimension accounting for the most variance of the data, and the subsequent ones account for decreasing amounts of variance in their respective dimensions. MDS is a nonlinear method that includes various multivariate data analysis techniques. It is developed to construct a set of low-dimensional embedding patterns which best preserve pairwise Euclidean distances in the original high-dimensional space.

In this work, we use the PCA and MDS modules of the python package Scikit-learn.[42] The inverse closest-heavy atom distances were used as input features and the simulation snapshots were then projected onto the first four-dimensional feature spaces.

Supervised and Unsupervised Feature Extraction and Learning. We used the supervised and unsupervised feature extraction module implemented in Demystifying,[25] which aims to identify molecular features that are important for a specific biological question. An artificial feed-forward neural network, a multilayer perceptron (MLP) classifier,[43,44] was first trained to find the important residues for discriminating β2AR systems in the absence and presence of Nb80. Another MLP was subsequently trained to distinguish all of the Nb80-bound systems bound to different ligands. Inverse closest-heavy atom distances were used as input features, of which the importance was normalized. Layerwise relevance propagation (LRP)[45] was then applied to the trained network to rank the importance of every input feature for classification. With these approaches, we obtained the importance of every protein residue for distinguishing Nb80-bound and -unbound conformational ensembles, and for distinguishing ensembles bound to different

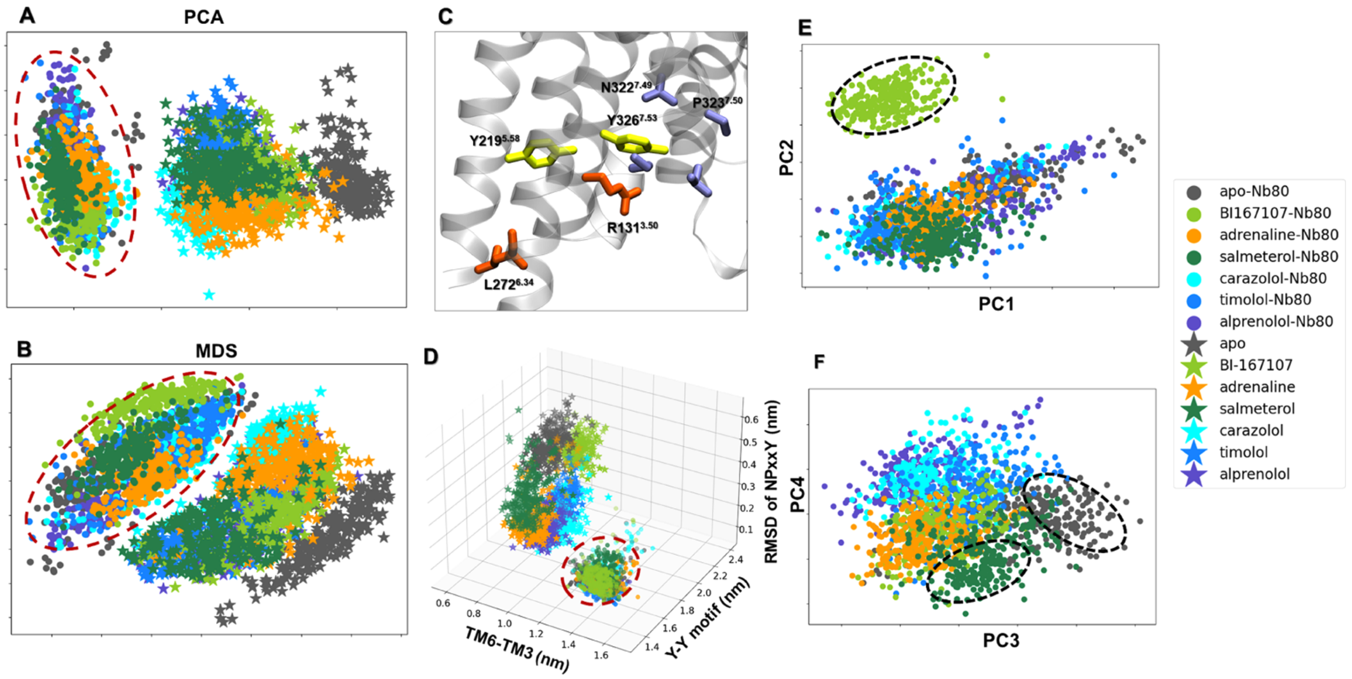


Figure 2. Dimensionality reduction analysis applied to the active-like simulation ensembles. Each point represents a snapshot and is depicted according to the ligand and Nb80-bound ensembles. Input features are the residue−residue Cα atom distances. The Nb80-bound ensembles are highlighted by a red dashed circle. (A) Principal component analysis (PCA) and (B) Multidimensional scaling (MDS) projection on the first two principal components extracted from all trajectories (combining trajectories with and without Nb80). (C) Conserved microswitches of the β2AR: R131 and L272 (orange) are located in the transmembrane 3 (TM3) and TM6, respectively. The outward displacement of TM6 is represented by the distance between the Cα atoms of R131$^{3.50}$ and L272$^{6.34}$. Y219$^{5.58}$ and Y326$^{7.53}$ (yellow) are part of TM5 and TM7, which are close to each other via a water-mediated interaction (Y-Y motif) in the β2AR active state. The N$^{7.49}$P$^{7.50}$xxY$^{7.53}$ motif (blue) is at the bottom of TM7. (D) Distributions of the distances between TM6 and TM3, and Y-Y and RMSD of NPxxY motifs. (E, F) PCA projection onto the first four principal components (PC) of the Nb80-bound trajectories only. BI167107-, salmeterol-bound, and apo snapshots are highlighted by black dashed lines.

ligands. In that way, we could identify which residues were most affected by the binding of the nanobody, and by the binding of different ligands, respectively. As a control, we also calculated the Kullback−Leibler (KL)[25,46] divergence to derive the important residues. In the KL divergence calculation, high divergences represent nonoverlapping inverse distance distributions in active-like states, highlighting thereby features important to distinguish the conformational ensembles. In addition, unsupervised learning PCA[40] was performed to capture important residues distinguishing β2AR ensembles with and without Nb80. The input distances were transformed into a set of principal components (PCs) by calculating the eigenvectors of the corresponding covariance matrix. We then estimated the importance of individual distances contributing to the PCs by multiplying the PCs with their eigenvalues and projecting them back onto the input features.[25]

Dynamic Network Analysis. Dynamic network analysis for the β2AR in the absence and presence of Nb80 was performed using the *NetworkView* plugin in VMD.[47,48] For each system, a network map with each protein residue defined as a node was generated. Edges were added between pairs of "in-contact" nodes whose heavy atoms interacted within 4.5 Å for more than 75% of the simulation time. Each edge is weighted by the correction values of two end nodes using the equation: $w_{ij} = -\log(|C_{ij}|)$, in which w_{ij} and C_{ij} are the weight and correlation values, respectively. The weight of an edge represents the potential for information transfer (betweenness) between two nodes, where a stronger cross-correlation results in a higher weight, then represented as a thicker edge. Each network was divided into communities of nodes with highly frequent and strong connection to each other using Girvan−Newman algorithm.[49] Critical nodes that connect the neighboring communities were then identified. The pathways describing the communication between Nb80 and the ligand-binding pocket were also identified based on edge betweenness derived from the correlation of nodes. The Floyd−Warshall algorithm[50] was used to determine the optimal path between two given nodes: a source and a sink. In general, after the source and sink are chosen, the optimal path is defined to be the connecting route between the two nodes (residues), which minimizes the number of intermediate nodes and maximizes the sum of edge betweenness of the connecting route. In addition, using the toolkit *subopt*, we identified suboptimal paths, *i.e.*, paths that are slightly longer than the optimal path.

■ RESULTS AND DISCUSSION

Global Structural Features Derived from Data-Driven Analysis. We used an adaptive sampling protocol to quantitatively sample the most stabilized active-like states of unliganded and ligand-bound β2AR in the absence and presence of Nb80 (Figure 1 and Table 1). We ran the sampling method for eight iterations, at which point the center points did not drift between iterations (Figure S1). These conformations are kinetically accessible from the initial fully active structure (PDB ID 3P0G[26]) and represent snapshots of the protein complex in the fully active-like and intermediate-active states. The ligands studied include the full agonists BI167107 and adrenaline, the G-protein-biased partial agonist salmeterol, the antagonists alprenolol and timolol, and the inverse agonist carazolol. To better understand the receptor

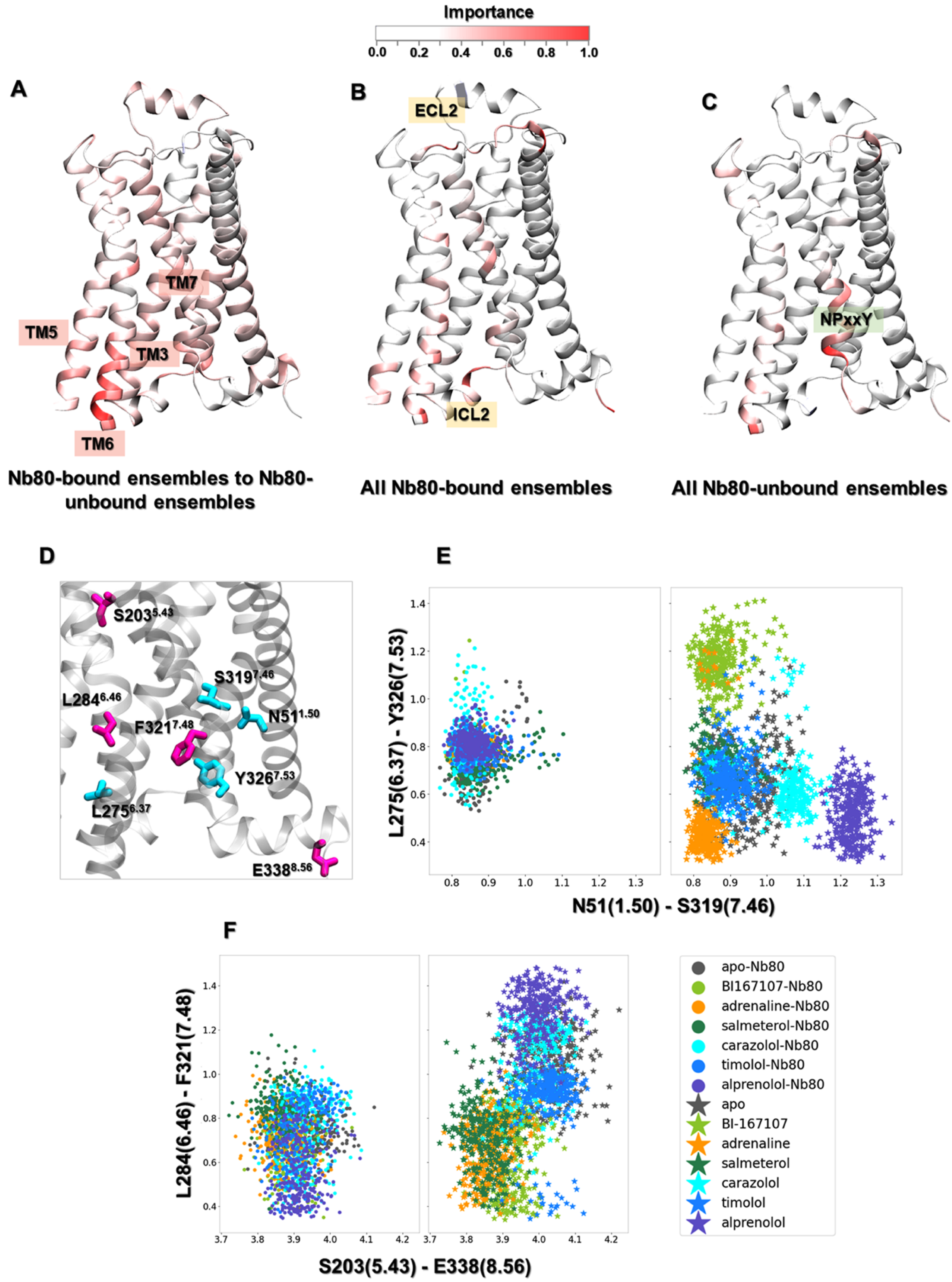


Figure 3. Residues important for discriminating Nb80-dependent activation mechanisms, derived from training a multilayer perceptron (MLP) classifier on equilibrated active-like ensembles. (A) Comparison between Nb80-bound and Nb80-unbound ensembles. The most important hotspot is located at the end of TM6. (B) Residues important to distinguish the Nb80-bound ensembles. (C) Residues important to distinguish the Nb80-unbound ensembles. (D) Residues important to differentiate Nb80-bound from Nb80-unbound ensembles. (E) Distribution of N51$^{1.50}$- S319$^{7.46}$ and L275$^{6.37}$-Y326$^{7.53}$ distances. (F) Distribution of S203$^{5.43}$-E338$^{8.56}$ and L284$^{6.46}$-F321$^{7.48}$ distances.

conformational changes triggered by binding of Nb80 and various ligands, we performed two different dimensionality reduction analyses: principal component analysis (PCA) and multidimensional scaling (MDS). This allowed us to project the results onto a low-dimensional space and clearly visualize overlap between the conformational ensembles. Each point represents a simulation snapshot, which is colored and marked according to the bound ligand and whether Nb80 is present, respectively (Figure 2).

Nb80-Stabilized Changes in the β2AR. The PCA and MDS analyses of the whole dataset (encompassing the trajectories with and without Nb80) revealed that the Nb80-bound and -unbound states are grouped into two distinct clusters (Figure 2A,B). The absence of overlap between the two clusters for both dimensional reduction methods suggests that Nb80 binding indeed induces conformational changes of the receptor that are independent of ligand binding. In addition, MD snapshots with Nb80 bound tended to be grouped together more compactly than those without Nb80, possibly implying a higher structural rigidity. Residue importance derived from PCA identified that part of the intracellular end of TM6 and TM7 contribute significantly to the different conformational distribution of Nb80-bound and -unbound β2AR ensembles in Figure 2A (Figure S2). To further illustrate the Nb80-induced conformational alterations, we further analyzed some traditional microswitches and measured the outward displacement of transmembrane helix 6 (TM6), the twist of the N$^{7.49}$ P$^{7.50}$ xxY$^{7.53}$ motif, and a water-mediated interaction between Y219$^{5.58}$ and Y326$^{7.53}$ (Y-Y motif), which all represent hallmarks of β2AR activation (Figures 2D and S3−S6). Here, the TM6 displacement was measured by the distance of Cα atoms between R131$^{3.50}$ and L272$^{6.34}$ (TM6-TM3), and Y-Y motif was represented by the closest-heavy atom distance between Y219$^{5.58}$ and Y326$^{7.53}$ (Figure 2C). As shown in Figure 2D, the Nb80-bound ensembles grouped together, away from the Nb80-free states in the 3D distribution space of the distances of TM6-TM3, Y-Y motif, and RMSD of NPxxY motif, in agreement with the dimensionality reduction results. Furthermore, for the intermediate-active ensembles without Nb80, binding of different ligands resulted in distinct structural spaces. In contrast, all of the snapshots with Nb80 occupied a similar region of the conformational space and gathered into a single cluster represented by an increase in the TM6-TM3 distance, a decrease in the distance of the Y-Y motif, and a decrease in the RMSD of the NPxxY motif. The changes in microswitches indicate that the receptor approaches the fully active state in the simulations. This is consistent with an Nb80-mediated enhancement of receptor activation, also reported in previous work.[20] Our analysis suggests that Nb80 binding triggers conformational changes in the receptor and favors fully active-like conformations, independent of the binding of various ligands. Similar effects were also observed in other GPCRs, such as angiotensin II type 1 receptor (AT1R), adenosine A2A receptor, and β1-adrenergic receptor.[51−53] However, no overlap in the projected conformational distribution does not mean that Nb80-bound snapshots are completely different from those without Nb80. There could indeed be overlap in a different projection space. Therefore, we further investigated the conformational space of the third and fourth PCs for all of the ensembles. The result indicates that all simulation snapshots still share common features, whether the Nb80 bound or not (Figure S7A,B).

Ligand-Dependent Stabilization of β2AR-Nb80 States. The first two-dimensional projections from PCA and MDS as well as the distribution of microswitches highlight the Nb80-induced effects on the receptor conformation, while ligand-mediated structural changes are not resolved in this subspace. To capture the conformational differences between the receptor bound to various ligands, we carried out the same dimensionality reduction methods on the Nb80-bound conformational ensembles only, and projected the conformational ensemble on the first four components, resulting in a different separation of the data (Figures 2E,F and S7C,D). In Figure 2E, BI167107, a full agonist with ultrahigh affinity to β2AR, segregates away from other ligands, which cluster together. This implies that the binding of BI167107 induces specific conformational changes in the Nb80-bound receptor. At the same time, the third and fourth components in the projection show many BI167107-bound snapshots sharing a similar conformational distribution with the others (Figure 2F). In addition, we find that most of the unliganded and salmeterol-bound snapshots deviate from the group at the center, indicating that different states are assumed for the two systems (Figure 2F).

These results, together with the microswitch conformational distribution, suggest that Nb80 binding promoted all simulation ensembles to share overall features of a fully active state, but the unliganded, BI167107 and salmeterol stabilized unique activation features. This is in agreement with previous experimental results, supporting the notion that small ligand-specific conformational changes contribute to different receptor activation and downstream signals.[54,55]

Nb80 and Ligand-Induced Local Structural Changes. Unsupervised data-driven analysis can provide insights into overall conformational differences of the receptor bound to different ligands in the absence and presence of Nb80, but fails to reveal specific Nb80- and ligand-induced activation signatures. To capture important features of receptor activation among fully active-like states controlled by Nb80 and ligands, we decided to resort to supervised learning methods. We trained classifiers to learn differences between simulation trajectory datasets, using as input inverse inter-residue Cα distances. With this approach, we derive residues that are important to distinguish different receptor-ligand-Nb80 systems, with the idea that these residues could play a substantial role in the receptor activation.

Nb80-Specific Local Conformational Changes. Importance profiles were calculated using layerwise relevance propagation on a multilayer perceptron (MLP) classifier trained to distinguish Nb80-bound and -unbound states (Figure 3A), as well as states modeled in the presence of different ligands from one another, in the presence and in the absence of Nb80 (Figure 3B,C). As a control, we also characterized important features by computing the Kullback−Leibler (KL) divergence, where residues with high KL divergences are defined as important features (Figure S8). Compared to KL divergence, the MLP classifier generates importance profiles with more peaks, as it can find all important features by performing nonlinear transformations of input features.[25] We observed that both methods identified the cytoplasmic end of TM6 as the most important region to discriminate states with and without Nb80 (Figures 3A and S8A). Recent studies have identified multiple inactive, intermediate and active receptor states with different degrees of conformational changes at the intracellular end of TM6, in

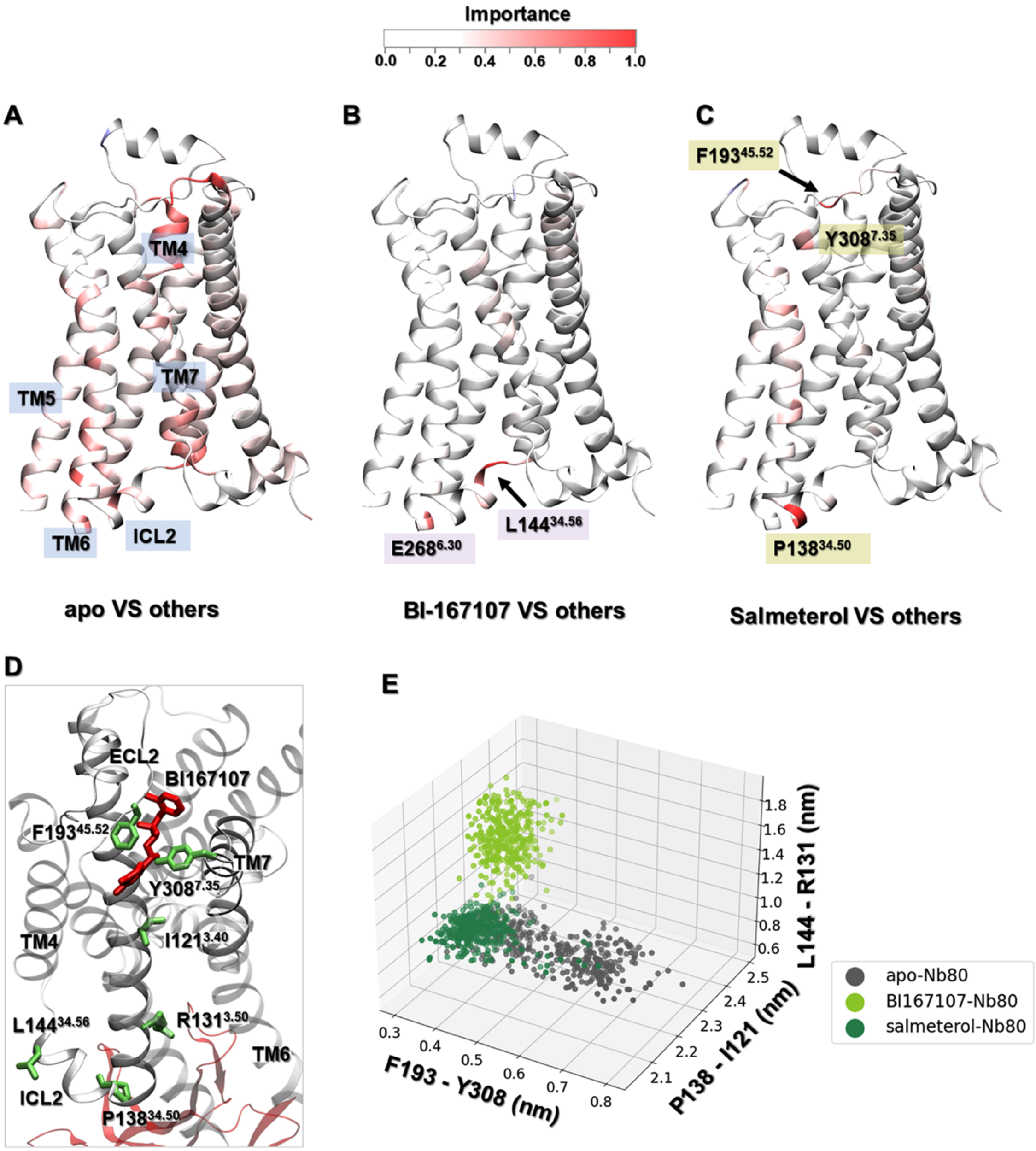


Figure 4. Important residues derived from the equilibrated active-like ensembles for discriminating ligand-dependent activation mechanisms using a multilayer perceptron (MLP) classifier. (A−C) Comparison of the apo, BI167107-, and salmeterol-bound ensembles to the others, respectively, in the presence of Nb80. (D) Important residues for differentiating apo, BI167107-, and salmeterol-bound β2AR-Nb80 ensembles. (E) Distances distribution between F193$^{45.52}$−Y308$^{7.35}$, P138$^{34.50}$−I121$^{3.40}$, and L144$^{34.56}$−R131$^{3.50}$ in apo, BI167107- and salmeterol-bound β2AR-Nb80 ensembles.

which complete receptor activation accompanied by an ∼14 Å outward movement of TM6 requires both agonist and G-protein or a mimetic nanobody such as Nb80.[20,56] Notably, this region does not differentiate the ligand-bound receptor modeled in the presence or absence of Nb80 (Figure 3B,C), suggesting that the cytoplasmic end of TM6 conformations are very similar within these two classes. This is also compatible with the conformational distribution of microswitches (Figure 2D). In addition, the MLP classifier also highlighted some residues on TM3, TM5, and TM7, which exhibited different conformations in the Nb80-bound and -unbound states.

In contrast, there are only a few identified hotspots for discriminating all β2AR-Nb80 complexes, illustrating that all of the bound ligands stabilized common structural rearrangements (Figure 3B). Among the few regions distinguishing ligand-bound ensembles, a few residues in the intracellular loop (ICL) 2 and extracellular loop (ECL) 2 showed up as important when comparing receptor ensembles when bound to the various ligands in the presence of Nb80. In agreement with this, several studies point out that ECL2 is involved in ligand specificity, and in determining the affinity of ligands toward the receptor.[57] Moreover, it should be stressed that ICL2 directly interacts with the N-terminus of G-protein and is responsible for the selectivity of receptor-G-protein interactions as well as the efficiency of G-protein activation.[58−60] Besides, for all receptor states without Nb80 bound, the NPxxY motif exhibited a ligand-specific conformation, in agreement with our previous study[29] (Figure 3C). In contrast, the NPxxY motif

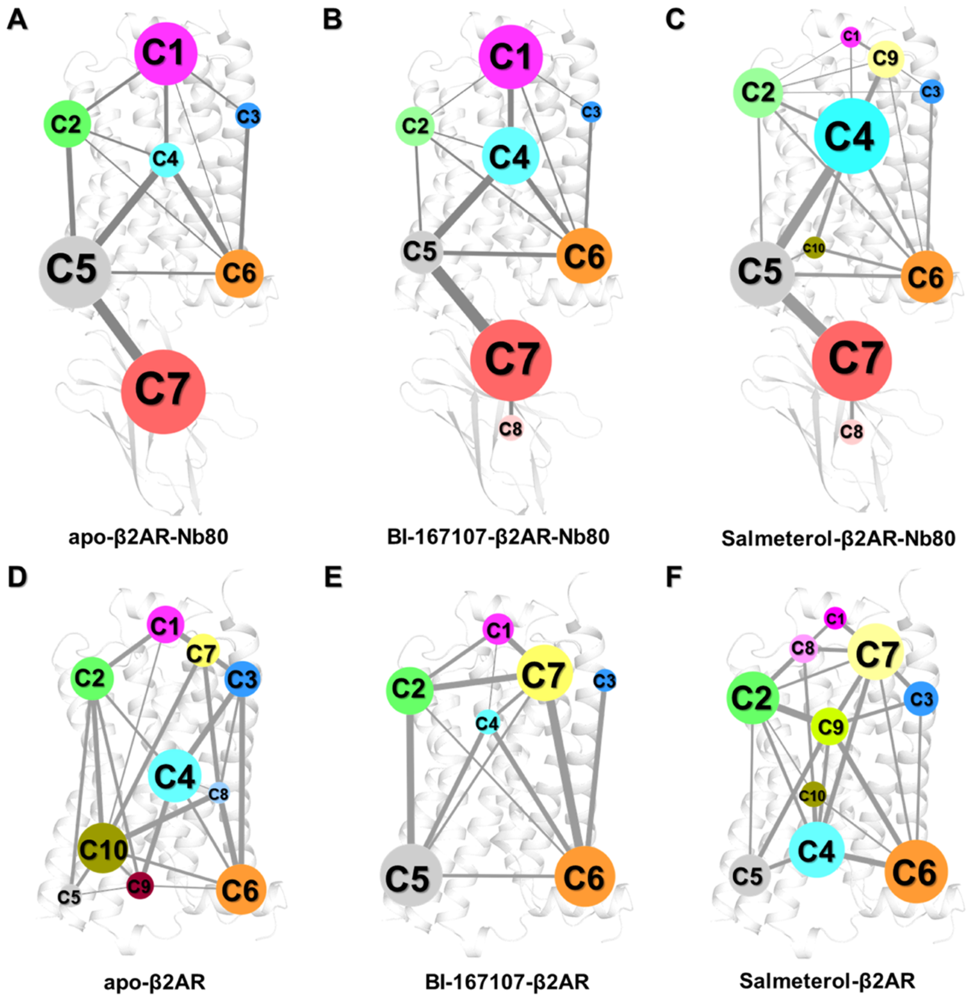


Figure 5. Dynamic networks of the apo, BI167107-, and salmeterol-bound *β*2AR with and without Nb80 bound are analyzed using community network analysis. (A−C) Two-dimensional (2D) networks of unliganded, BI167107- and salmeterol-bound forms with Nb80. (D−F) Two-dimensional (2D) networks of apo, BI167107-, and salmeterol-bound forms without Nb80. Network communities are colored separately according to their ID number. A community represents a set of highly intra-connected nodes (residues), its size being determined by the number of nodes included in a community. Edges connecting two communities are represented by lines, of which the width is proportional to the strength of the information flow between the connected communities.

adopts a similar conformation for all ligand-bound ensembles in the presence of Nb80 (Figures 3B and 2D).

In addition to conformational differences in the cytoplasmic region induced by Nb80 binding, structural changes through the TM domain were also captured by this analysis. Several key residues with a higher importance for distinguishing Nb80-stabilized active conformations were extracted for further investigation (Figure 3D). For the Nb80-free ensembles, agonists governed different TM6 and TM7 orientations near the NPxxY motif, leading to distinct distances between Y326$^{7.53}$ and L275$^{6.37}$. In the same region, a hydrogen bond formed between S319$^{7.46}$ and N51$^{1.50}$, one of the most conserved residues in the class A GPCRs, only in the agonist-bound receptor. However, we notice that Nb80 binding stabilized similar conformation between Y326$^{7.53}$ and L275$^{6.37}$ and maintained the hydrogen contact of S319$^{7.46}$ with N51$^{1.50}$ regardless of ligand bound (Figure 3E). Moreover, our analysis indicated that agonists induced a local contraction between L284$^{6.46}$ and F321$^{7.48}$ and a long-range contraction between S203$^{5.43}$ and E338$^{8.56}$ compared to nonagonists in the absence of Nb80. These residues are located around TM5 bulge, PIF motif, and NPxxY motif, and play an important role in the receptor activation (Figure 3D). However, from the data-driven analysis, the binding of Nb80 could make the distribution for the above four residues overlap for nonagonist-bound receptor features (Figure 3F). Such comparison further supports the finding that Nb80 binding induces some structural rearrangements throughout the protein and stabilizes a fully active-like conformation of the *β*2AR independently of the chemical nature of the ligand bound in the extracellular site. This also suggests a higher free energy barrier for Nb binding for nonagonist ligands.

Ligand-Specific Local Conformational Changes in the Presence of Nb80. To better understand the different ligand-induced conformational changes in *β*2AR-Nb80 complexes, the receptor ensembles of the apo, BI167107-bound, and salmeterol-bound, which all occupied a distinct region of the conformational space in the dimensional reduction analysis (Figure 2E,F), were labeled as separate datasets for further classification. Salmeterol is a functionally selective *β*2AR partial agonist with a 5- to 20-fold bias toward the activation of Gs over arrestin.[27,28] In addition, its high selectivity and long-acting properties contribute to it being one of the most prescribed drugs for treating asthma and chronic obstructive

https://doi.org/10.1021/acs.jcim.1c00826
J. Chem. Inf. Model. 2021, 61, 6024−6037

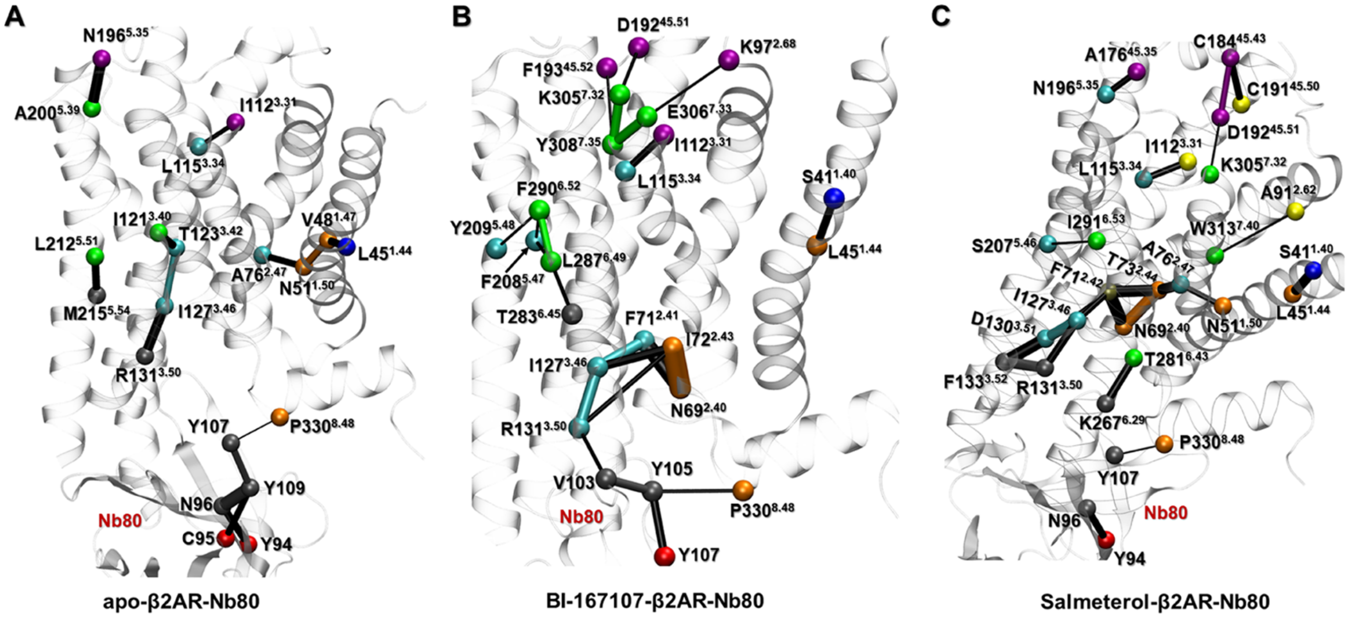


Figure 6. Critical nodes in the apo, BI167107- and salmeterol-bound β2AR-Nb80 structures. Each critical node is located at the interface of neighboring communities and corresponds to the edge with the highest score in terms of connectivity. Critical nodes are colored consistently with the communities of dynamic network models of Figure 5, and the connecting edges are represented by lines with their width weighted by betweenness.

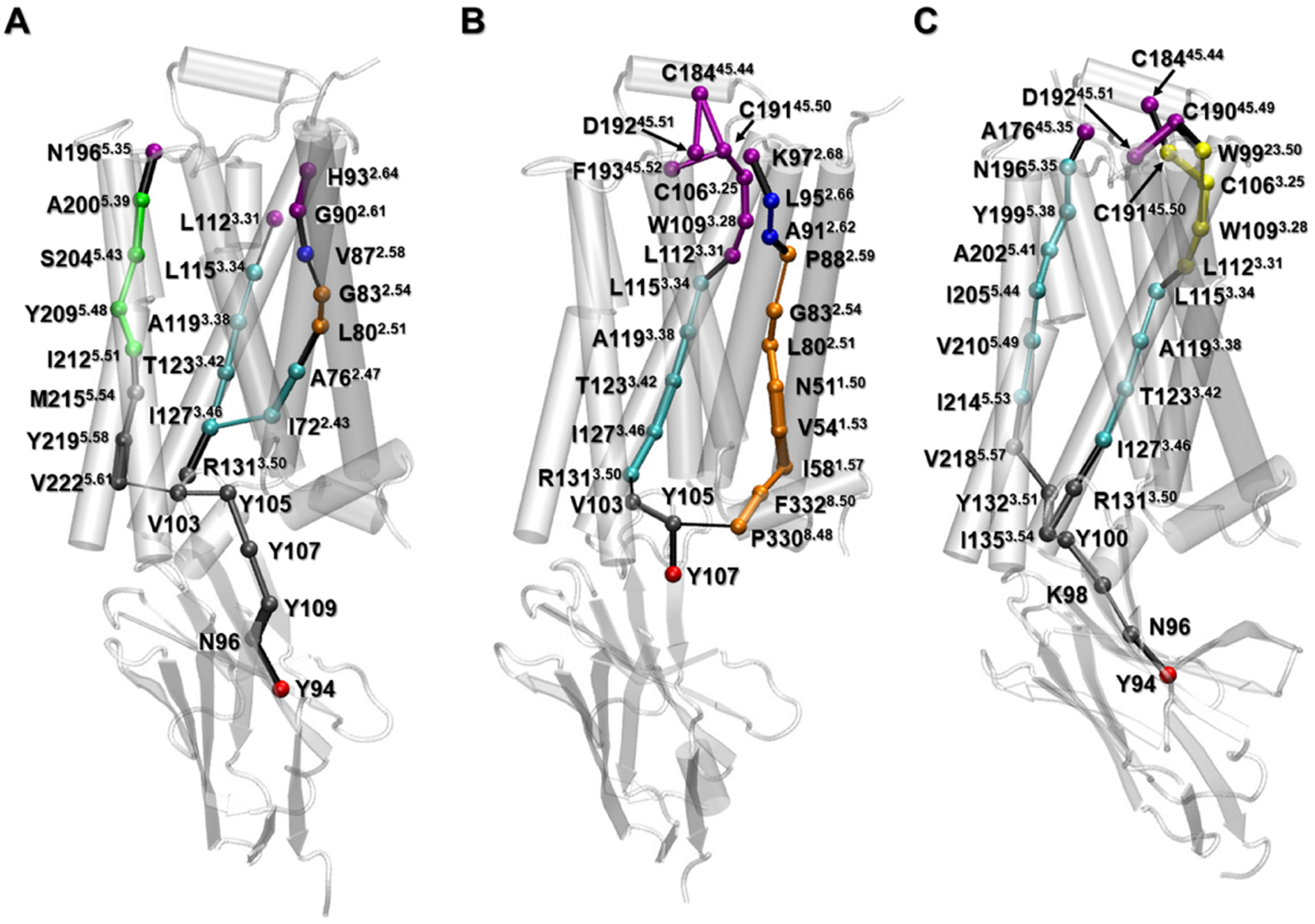
pulmonary disease (COPD).[61] Using a similar protocol as above, we identified features specific to the three chosen ensembles against the others in the presence of Nb80 (Figure 4A−C). Compared to the two ligand-bound ensembles, there are more important residues located in the TM domain in the unliganded ensemble (Figure 4A−C). This may originate from the large flexibility of the apo receptor (Figure S9), in agreement with spectroscopy experiments suggesting that the β2AR could not be stabilized in its fully active state in the absence of agonist binding.[20]

Notably, only a few residues in the BI167107- and salmeterol-β2AR-Nb80 complexes were captured as important, which indicates that there were only subtle differences between their conformational ensembles. Those corresponded to, for example, L144$^{34.56}$ and E268$^{6.30}$ in the BI167107-bound state, and P138$^{34.50}$, F193$^{45.52}$, and Y308$^{7.35}$ in the salmeterol-bound state. Among them, L144$^{34.56}$ and P138$^{34.50}$ are located in ICL2, which is associated with distinct ligand-dependent conformational changes to recognize G-proteins or β-arrestin.[59,62] Furthermore, mutational and biophysical analysis suggested that F193$^{45.52}$ and Y308$^{7.35}$ are closer to each other in the agonist-bound β2AR-Nb80 complex and form a lid-like structure over the orthosteric ligand-binding pocket, which slowed down the rate of ligand dissociation, and accordingly contributed to the enhancement of the ligand affinity.[55] As shown in Figure 4E, the distance between F193$^{45.52}$ and Y308$^{7.35}$ in unliganded simulation snapshots ranged from 3 Å to 7 Å, while in BI167107- and salmeterol-bound states it was stabilized around 3−4 Å, which provides a structural explanation for the agonist-induced enhancement of receptor activation observed in experiments.[63,64] Meanwhile, we noticed that the F193$^{45.52}$-Y308$^{7.35}$ distance in salmeterol-β2AR-Nb80 complex is slightly larger than that in BI167107-β2AR-Nb80 complex, which can be related to the lower affinity and partial activation effect of salmeterol. Moreover, mutagenesis studies[65,66] have reported that the hydrogen bond between F193$^{45.52}$ and the aryl-oxy-alkyl tail of salmeterol contributed to its high selectivity for β2AR over β1AR.

Furthermore, we also identified other residue pairs with high importance profiles, such as P138$^{34.50}$-I121$^{3.40}$ and L144$^{34.56}$-R131$^{3.50}$, indicative of BI167107 and salmeterol binding resulting in different conformations (Figures 4D,E and S10). Among them, I121$^{3.40}$ and R131$^{3.50}$ are part of the PIF motif and the "ionic lock", respectively, which are hallmarks of GPCR activation.[35] We observe shorter distances of the residues pairs P138$^{34.50}$-I121$^{3.40}$ and L144$^{34.56}$-R131$^{3.50}$ in the salmeterol complex than those in the BI167107 complex, indicating a loose interaction connecting the intracellular region and ligand-binding site in the BI167107-bound structure. We also did computational alanine scanning analysis on unliganded, BI167107- and salmeterol-bound structures through Rosetta alanine scan serve[67] to further demonstrate the importance of residues identified above in the receptor activation (Figure S11). Overall, there are indeed distinct structural features associated with the receptor activation presenting in Nb80-stabilized β2AR bound to ligands with varying efficacies.

In general, our approach has succeeded in identifying important features distinguishing Nb80-bound and -unbound states. In addition to the intracellular end of TM6, some highly conserved residues such as N51$^{1.50}$, S319$^{7.46}$, S203$^{5.43}$, and Y326$^{7.53}$ were identified to play crucial roles in the receptor activation. ICL2, involved in G-protein activation, was also highlighted to be important in different ligand-bound β2AR-Nb80 structures. Interestingly, F193$^{45.52}$ was captured as a key factor for the selectivity of salmeterol in β2AR activation.

Dynamic Allosteric Network in the β2AR. In recent years, computational dynamic network models have been widely applied to biomolecule systems to decipher residue−residue interactions and elucidate allosteric communication.[68,69] In this study, to gain insights into the allosteric communication pathways modulated by Nb80 and ligands, we



apo-β2AR-Nb80 BI-167107-β2AR-Nb80 Salmeterol-β2AR-Nb80

Figure 7. Optimal paths connecting the intracellular (C7) and extracellular binding sites (C1) in apo (A), BI167107 (B), and salmeterol (C)-bound *β*2AR-Nb80 structures. Residues are rendered as spheres and colored consistently with the communities they belong to in Figure 5, and the connecting edges are represented by lines with their width weighted by betweenness.

constructed a residue interaction network model using the MD simulation ensembles of the *β*2AR and analyzed it using community network analysis (see the Methods section). We specifically focused on six ensembles, including unliganded, BI167107- and salmeterol-bound *β*2AR in the presence and absence of Nb80 (Figures 5 and S12). As shown in Figure 5, there are distinct intercommunity flows in these different receptor states. Overall, a smaller number of communities are identified in the Nb80-bound structures compared to those in Nb80-free states, indicating that Nb80 binding induces tighter and stronger local communication networks and consequently bigger communities. For example, in the unliganded structures, Nb80 binding promoted the grouping of communities 5, 9, and 10 (C5, C9, and C10), which are mainly found at the intracellular end of TM3, TM5, TM6, and ICL2, which correspond to a large community (C5) in the apo-*β*2AR-Nb80 complex (Figure 5A,D). Similarly, communities C1 and C7 located at the extracellular domain of the BI167107-bound structure merged into a single cluster C1 upon binding of Nb80 (Figure 5B,E). However, we observed that the dynamic network of the salmeterol-bound receptor is different from the others in the presence of Nb80, especially in the extracellular region (Figure 5C). Community C1 in the unliganded and BI167107-bound states was split into C1 and C9 in the salmeterol-bound state. This might be attributed to the long aryl-oxy-alkyl tail of salmeterol, which led to the generation of an exosite consisting of residues around ECL2, ECL3, and the extracellular ends of TM6 and TM7. Interestingly, the exosite is associated with high receptor selectivity and ligand affinity.[35]

We also explored conformational changes around the ligand-binding pocket induced by Nb80 binding. Figure S13 displays residues within a 4.0 Å cutoff of ligands BI167107 and salmeterol. In the BI167107-*β*2AR ensemble, residues at the orthosteric site are involved in communities C1, C2, and C7, which were redistributed in C1, C2, and C4 after Nb80 binding (Figure S14A,B). Especially, F193$^{45.52}$ in ECL2 was merged into one group with W109$^{3.28}$, D113$^{3.32}$, and Y316$^{7.43}$ in TM3 and TM7, respectively, suggesting that stronger interactions formed between these residues, which presumably contributes to the slower dissociation of BI167107 from the orthosteric site. Compared to BI167107, more residues form the salmeterol-binding pocket due to its long aryl-oxy-alkyl tail (Figure S13). In contrast to BI167107, there was no big difference in the interaction network around the exosite, but stronger communications occurred in F289$^{6.51}$, Y308$^{7.35}$, and N312$^{7.39}$ when Nb80 was bound, resulting in the extracellular end of TM6 and TM7 being closer to the ligand (Figure S14C,D).

In addition, nodes (residues) critical for the communication across communities were identified for the three receptor-

Nb80 ensembles (Figure 6). Residues R131$^{3.50}$, I127$^{3.46}$, I112$^{3.31}$, and L115$^{3.34}$ located on TM3, which is an important signal transduction domain across class A GPCRs,[70] were identified as critical residues in all of the Nb80-bound structures. Furthermore, other critical residues like D192$^{45.51}$ and K305$^{7.32}$ in BI167107- and salmeterol-bound states have been reported to contribute to the formation of a closed conformation over the ligand-binding pocket, in part responsible for enhanced ligands binding affinity.[55] We also found some ligand-specific critical residues, such as K97$^{2.68}$, E306$^{7.33}$, F290$^{6.52}$, and Y209$^{5.48}$ in the BI167107-bound state, and C184$^{45.43}$, C191$^{45.50}$, W313$^{7.40}$, and S207$^{5.46}$ in the salmeterol-bound state, which also have an effect on the receptor activity and ligand affinity, supported by the mutagenesis data reported in the G-protein-coupled receptor data bank (GPCRdb, http://gpcrdb.org).[71,72]

We further analyzed the optimal pathways in the three Nb80-bound structures, to identify residues involved in information transfer from the Nb80-binding site to the extracellular domain of the receptor (Figure 7). For each network model, we selected critical nodes in communities C7 (Nb80) and C1 (extracellular binding site) as start- and end-points, respectively, for pathways calculation. Those were Y94, N196$^{5.35}$, L112$^{3.31}$, and H93$^{2.64}$ in the unliganded state; Y107, F193$^{45.53}$, C192$^{45.51}$, and K97$^{2.68}$ in the BI167107-bound state; and Y94, A176$^{45.35}$, C184$^{45.44}$, and C192$^{45.51}$ in the salmeterol-bound state (Figure 6). The C5 community contains residues from both the Nb80 and intracellular ends of TM3, TM5, and TM6, forming an interfacial community. More residues of Nb80 merged into C5 in the unliganded and salmeterol-bound networks than in the BI167107-bound network. The unliganded and salmeterol-β2AR-Nb80 states share the inner Nb residue Y94 as an important residue for communication between the Nb80-only community C7 with the mixed receptor-Nb80 community C5. In contrast, in the BI167107-bound ensemble, the surface residue Y107 fulfills this role. This suggests a comparatively loose interaction induced by BI167107 at the β2AR-Nb80 interface (Figures 7 and S15).

In addition, the major difference in the three models is that there are three optimal pathways connecting the extra- and intracellular binding sites in the unliganded structure, while there are only two in ligand-bound structures. This could be expected from the more prominent fluctuations in the ligand-free receptor (Figure S9). Indeed, we observed one pathway going primarily along TM3 in all three network models, which used R131$^{3.50}$ of the ionic lock as a bridge node connecting the Nb80-binding and ligand-binding sites. This is supported by previous studies emphasizing the significant role of TM3 in signal transduction between the intracellular and extracellular binding sites.[70,73] Notably, the BI167107-specific pathway sent signals mainly along H8, TM1, and TM2, whereas the salmeterol-bound receptor's pathway prominently involved TM5 (Figure 7B,C). Several studies pointed out that TM2 might be regarded as a pivot for activating conformational change of GPCRs, in which the Pro residue at 2.58, 2.59, or 2.60 may contribute to specialize GPCRs binding of different ligand types (P88$^{2.59}$ in the β2AR).[70,74] In addition to P88$^{2.59}$, N51$^{1.50}$ in BI167107-specific optimal pathway is associated with water-mediated interactions around the cytoplasmic halves of TM2, TM6, and TM7, playing a crucial role in GPCR activation.[75] In the case of the salmeterol-specific pathway, I205$^{5.44}$ and V210$^{5.49}$ are located near S207$^{5.46}$ (the TM5 bulge) and I211$^{5.50}$ (PIF motif), which are involved in highly conserved microswitches.[35] The hydrophobic interactions involving I205$^{5.44}$ and V210$^{5.49}$ may indirectly help stabilize the inward conformations of S207$^{5.46}$ and I211$^{5.50}$. Moreover, V218$^{5.57}$ and N196$^{5.35}$ contributed to signal transmission connecting the intracellular end of TM3 and ECL2 region (Figure 7C).

To summarize, network analysis revealed that Nb80 induced high levels of communication especially in the intracellular domains of TM3, TM5, TM6, and ICL2, and in the extracellular domains of TM2, TM3, TM5, TM7, and ECL2. With this approach, we also identified critical residues that had important effects on the receptor activity and ligand affinity. In addition, ligand-specific allosteric signaling pathways highlighted different conformational changes controlled by the ligands.

CONCLUSIONS

Many studies[54,55,76,77] have shown that nanobodies, functioning as G-protein mimetics, succeed in stabilizing different GPCR conformations and further affect the affinity of ligands by allosteric modulation. Nb80, the first reported nanobody, bound intracellularly to β2AR not only fully stabilizes the active agonist-bound receptor conformation but also highly improves the agonist affinity. In this work, on the basis of data-driven methods and dynamic network analysis for active-like β2AR ensembles bound to ligands with varying efficacies in the absence and presence of Nb80, we propose a molecular interpretation of the allosteric modulation mechanism due to Nb80 binding. Nb80 binding was found to stabilize the same conformational rearrangements for different systems, especially the larger intracellular outward movement of TM6 and the decrease in the distance of the Y-Y motif and the RMSD of NPxxY motif. Highly conserved residues N51$^{1.50}$, S319$^{7.46}$, S203$^{5.43}$, L284$^{6.46}$, and Y326$^{7.53}$ are identified to be important in the Nb80-stabilized active β2AR conformation. Network analysis further reveals Nb80-induced stronger interactions in the intracellular and extracellular domains of the receptor. In addition, apo, BI167107-, and salmeterol-bound states exhibit subtle differences in TM3, ECL2, and ICL2 induced by resides such as I121$^{3.40}$, R131$^{3.50}$, F193$^{45.52}$, P138$^{34.50}$, and L144$^{34.56}$, some of which are also identified as critical nodes in dynamical network models and proved to be important for the receptor activity by previous mutagenesis experiments.[35,70,71] Interestingly, we observed that the BI167107- and salmeterol-specific optimal pathways contribute to the signal transmission connecting Nb80 and ligand-binding sites mainly via TM1, TM2, and TM5, respectively.

Thus, enhanced sampling MD simulations combined with data-driven analysis methods were useful to probe the allosteric effect of Nb80 binding. Our results shed light on ligands-specific subtle structural differences and signal transmission pathways. This work provides structural insights underlying the enhanced β2AR activation activity and ligand affinity modulated by Nb80. These findings could be helpful for structure-based drug discovery targeting GPCRs, taking into account the effect of intracellular binding partners.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00826.

Distance between the center points of replicas over iterations; important residues for discriminating Nb80-bound and -unbound ensembles, derived from PCA on equilibrated active-like ensembles; overlay of representative snapshots of all Nb80-bound and -unbound states; TM6-TM3 distance, Y-Y motif distance and RMSD of NPxxY motif in all Nb80-bound and -unbound ensembles along the MD trajectory; dimensionality reduction analysis applied to the active-like simulation ensembles; important residues derived from the equilibrated active-like ensembles for discriminating Nb80- and ligand-dependent activation mechanisms by computing Kullback−Leibler divergence (KL); residue average fluctuations measured as root-mean-square fluctuation (RMSF) in the active-like simulation ensembles of apo, BI167107-, and salmeterol-bound *β*2AR-Nb80 structures; superimposition of representative snapshots of apo, BI167107-, and salmeterol-bound *β*2AR in the presence of Nb80; computational alanine scanning of *β*2AR residues in apo-, BI167107-, and salmeterol-*β*2AR-Nb80 systems; dynamic networks are identified in the apo, BI167107-, and salmeterol-bound *β*2AR with and without Nb80 bound through community network analysis; overall ribbon representations of the BI167107- and salmeterol- binding pocket of DRD2; BI167107- and salmeterol-binding pockets with and without Nb80 are highlighted in the dynamic networks through community network analysis; and local network communities involving Nb80 and the intracellular domain of *β*2AR (PDF)

AUTHOR INFORMATION

Corresponding Author

Lucie Delemotte − *Science for Life Laboratory, Department of Applied Physics, KTH Royal Institute of Technology, SE-17121 Solna, Sweden;* orcid.org/0000-0002-0828-3899; Email: lucied@kth.se

Authors

Yue Chen − *Science for Life Laboratory, Department of Applied Physics, KTH Royal Institute of Technology, SE-17121 Solna, Sweden*

Oliver Fleetwood − *Science for Life Laboratory, Department of Applied Physics, KTH Royal Institute of Technology, SE-17121 Solna, Sweden;* orcid.org/0000-0002-4277-2661

Sergio Pérez-Conesa − *Science for Life Laboratory, Department of Applied Physics, KTH Royal Institute of Technology, SE-17121 Solna, Sweden*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.1c00826

Notes

The authors declare no competing financial interest.

The data necessary to reproduce the findings presented in this paper can be found on OSF (DOI 10.17605/OSF.IO/6XPYV). The code used to run and analyze simulations has been deposited on GitHub (https://github.com/delemottelab/demystifying, and https://github.com/delemottelab/state-sampling). The software used in this study includes GROMACS 2018.6 (open source, see https://www.gromacs.org/Downloads/), CHARMM-GUI (open source, see https://charmm-gui.org/), Cytoscape (open source, see https://cytoscape.org/), and VMD (https://www.ks.uiuc.edu/Research/vmd/vmd-1.9.3/).

ACKNOWLEDGMENTS

L.D. thanks the support of Science for Life Laboratory, the Göran Gustafsson Foundation, the Knut and Alice Wallenberg Foundation, and the Swedish research council (grant no. VR-2018-04905). The simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at PDC Centre for High Performance Computing (PDC-HPC).

REFERENCES

(1) Katritch, V.; Cherezov, V.; Stevens, R. C. Structure-Function of the G Protein−Coupled Receptor Superfamily. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53*, 531−556.

(2) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discovery* **2017**, *16*, 829−842.

(3) Hauser, A. S.; Chavali, S.; Masuho, I.; Jahn, L. J.; Martemyanov, K. A.; Gloriam, D. E.; Babu, M. M. Pharmacogenomics of GPCR Drug Targets. *Cell* **2018**, *172*, 41−54.

(4) Du, Y.; Duc, N. M.; Rasmussen, S. G. F.; Hilger, D.; Kubiak, X.; Wang, L.; Bohon, J.; Kim, H. R.; Wegrecki, M.; Asuru, A.; Jeong, K. M.; Lee, J.; Chance, M. R.; Lodowski, D. T.; Kobilka, B. K.; Chung, K. Y. Assembly of a GPCR-G Protein Complex. *Cell* **2019**, *177*, 1232.

(5) Latorraca, N. R.; Wang, J. K.; Bauer, B.; Townshend, R. J. L.; Hollingsworth, S. A.; Olivieri, J. E.; Xu, H. E.; Sommer, M. E.; Dror, R. O. Molecular mechanism of GPCR-mediated arrestin activation. *Nature* **2018**, *557*, 452−456.

(6) Filipek, S. Molecular switches in GPCRs. *Curr. Opin. Struct. Biol.* **2019**, *55*, 114−120.

(7) Schönegge, A.-M.; Gallion, J.; Picard, L.-P.; Wilkins, A. D.; Le Gouill, C.; Audet, M.; Stallaert, W.; Lohse, M. J.; Kimmel, M.; Lichtarge, O.; Bouvier, M. Evolutionary action and structural basis of the allosteric switch controlling *β*2AR functional selectivity. *Nat. Commun.* **2017**, *8*, No. 2169.

(8) Ballesteros, J. A.; Weinstein, H. Integrated Methods for the Construction of Three-dimensional Models and Computational Probing of Structure-Function Relations in G Protein-coupled Receptors. In *Methods in Neurosciences*; Sealfon, S. C., Ed.; Academic Press, 1995; Vol. *25*, pp 366−428.

(9) Inoue, A.; Raimondi, F.; Kadji, F. M. N.; Singh, G.; Kishi, T.; Uwamizu, A.; Ono, Y.; Shinjo, Y.; Ishida, S.; Arang, N.; Kawakami, K.; Gutkind, J. S.; Aoki, J.; Russell, R. B. Illuminating G-Protein-Coupling Selectivity of GPCRs. *Cell* **2019**, *177*, 1933.

(10) Glukhova, A.; Draper-Joyce, C. J.; Sunahara, R. K.; Christopoulos, A.; Wootten, D.; Sexton, P. M. Rules of Engagement: GPCRs and G Proteins. *ACS Pharmacol. Transl. Sci.* **2018**, *1*, 73−83.

(11) McCorvy, J. D.; Butler, K. V.; Kelly, B.; Rechsteiner, K.; Karpiak, J.; Betz, R. M.; Kormos, B. L.; Shoichet, B. K.; Dror, R. O.; Jin, J.; Roth, B. L. Structure-inspired design of *β*-arrestin-biased ligands for aminergic GPCRs. *Nat. Chem. Biol.* **2018**, *14*, 126−134.

(12) Smith, J. S.; Lefkowitz, R. J.; Rajagopal, S. Biased signalling: from simple switches to allosteric microprocessors. *Nat. Rev. Drug Discovery* **2018**, *17*, 243−260.

(13) Jaeger, K.; Bruenle, S.; Weinert, T.; Guba, W.; Muehle, J.; Miyazaki, T.; Weber, M.; Furrer, A.; Haenggi, N.; Tetaz, T.; Huang, C.-Y.; Mattle, D.; Vonach, J.-M.; Gast, A.; Kuglstatter, A.; Rudolph, M. G.; Nogly, P.; Benz, J.; Dawson, R. J. P.; Standfuss, J. Structural Basis for Allosteric Ligand Recognition in the Human CC Chemokine Receptor 7. *Cell* **2019**, *178*, 1222.

(14) Hollingsworth, S. A.; Kelly, B.; Valant, C.; Michaelis, J. A.; Mastromihalis, O.; Thompson, G.; Venkatakrishnan, A. J.; Hertig, S.; Scammells, P. J.; Sexton, P. M.; Felder, C. C.; Christopoulos, A.; Dror, R. O. Cryptic pocket formation underlies allosteric modulator selectivity at muscarinic GPCRs. *Nat. Commun.* **2019**, *10*, No. 3289.

(15) Chen, S.; Lu, M.; Liu, D.; Yang, L.; Yi, C.; Ma, L.; Zhang, H.; Liu, Q.; Frimurer, T. M.; Wang, M.-W.; Schwartz, T. W.; Stevens, R. C.; Wu, B.; Wüthrich, K.; Zhao, Q. Human substance P receptor binding mode of the antagonist drug aprepitant by NMR and crystallography. *Nat. Commun.* **2019**, *10*, No. 638.

(16) Cheng, Y. Single-particle cryo-EM—How did it get here and where will it go. *Science* **2018**, *361*, 876−880.

(17) Dror, R. O.; Arlow, D. H.; Maragakis, P.; Mildorf, T. J.; Pan, A. C.; Xu, H.; Borhani, D. W.; Shaw, D. E. Activation mechanism of the β2-adrenergic receptor. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18684−18689.

(18) Dror, R. O.; Mildorf, T. J.; Hilger, D.; Manglik, A.; Borhani, D. W.; Arlow, D. H.; Philippsen, A.; Villanueva, N.; Yang, Z.; Lerch, M. T.; Hubbell, W. L.; Kobilka, B. K.; Sunahara, R. K.; Shaw, D. E. Structural basis for nucleotide exchange in heterotrimeric G proteins. *Science* **2015**, *348*, 1361−1365.

(19) Nygaard, R.; Zou, Y.; Ron, O.; Thomas, J.; Daniel, H.; Manglik, A.; Albert, C.; Corey, W.; Fung, J. J.; Bokoch, M. P.; Thian, F. S.; Kobilka, T. S.; Shaw, D. E.; Mueller, L.; Prosser, R. S.; Kobilka, B. K. The Dynamic Process of β2-Adrenergic Receptor Activation. *Cell* **2013**, *152*, 532−542.

(20) Manglik, A.; Tae, H.; Masureel, M.; Altenbach, C.; Yang, Z.; Hilger, D.; Michael, T.; Tong, S.; Foon, S.; Wayne, L.; Prosser, R. S.; Brian, K. Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **2015**, *161*, 1101−1111.

(21) Strachan, R. T.; Sun, J.-p.; Rominger, D. H.; Violin, J. D.; Ahn, S.; Rojas Bie Thomsen, A.; Zhu, X.; Kleist, A.; Costa, T.; Lefkowitz, R. J. Divergent Transducer-specific Molecular Efficacies Generate Biased Agonism at a G Protein-coupled Receptor (GPCR)*. *J. Biol. Chem.* **2014**, *289*, 14211−14224.

(22) De Lean, A.; Stadel, J. M.; Lefkowitz, R. J. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *J. Biol. Chem.* **1980**, *255*, 7108−7117.

(23) Lee, S.; Nivedha, A. K.; Tate, C. G.; Vaidehi, N. Dynamic Role of the G Protein in Stabilizing the Active State of the Adenosine A2A Receptor. *Structure* **2019**, *27*, 703.

(24) Kim, K.; Paulekas, S.; Sadler, F.; Gupte, T. M.; Ritt, M.; Dysthe, M.; Vaidehi, N.; Sivaramakrishnan, S. β2-adrenoceptor ligand efficacy is tuned by a two-stage interaction with the Gαs C terminus. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, No. e2017201118.

(25) Fleetwood, O.; Kasimova, M. A.; Westerlund, A. M.; Delemotte, L. Molecular Insights from Conformational Ensembles via Machine Learning. *Biophys. J.* **2020**, *118*, 765−780.

(26) Rasmussen, S. G. F.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature* **2011**, *469*, 175−180.

(27) Gimenez, L. E.; Baameur, F.; Vayttaden, S. J.; Clark, R. B. Salmeterol efficacy and bias in the activation and kinase-mediated desensitization of β2-adrenergic receptors. *Mol. Pharmacol.* **2015**, *87*, 954.

(28) Fleetwood, O.; Matricon, P.; Carlsson, J.; Delemotte, L. Energy Landscapes Reveal Agonist Control of G Protein-Coupled Receptor Activation via Microswitches. *Biochemistry* **2020**, *59*, 880−891.

(29) Fleetwood, O.; Carlsson, J.; Delemotte, L. Identification of ligand-specific G protein-coupled receptor states and prediction of downstream efficacy via data-driven modeling. *eLife* **2021**, *10*, No. e60715.

(30) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405−413.

(31) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.

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

(33) Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved Binding Mode of Human β2 Adrenergic Receptor Inverse Agonists and Antagonist Revealed by X-ray Crystallography. *J. Am. Chem. Soc.* **2010**, *132*, 11443−11445.

(34) Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor. *Structure* **2008**, *16*, 897−905.

(35) Masureel, M.; Zou, Y.; Picard, L.-P.; van der Westhuizen, E.; Mahoney, J. P.; Rodrigues, J. P. G. L. M.; Mildorf, T. J.; Dror, R. O.; Shaw, D. E.; Bouvier, M.; Pardon, E.; Steyaert, J.; Sunahara, R. K.; Weis, W. I.; Zhang, C.; Kobilka, B. K. Structural insights into binding specificity, efficacy and bias of a β2AR partial agonist. *Nat. Chem. Biol.* **2018**, *14*, 1059−1066.

(36) Ring, A. M.; Manglik, A.; Kruse, A. C.; Enos, M. D.; Weis, W. I.; Garcia, K. C.; Kobilka, B. K. Adrenaline-activated structure of β2-adrenoceptor stabilized by an engineered nanobody. *Nature* **2013**, *502*, 575−579.

(37) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830−7843.

(38) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(39) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19−25.

(40) Tipping, M. E.; Bishop, C. M. Probabilistic Principal Component Analysis. *J. R. Stat. Soc. Ser. B* **1999**, *61*, 611−622.

(41) Borg, I.; Groenen, P. J. *Modern Multidimensional Scaling: Theory and Applications*; Springer Science & Business Media, 2005.

(42) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

(43) Bishop, C. M. *Pattern Recognition and Machine Learning*; Springer, 2006.

(44) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980 2014.

(45) Montavon, G.; Samek, W.; Müller, K.-R. Methods for interpreting and understanding deep neural networks. *Digital Signal Process.* **2018**, *73*, 1−15.

(46) Kullback, S.; Leibler, R. A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79−86.

(47) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6620−6625.

(48) Eargle, J.; Luthey-Schulten, Z. NetworkView: 3D display and analysis of protein·RNA interaction networks. *Bioinformatics* **2012**, *28*, 3000−3001.

(49) Girvan, M.; Newman, M. E. J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821−7826.

(50) Floyd, R. W. Algorithm 97: shortest path. *Commun. ACM* **1962**, *5*, 345.

(51) Grahl, A.; Abiko, L. A.; Isogai, S.; Sharpe, T.; Grzesiek, S. A high-resolution description of β1-adrenergic receptor functional dynamics and allosteric coupling from backbone NMR. *Nat. Commun.* **2020**, *11*, No. 2216.

(52) Ye, L.; Van Eps, N.; Zimmer, N.; Ernst, O. P.; Prosser, R. Activation of the A2A adenosine G-protein-coupled receptor by conformational selection. *Nature* **2016**, *533*, 265−268.

(53) Wingler, L. M.; Skiba, M. A.; McMahon, C.; Staus, D. P.; Kleinhenz, A. L. W.; Suomivuori, C.-M.; Latorraca, N. R.; Dror, R. O.; Lefkowitz, R. J.; Kruse, A. C. Angiotensin and biased analogs induce structurally distinct active conformations within a GPCR. *Science* **2020**, *367*, 888−892.

(54) Staus, D. P.; Strachan, R. T.; Manglik, A.; Pani, B.; Kahsai, A. W.; Kim, T. H.; Wingler, L. M.; Ahn, S.; Chatterjee, A.; Masoudi, A.; Kruse, A. C.; Pardon, E.; Steyaert, J.; Weis, W. I.; Prosser, R. S.; Kobilka, B. K.; Costa, T.; Lefkowitz, R. J. Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **2016**, *535*, 448−452.

(55) DeVree, B. T.; Mahoney, J. P.; Vélez-Ruiz, G. A.; Rasmussen, S. G. F.; Kuszak, A. J.; Edwald, E.; Fung, J.-J.; Manglik, A.; Masureel, M.; Du, Y.; Matt, R. A.; Pardon, E.; Steyaert, J.; Kobilka, B. K.; Sunahara, R. K. Allosteric coupling from G protein to the agonist-binding pocket in GPCRs. *Nature* **2016**, *535*, 182−186.

(56) Rosenbaum, D. M.; Zhang, C.; Lyons, J. A.; Holl, R.; Aragao, D.; Arlow, D. H.; Rasmussen, S. G. F.; Choi, H.-J.; Devree, B. T.; Sunahara, R. K.; Chae, P. S.; Gellman, S. H.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Caffrey, M.; Gmeiner, P.; Kobilka, B. K. Structure and function of an irreversible agonist-$\beta(2)$ adrenoceptor complex. *Nature* **2011**, *469*, 236−240.

(57) Bang, I.; Choi, H.-J. Structural features of β2 adrenergic receptor: crystal structures and beyond. *Mol. Cells* **2015**, *38*, 105−111.

(58) Liu, X.; Masoudi, A.; Kahsai, A. W.; Huang, L.-Y.; Pani, B.; Staus, D. P.; Shim, P. J.; Hirata, K.; Simhal, R. K.; Schwalb, A. M.; Rambarat, P. K.; Ahn, S.; Lefkowitz, R. J.; Kobilka, B. Mechanism of β2-AR regulation by an intracellular positive allosteric modulator. *Science* **2019**, *364*, 1283−1287.

(59) Shan, J.; Weinstein, H.; Mehler, E. L. Probing the Structural Determinants for the Function of Intracellular Loop 2 in Structurally Cognate G-Protein-Coupled Receptors. *Biochemistry* **2010**, *49*, 10691−10701.

(60) Moro, O.; Lameh, J.; Högger, P.; Sadée, W. Hydrophobic amino acid in the i2 loop plays a key role in receptor-G protein coupling. *J. Biol. Chem.* **1993**, *268*, 22273−22276.

(61) Cazzola, M.; Donner, C. F. Long-Acting β2 Agonists in the Management of Stable Chronic Obstructive Pulmonary Disease. *Drugs* **2000**, *60*, 307−320.

(62) Moreira, I. S. Structural features of the G-protein/GPCR interactions. *Biochim. Biophys. Acta, Gen. Subj.* **2014**, *1840*, 16−33.

(63) Yao, X. J.; Vélez Ruiz, G.; Whorton, M. R.; Rasmussen, S. G. F.; DeVree, B. T.; Deupi, X.; Sunahara, R. K.; Kobilka, B. The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9501−9506.

(64) Kim, H.; Lee, H. N.; Choi, J.; Seong, J. Spatiotemporal Characterization of GPCR Activity and Function during Endosomal Trafficking Pathway. *Anal. Chem.* **2021**, *93*, 2010−2017.

(65) Thal, D. M.; Glukhova, A.; Sexton, P. M.; Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature* **2018**, *559*, 45−53.

(66) Carter, A. A.; Hill, S. J. Characterization of Isoprenaline- and Salmeterol-Stimulated Interactions between β2-Adrenoceptors and β-Arrestin 2 Using β-Galactosidase Complementation in C2C12 Cells. *J. Pharmacol. Exp. Ther.* **2005**, *315*, 839−848.

(67) Kortemme, T.; Kim, D. E.; Baker, D. Computational Alanine Scanning of Protein-Protein Interfaces. *Sci. STKE* **2004**, *2004*, pl2.

(68) Bhattacharya, S.; Vaidehi, N. Differences in Allosteric Communication Pipelines in the Inactive and Active States of a GPCR. *Biophys. J.* **2014**, *107*, 422−434.

(69) Vaidehi, N.; Bhattacharya, S. Allosteric communication pipelines in G-protein-coupled receptors. *Curr. Opin. Pharmacol.* **2016**, *30*, 76−83.

(70) Hulme, E. C. GPCR activation: a mutagenic spotlight on crystal structures. *Trends Pharmacol. Sci.* **2013**, *34*, 67−84.

(71) Munk, C.; Harpsøe, K.; Hauser, A. S.; Isberg, V.; Gloriam, D. E. Integrating structural and mutagenesis data to elucidate GPCR ligand binding. *Curr. Opin. Pharmacol.* **2016**, *30*, 51−58.

(72) Kooistra, A. J.; Mordalski, S.; Pándy-Szekeres, G.; Esguerra, M.; Mamyrbekov, A.; Munk, C.; Keserű, G. M.; David, E. GPCRdb in 2021: integrating GPCR sequence, structure and function. *Nucleic Acids Res.* **2021**, *49*, D335−D343.

(73) Zhou, Q.; Yang, D.; Wu, M.; Guo, Y.; Guo, W.; Zhong, L.; Cai, X.; Dai, A.; Jang, W.; Shakhnovich, E. I.; Liu, Z.-J.; Stevens, R. C.; Lambert, N. A.; Babu, M. M.; Wang, M.-W.; Zhao, S. Common activation mechanism of class A GPCRs. *eLife* **2019**, *8*, No. e50279.

(74) Chabbert, M.; Castel, H.; Pele, J.; Deville, J.; Legendre, R.; Rodien, P. Evolution of class A G-protein-coupled receptors: implications for molecular modeling. *Curr. Med. Chem.* **2012**, *19*, 1110−1118.

(75) Venkatakrishnan, A. J.; Ma, A. K.; Fonseca, R.; Latorraca, N. R.; Kelly, B.; Betz, R. M.; Asawa, C.; Kobilka, B. K.; Dror, R. O. Diverse GPCRs exhibit conserved water networks for stabilization and activation. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3288−3293.

(76) Heukers, R.; De Groof, T. W. M.; Smit, M. J. Nanobodies detecting and modulating GPCRs outside in and inside out. *Curr. Opin. Cell Biol.* **2019**, *57*, 115−122.

(77) De Groof, T. W. M.; Bobkov, V.; Heukers, R.; Smit, M. J. Nanobodies: New avenues for imaging, stabilizing and modulating GPCRs. *Mol. Cell. Endocrinol.* **2019**, *484*, 15−24.