Received: 16 February 2022 | Revised: 9 August 2022 | Accepted: 16 September 2022

DOI: 10.1002/pro.4456

FULL-LENGTH PAPER

THE PROTEIN SOCIETY WILEY

Mechanistic basis of GPCR activation explored by ensemble refinement of crystallographic structures

Jesper J. Madsen[1,2] | **Libin Ye**[3,4] | **Thomas M. Frimurer**[5] | **Ole H. Olsen**[5]

[1]Global and Planetary Health, College of Public Health, University of South Florida, Tampa, Florida, USA

[2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, Florida, USA

[3]Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, Tampa, Florida, USA

[4]H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida, USA

[5]Section for Metabolic Receptology, Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark

Correspondence
Ole H. Olsen, Section for Metabolic Receptology, Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Blegdamsvej 3b, DK-2200 Copenhagen, Denmark.
Email: oho@sund.ku.dk

Review Editor: Nir Ben-Tal

Abstract

G protein-coupled receptors (GPCRs) are important drug targets characterized by a canonical seven transmembrane (TM) helix architecture. Recent advances in X-ray crystallography and cryo-EM have resulted in a wealth of GPCR structures that have been used in drug design and formed the basis for mechanistic activation hypotheses. Here, ensemble refinement (ER) of crystallographic structures is applied to explore the impact of binding of agonists and antagonist/inverse agonists to selected structures of cannabinoid receptor 1 (CB1R), β_2 adrenergic receptor (β_2AR), and A_{2A} adenosine receptor (A_{2A}AR). To assess the conformational flexibility and its role in GPCR activation, hydrogen bond (H-bond) networks are analyzed by calculating and comparing H-bond propensities. Mapping pairwise propensity differences between agonist- and inverse agonist/antagonist-bound structures for CB1R and β_2AR shows that agonist binding destabilizes H-bonds in the intracellular parts of TM 5–7, forming the G protein binding cavity, while H-bonds of the extracellular segment of TMs surrounding the orthosteric site are conversely stabilized. Certain class A GPCRs, for example, A_{2A}AR, bind an allosteric sodium ion that negatively modulates agonist binding. The impact of sodium-excluding mutants ($D52^{2.50}N$, $S91^{3.39}A$) of A_{2A}AR on agonist binding is examined by applying ER analysis to structures of wildtype and the two mutants in complex with a full agonist. While $S91^{3.39}A$ exhibits normal activity, $D52^{2.50}N$ quenches the downstream signaling. The mainchain H-bond pattern of the latter is stabilized in the intracellular part of TM 7 containing the NPxxY motif, indicating that an induced rigidity of the mutation prevents conformational selection of G proteins resulting in receptor inactivation.

KEYWORDS

allosteric regulation, ensemble refinement, G protein-coupled receptor (GPCR), hydrogen bond, molecular dynamics, protein conformation

1 | INTRODUCTION

G protein-coupled receptors (GPCRs) are the largest class of cell surface receptors and responsible for the majority of signal transduction across cell membranes in response to extracellular ligands.[1] Many key physiological responses are regulated by GPCRs and as such they represent a prominent group of therapeutic targets for a wide range of diseases.[2–6] The binding of agonists induce conformational changes at the intracellular cytoplasmic side of the receptor, which promote downstream signaling cascades mediated by binding of G proteins and other intracellular effectors such as G protein-coupled receptor kinases (GRKs) and β-arrestins.[7,8] X-ray crystallographic structures of genetically modified GPCRs co-crystallized with agonists, inverse agonist, antagonists, and co-factors have paved the way for understanding the receptor activation process.[7,9–14] The GPCR transmembrane domain is composed of a canonical transmembrane (TM) architecture with seven alpha helices in which particular structural changes can be observed accompanying activation in the intracellular domain such as elongation and rotation of TM 5, outward movement of TM 6, and inward movement of TM 7, resulting in the formation of a large intracellular crevice.[15,16] It is hypothesized that these conformational changes are mediated by a conserved network of non-covalent contacts whose allosteric rearrangements define the activation pathways. Of additional interest are three groups of highly conserved amino acids (called microswitches) that appear to have special significance in the activation pathways.[17–22] The microswitch motifs reside in TM 3 (DRY), TM 6 (CWxP), and TM 7 (NPxxY), respectively. Lastly, growing emphasis is put on delineating the ensemble nature of receptor proteins such as GPCRs in order to understand how their conformational ensembles are shifted between multiple states responsible for downstream signaling by various ligands.[23,24] Numerous biophysical explorations have supported these hypotheses.[6,25–34]

A central glimpse into the molecular basis of GPCR activation is provided by the wealth of structures that is becoming available. Unfortunately, static X-ray structures only reveal part of the biochemical puzzle since a comprehensive understanding of GPCR function requires the grasp of conformational flexibility regulated by ligands, allosteric modulators, co-factors, etc. Single structure models derived from X-ray data, "snapshots," do not adequately account for the inherent, functionally important dynamics of protein molecules.[35] The B-factor is a measure of the thermal motion of individual atoms—low B-factors indicate well-ordered parts while large B-factors indicate less-ordered or flexible parts—but they do not directly allow for analysis of nonbonded contacts such as hydrogen bond (H-bond) networks, which are highly relevant to receptor activation and allostery. Thus, accurate measurements of dynamic properties of GPCRs are essential in complementing structural information. Experimental techniques such as nuclear magnetic resonance (NMR) spectroscopy can be used to determine ensembles of conformational states,[27,28,36–42] and hydrogen-deuterium exchange-mass spectroscopy (HDX-MS) has been used to determine changes in the relative solvent accessibility of amide protons to reveal specific alterations in the energy landscapes of the receptors conformational states.[29,43–45]

An informative way to model the flexibility present in a crystallized protein is to represent the structure as a set of overlapping, noninteracting conformers each accounting for a fraction of the total electron density. This can be accomplished by applying ensemble refinement (ER)[46–48] to model both structure and flexibility based on X-ray diffraction data sets.[49] One inherent drawback is that the dynamics can be restricted by the crystal packing, which in turn can complicate comparisons across different crystallization conditions (e.g., crystal space groups). Another drawback is that sometimes-unnatural crystallization conditions (e.g., low temperature), receptor modifications, and thermo-stabilization strategies (e.g., co-crystallization of proteins or antibodies) introduced to facilitate crystallization. However, it is worth stressing that previous studies strongly support the reliability of using ER for evaluation of protein flexibility, potential improvement of Rfree/Rwork measures of model refinement quality,[47] and that good agreement has been established between both ER and NMR results as well as ER and biochemical results for highly allosterically regulated proteins.[35,50] We remark that information about (de)stabilization of the receptor structure with direct influence on the GPCR activation mechanism is retained to a larger extent than what is typically appreciated. Furthermore, the (de)stabilization patterns reveal quantifiable footprints of receptor activation, allostery, and biased signaling through their manifestation in the rigidity (plasticity) character of the receptor structure and dynamics.

The aim of the present study is to identify quantifiable footprints of altered dynamics or allostery in GPCR X-ray structures that have been crystallized in active (agonist-bound) versus inactive (antagonist/inverse agonist-bound) conformations by means of ER analysis. Such an analysis typically produces an ensemble of 30–100 structures for each corresponding static crystal structure that will be scrutinized in terms of interaction patterns that can facilitate activation or allostery: For instance, it is known that main chain hydrogen-bonded structural elements (like α-helices, β-sheets, or β-turns) can be attributed to 75% of the protein conformation,

with the remaining residues participating in both additional intramolecular hydrogen bonding and hydrogen bonding to water.[51] In addition, embedded water and some of the networks formed by water molecules in turn appear to be evolutionarily conserved among diverse GPCRs, which possibly contribute to mediating both state-dependent and state-independent interactions as well as swelling upon activation.[52–62] Due to their inherent importance in not only stabilizing the receptor structure, but also governing the ensemble dynamics upon ligand-binding, changes in H-bond strength and flexibility upon ligand binding are likely to reveal structural fingerprints of changed dynamics or allostery.

The paper is organized as follows. First, ER analyses of the differences in dynamics induced by binding of agonists and tight-binding antagonists to the cannabinoid receptor 1 GPCR (CB1R) and β_2 adrenergic GPCR (β_2AR) show that agonist binding destabilizes the H-bonds of the intracellular parts of TM 5, 6, and 7, allowing conformational selection of G proteins. The H-bonds of the extracellular segment of the TMs surrounding the orthosteric site are conversely stabilized. Further, the H-bond pattern is found altered in the vicinity of the microswitch motifs DRY, CWxP, and NPxxY when an agonist-bound CB1R ensemble is compared to that of antagonist bound. Second, certain class A GPCRs, for example, A_{2A}AR, bind an allosteric sodium ion which negatively modulates agonist binding. The impact of sodium-excluding mutants ($D52^{2.50}N$, $S91^{3.39}A$) of A_{2A}AR on agonist binding is examined by applying ER analysis to structures of wildtype and two mutants in complex with the same full agonist.[12] While $S91^{3.39}A$ exhibits normal activity, $D52^{2.50}N$ quenches the downstream signaling. The main chain H-bond pattern of the latter is stabilized in the intracellular part of TM 7, which contains the NPxxY motif, indicating that an induced rigidity of the mutation prevents conformational selection of G proteins resulting in receptor inactivation. The stabilized H-bond patterns of the mutants are consistent with their increased temperature stability.

2 | RESULTS AND DISCUSSION

The structural information obtained for GPCRs is largely based on X-ray structures of receptor molecules that are slightly modified (e.g., thermostabilizing mutants, fusion to helper molecules) which in turn may influence their static X-ray conformation. However, recall that these modified GPCRs display similar ligand binding properties as those of their parent receptor,[7] strongly indicating that conformations observed in X-ray structures are similar to their native reference and hence functionally relevant. As always when discussing the structures resolved in a crystal environment, close contacts caused by packing represents an inherent problem. However, the TM regions of the GPCRs that we are investigating are distant from neighboring molecules in the crystal as we shall illustrate in upcoming sections.

The ER analysis attempts to model structures as well as flexibility based on X-ray diffraction data by representing the structure as a set of overlapping, noninteracting conformers.[46–48] In the present work, ER analysis has been applied to three class A GPCRs: CB1R,[1,63,64] β_2AR,[65–68] and A_{2A}AR[12,28] to elucidate differences in flexibility and allostery induced by agonist- and inverse agonist-binding. The observations from ER analyses are supported by principal component analysis (PCA) on multiple independent molecular dynamic (MD) trajectories.

First, we explore structures of CB1R and β_2AR in complex with agonists and antagonist/inverse agonist to identify differences in H-bond patterns between active and inactive receptor conformational ensembles. Structures of the same construct of CB1R bound to agonist[63] and a tight-binding antagonist[64] are undertaken. The CB1R structures were resolved by the same laboratory and have the same resolution (2.8 Å), albeit crystallized in different space groups (Table S1). The CB1R construct was genetically modified and fused with flavodoxin as fusion protein-partner.

Next, we examine β_2AR structures in complex with an agonist[69] and a partial inverse agonist.[68] These structures represent an example of some complexity. Both modified receptor constructs are fused with T4-lysozyme, however the agonist-bound complex contains an additional stabilizing nanobody. Further, the space groups are different.

Finally, we study the impact of sodium ion-excluding mutants in A_{2A}AR by analyzing structures of both the wildtype and the two sodium ion-excluding mutants of A_{2A}AR in complex with the same agonist.[12] An identical A_{2A}AR construct (apart from the single point mutation in each mutant structure) is used and crystallization occurs in the same space group.

2.1 | H-bond network change is key in the activation mechanism of CB1R

In Figure 1a-d (left half), data from the static X-ray structure, ER analysis, crystallographic B-factors, and crystal contacts for the active conformation of CB1R with a bound agonist, AM11542 (PDB ID: 5xra) are reviewed. In the right half of Figure 1e-h, similar data for the inactive conformation of CB1R with a tight-binding antagonist

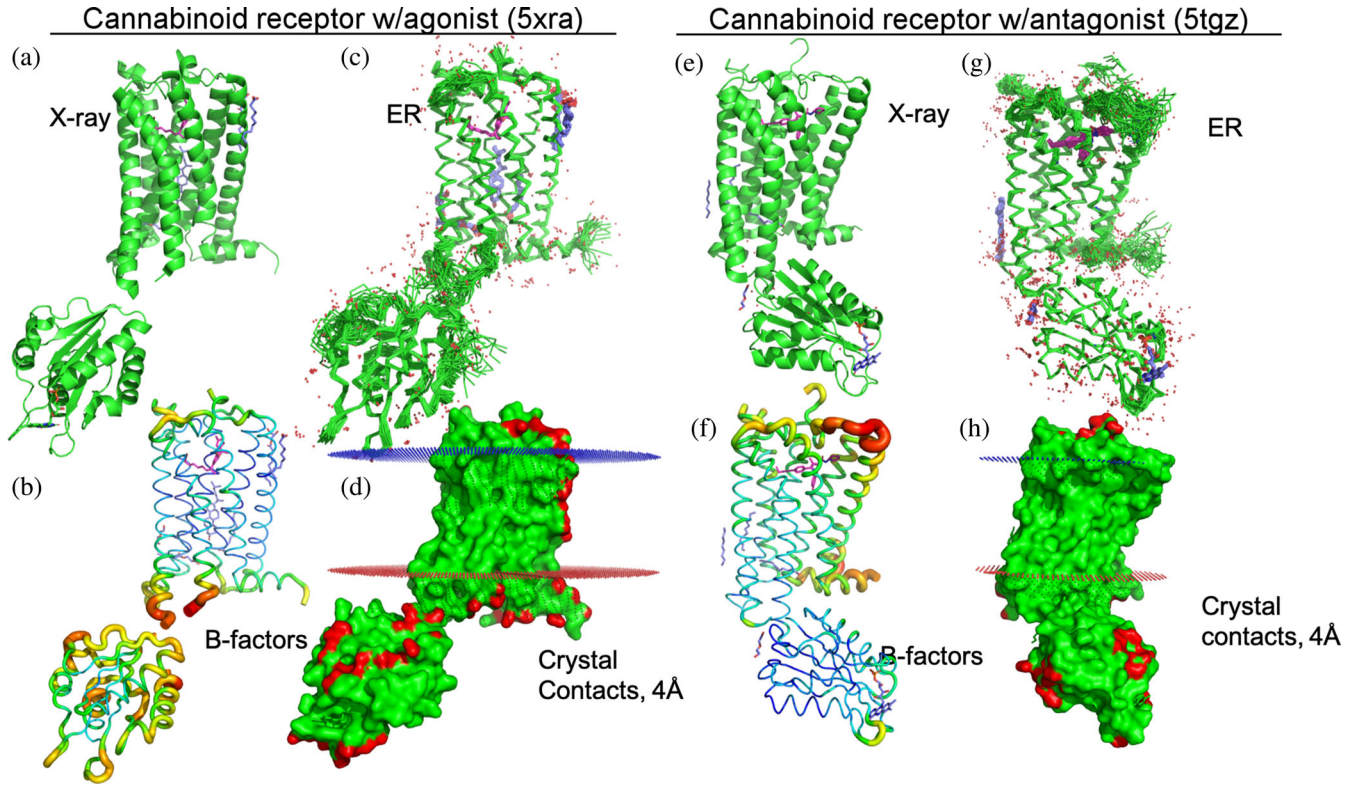

FIGURE 1 Overview of structures from ER analysis, static X-ray structure, crystallographic B-factors, and crystal contacts for CB1R with a bound agonist, AM11542 (PDB ID: 5xra) (a–d) and for CB1R with a bound antagonist, AM6538 (PDB ID: 5tgz) (e–h). (a and e) The static X-ray structures in green (cartoon), agonist/antagonist in purple, and lipids in blue. (b and f) Crystallographic B-factors shown as putty models, agonist/antagonist in purple, and lipids in blue. (c and g) Structures from ER analysis (green, in ribbon presentation), agonist/antagonist in purple, and lipids in blue. (d and h) Surface presentations of close crystal contacts, surface of atoms closer than 4 Å to a neighbor in the crystal lattices in red. The position of the membrane bilayer has been indicated by a blue (red) grid depicting extracellular (intracellular) membrane surface calculated using the PPM server.[70] The intercellularly bound domain is flavodoxin. CB1R, cannabinoid receptor 1; ER, ensemble refinement

AM6538 (PDB ID: 5tgz) are shown. The position of the membrane bilayer has been indicated by a blue (red) grid depicting extracellular (intracellular) membrane surface estimated using the PPM server.[70]

The presence of different ligands induces formation of the dissimilar crystal space groups, crystal contacts, and relative orientations of the receptor fused with flavodoxin (Figure 1d,h). The changed crystal contacts may potentially influence the H-bond patterns. However, the main chain H-bonds, which are embedded in the receptor structure and are of primary importance to our analyses, are in little to no contact with neighboring molecules particularly in the TM regions. Thus, the potential of crystal-contact artifacts are minimal and can be considered largely inconsequential to the main conclusions of this work. The precise details of this argument are provided for the individual GPCR systems below in their corresponding sections.

The calculated main chain H-bond propensities, P, are shown for the CB1R ER structures bound with agonist (PDB ID: 5xra, -P, red dots) and antagonist (PDB ID: 5tgz, P, black dots) versus residue number (Figure 2). The propensity of a H-bond was determined as the number of ER structures in which it was present relative to the total number of structures in the ER ensemble to yield a propensity value in the range from 0 to 1, with the latter signifying its presence in all structures. Further, in Figure 2 the H-bond propensity differences ΔP ($=P_{agonist} - P_{antagonist}$) versus residue number for CB1R are shown as green dots. Negative (positive) values of ΔP indicate weakened (enhanced) H-bonds for agonist structures compared to antagonist structures. TMs are indicated by vertical black, dotted lines. The histogram to the right in Figure 2 depicts the ΔP densities. The ΔP density distribution resembles a normal distribution, possibly with long tails. However, the outliers H-bond ΔPs, as determined by a cut-off that can be set to $|\Delta P| > 0.5$ (above (below) the upper (lower) dotted horizontal lines) are of particular interest. This cut-off value for $|\Delta P|$ is appropriate for the ER ensembles, consistent with our

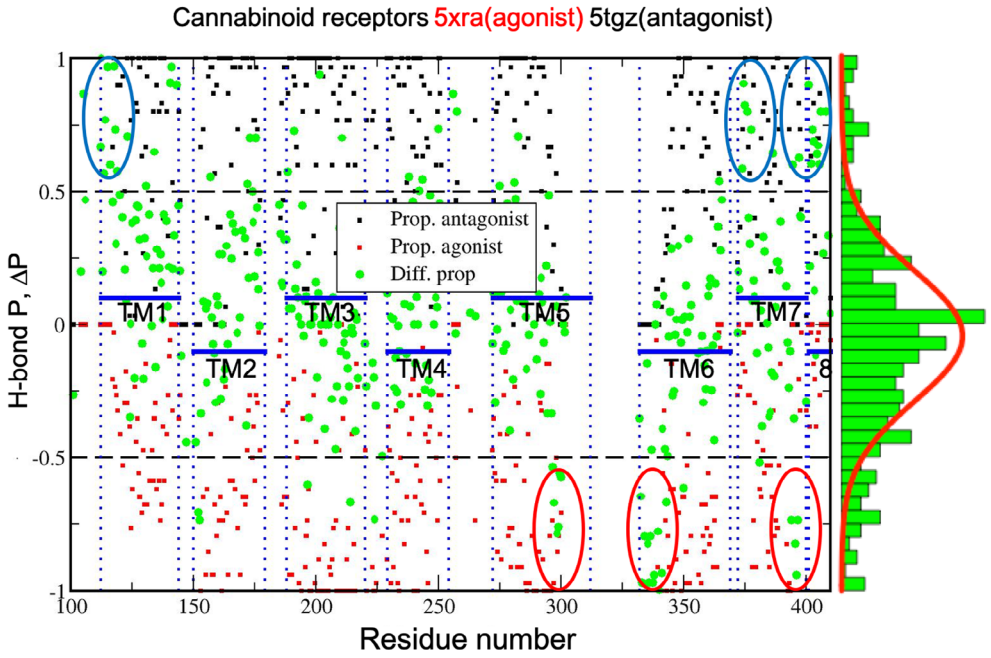

FIGURE 2 Comparison of H-bond propensities P, for CB1R ER structures with agonist (PDB ID: 5xra, -P, red dots) and antagonist (PDB ID: 5tgz, P, black dots) versus residue number. H-bond propensity differences ΔP ($=P_{agonist} - P_{antagonist}$) versus residue number for CB1R in green dots. Negative (positive) values of ΔP indicate weakened (enhanced) H-bonds for agonist structures compared to antagonist structures. Transmembrane helices are indicated by vertical black, dotted lines. The red and blue ovals are added to guide the eyes towards important clusters. |ΔP| values larger than 0.5 are mapped to the X-ray structure (PDB ID: 5xra) in Figure 3. The histogram to the right depicts the ΔP densities. CB1R, cannabinoid receptor 1; ER, ensemble refinement

earlier work[50] and it can be justified by performing a visual normality test (the quantile-quantile, or Q-Q, plot). We remark that a similar analysis of MD simulation results appears to necessitate setting a lower cut-off according to the same normality criterium. It is plausible that this either could be because of limited sampling in usual trajectory lengths or the force field's tendency to over-structure secondary structure elements, such as helices.[71,72] The red and blue ovals in Figure 2 are added to guide the eyes towards important clusters. Interestingly, the main chain H-bonds within red ovals are confined to the intracellular regions of TMs 5, 6, and 7, while those within blue ovals are confined to the extracellular regions of TMs 1 and 7, and helix 8. In order to compare ensembles obtained from different structures, the difference in propensities is mapped to the B-factor column in an appropriate structure and color-coded accordingly (from blue, $\Delta P = 1$, to red, $\Delta P = -1$), using the python script data2bfactor.py executed from PyMOL.

In Figure 3, the transition of CB1R from antagonist to agonist binding mode is illustrated in terms of propensity differences (ΔP) mapped to the agonist-bound structure of CB1R (CB1R is shown as the green cartoon ribbon and the agonist as a purple sticks). In Figure 3a, main chain donor atoms (N) and acceptor atoms (O) are colored according to ΔP in the range from −1 (red) to 1 (blue), corresponding to stabilization or appearance of a hydrogen bond ($\Delta P > 0$, blue) or destabilization or disappearance of a hydrogen bond ($\Delta P < 0$, red). Only values of $|\Delta P| > 0.5$ are mapped onto the receptor structure since they are deemed unlikely to be due to random chance given the ΔP distribution as argued above. The upper, extracellular part of TMs 1, and 7 (forming parts of the orthosteric site) as well as helix 8 are stabilized, $\Delta P > 0.5$ (blue spheres), $\Delta P < -0.5$ (red spheres). Destabilizations are observed in the lower, intracellular part of the receptor, particularly in TMs 5, 6, and 7. In contrast helix 8 is stabilized. In Figure 3b, the changes in main chain/side chain H-bond are depicted. The stabilizations are in the upper part of the receptor, whereas the destabilizations are in the lower part of the receptor, both corresponding to a |ΔP| greater than 0.5. Changes in H-bond propensities for side chain/side chain interactions are elucidated in Figure 3c. Note that the rms deviations for ligands in the orthosteric sites (agonists and antagonists) are, in general, low (less than 0.2 Å). It is remarkable that most propensity changes are observed for mainchain H-bonds that are primarily confined to TMs. This reflects the rearrangements of the helix orientations in the active versus inactive receptor conformations. The destabilization or increased mobility of main chain H-bonds of the intracellular parts of TMs 5, 6, and 7 is interpreted as a footprint

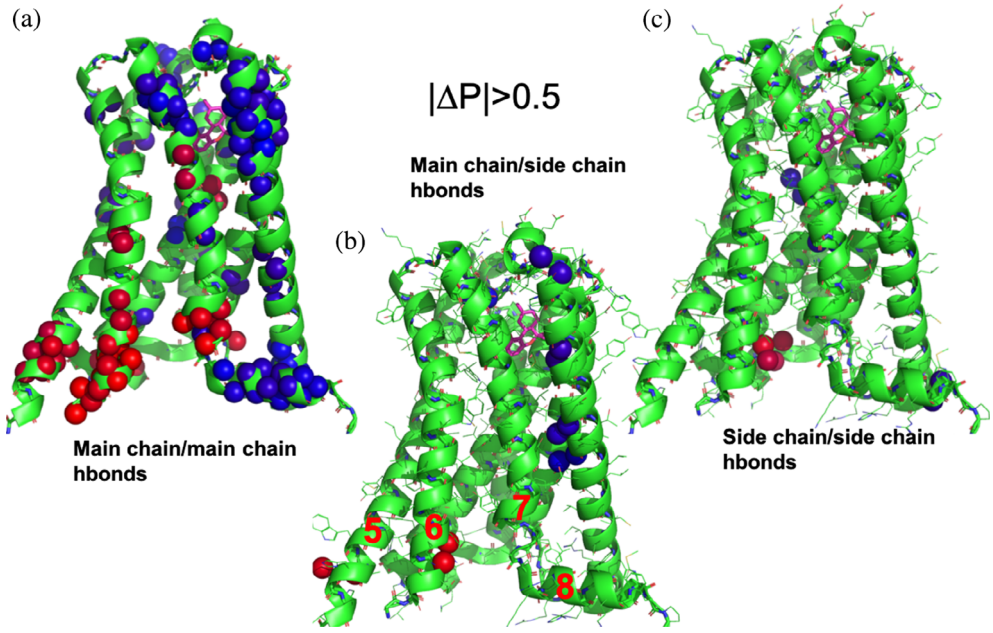

FIGURE 3 Transition of CB1R from antagonist mode to agonist mode in terms of propensity differences (ΔP) mapped to agonist-bound structure of CB1R (PDB ID: 5xra, green cartoon, agonist purple stick model). (a) Main chain donor atoms (N) and acceptor atoms (O) are colored according to ΔP in the range from −1 (red) to 1 (blue), corresponding to stabilization or generation of a hydrogen bond (ΔP > 0, blue) or destabilization or disappearance of a hydrogen bond (ΔP < 0, red). Only values of |ΔP| > 0.5 are mapped since they unlikely to be due to chance. The upper, extracellular part of the receptor as well as helix 8 are stabilized, ΔP > 0.5 (blue spheres), ΔP < −0.5 (red spheres). Destabilizations are observed in the lower, intracellular part of the receptor, particularly in TM3, 5, and 6. (b) Main chain/side chain H-bond changes. Stabilizations in upper and destabilizations in lower part of the receptor, |ΔP| > 0.5. (c) Side chain/side chain H-bond changes. CB1R, cannabinoid receptor 1; TM, transmembrane

of the allosteric action of the agonist and suggests increased flexibility of this part of the receptor, allowing various heterotrimeric G proteins or effectors such as GRKs and further β-arrestins bindings to cytoplasmic side, conveying the signaling diversity. Interestingly, analyses and comparisons of the H-bonds in the static agonist and antagonist-bound structures lead to a somewhat different picture. In Figure S1 the H-bond differences between the static structures have been mapped to the agonist-bound structure and compared to the data presented above. The clustering of the main chain H-bonds observed resulting from the ER analysis is not present in the X-ray analyses where differences are evenly distributed without any clusters.

According to the prevailing activation hypotheses, the conformational changes observed at the cytoplasmic side of the receptor are orchestrated by rearrangements in the microswitch motifs: CWxP in TM6, DRY in TM 3, and NPxxY in TM 7, where the last two microswitches are engaged in forming the G protein-binding site. TM 2 has recently been implicated in the allosteric modulation of CB1R.[73] The changed H-bond patterns in the vicinity of the microswitch motifs are visualized in Figure 4, where the DRY, CWxP and NPxxY are highlighted in gray rectangles. The active structure exhibits changed H-bond patterns close to the microswitches as compared with those in the inactive structure.[63,64] In the DRY motif, the canonical salt bridge D213$^{3.49}$-R214$^{3.50}$ is absent in the active X-ray structure with a corresponding H-bond propensity of 0 in the ER ensemble, but even though present in the inactive X-ray structure, it has a reduced H-bond propensity of 0.7 in the ER ensemble (the numbers in superscript denote the amino acid positions according to Ballesteros and Weinstein notation[74]). The side chain H-bond R214$^{3.50}$NH2-Y224$^{3.60}$OH is absent in active X-ray structure with propensity of 0, while it is present in inactive X-ray structure with propensity of 0.53. The main chain H-bond R214$^{3.50}$O-I218$^{3.54}$N is destabilized: present in active and inactive X-ray structures, however with a propensity of 0.33 for the active ensemble and propensity of 1 for the inactive ensemble. In contrast, the main chain H-bond A210$^{3.46}$O-D213$^{3.49}$N is stabilized in the active structure (present in X-ray with a propensity of 0.7 in ER ensemble) while it is missing in the inactive X-ray structure with a propensity of 0 in the ensemble. In the neighborhood of the toggle switch motif CWxP (including F200$^{3.36}$) the H-bonds for V350$^{6.42}$O-I353$^{6.45}$N and S199$^{3.35}$O-S203$^{3.39}$OG are destabilized in the active structure (missing in the X-ray structure, H-bond propensity in ER is 0), but present in the inactive X-ray structure (H-

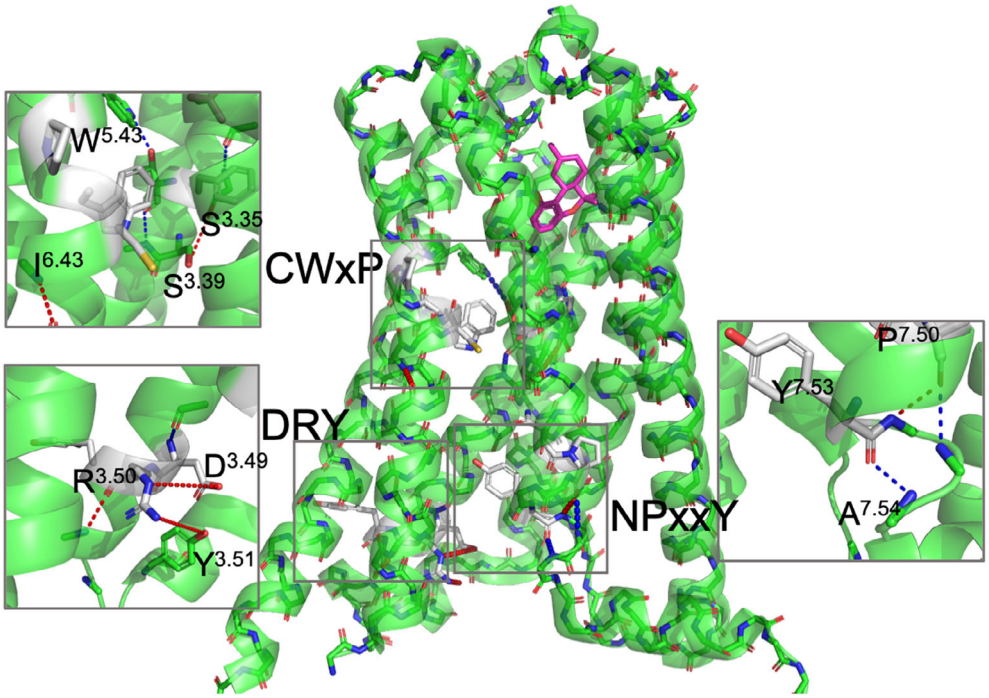

FIGURE 4 The H-bond propensity differences (ΔP) mapped to active structure of CB1R (PDB ID: 5xra). The vicinity of the microswitch motifs DRY, CWxP, and NPxxY highlighted in rectangles. Active structure has changed H-bond pattern close to microswitches as compared to inactive structure. In the DRY motif the canonical salt bridge D213$^{3.49}$-R214$^{3.50}$, the side chain H-bond R214$^{2.50}$NH2-Y215$^{3.51}$OH and main chain H-bond D213$^{3.49}$-O-I218$^{3.54}$N are destabilized while main chain H-bond A210$^{3.46}$O-D213$^{3.49}$N is stabilized. The toggle switch motif CWxP including F200$^{3.36}$ exhibits main chain H-bond destabilization for I353$^{6.45}$N-V350$^{6.42}$O and V196$^{3.32}$O-S199$^{3.35}$N, and for side chain H-bond S199$^{3.35}$O-S203$^{3.39}$OG while main chain H-bond T201$^{3.36}$O-V204$^{3.39}$N and side chain H-bond T201$^{3.36}$OG1-W279$^{5.43}$NE1 are stabilized. For the NPxxY the main chain H-bond P394$^{7.50}$O-A398$^{7.54}$N has shifted to P394$^{7.50}$-L399$^{7.55}$N destabilized while Y397$^{7.53}$O-R400$^{7.56}$N is stabilized. CB1R, cannabinoid receptor 1

bond propensity 0.6). The main chain H-bond V196$^{3.32}$O-S199$^{3.35}$N is present in both X-ray structures but H-bond propensity is 0.13 in the active structure and 0.735 in the inactive structure. The main chain H-bond T201$^{3.36}$O-V204$^{3.39}$N and side chain H-bond T201$^{3.36}$OG1-W279$^{5.43}$NE1 are stabilized that is, present in the active X-ray structure (H-bond propensities 0.97 and 0.73, respectively) and missing in the inactive structure. For the NPxxY, the main chain H-bond P394$^{7.50}$O-A398$^{7.54}$N has shifted to P394$^{7.50}$O-L399$^{7.55}$N while Y397$^{7.53}$O-R400$^{7.56}$N is stabilized. Qualitatively, the observations based on X-ray structures and ER are similar, but ER analysis reflects the plasticity of the protein.

2.2 | The influence of microswitch motifs on receptor conformational transition activation of β_2AR

The results obtained for CB1R encouraged us to look into another system where the GPCR has been crystallized in complex with agonist as well as antagonist. We choose to consider one of the most intensively studied receptors, the β_2AR. The agonist-bound structure (PDB ID: 4lde, resolution: 2.8 Å) has a T4-lysozyme fused to ICL3 and was co-crystallized with a nanobody, while the antagonist-bound structure only has T4-lysozyme fused to ICL3 (PDB ID: 2rh1, resolution: 2.4 Å). From Figures S2d,h, it can be seen that the TM regions does not have close contacts to their neighbors in the crystal. The same ER procedure as used for CB1R was applied. A summary comparing structural aspects of available X-ray structures can be seen in Figure S2.

The calculated main chain H-bond propensities, P, are shown for β_2AR ER structures with agonist (PDB ID: 4lde, –P, red dots) and antagonist (PDB ID: 2rh1, P, black dots) plotted versus residue number (Figure S3). The H-bond propensity differences ΔP versus residue number are shown in green dots. Negative (positive) values of ΔP indicate weakened (enhanced) H-bonds for agonist structures compared to antagonist structures. TMs are indicated by vertical black, dotted lines. The histogram to the right in Figure S3 depicts the ΔP densities. The red and blue ovals in Figure S3 are added to guide the eyes towards important clusters. Interestingly, the main chain H-bonds within red ovals are confined to the intracellular

regions of TMs 5, 6, and 7, thus similar to what was observed for CB1R, while those within blue ovals are confined to the extracellular regions of TMs 4 and 7. In Figure S4, the ΔP values are mapped to the active and agonist-bound structures. The similarities of the patterns seen in Figure 3 for the intracellular parts of TMs 5, 6, and 7 are striking.

In Figure S5, the H-bond propensity differences (ΔP) are mapped to active structure of β_2AR. Again, the vicinity of the microswitch motifs DRY, CWxP, and NPxxY are highlighted in gray rectangles. The active structure has changed H-bond pattern close to the microswitches as compared to the inactive structure. In the DRY motif, the salt bridge D130$^{3.49}$-R131$^{3.50}$ is absent, main chain H-bond V129$^{3.48}$O-Y132$^{3.51}$N and side chain H-bond D130$^{3.49}$OD1-S1433$^{4.55}$OG are destabilized while side chain H-bond T68$^{2.39}$OG1-R131$^{3.50}$NH1 is stabilized. In the CWxP motif, the two main chain H-bonds F282$^{6.44}$O-C285$^{6.47}$N and P288$^{6.50}$O-I291$^{6.53}$N are present in the inactive X-ray structure, but missing in the active X-ray structure, while those of T281$^{6.43}$O-C285$^{6.47}$N, T283$^{6.45}$O-W286$^{6.48}$N, and W286$^{6.48}$O-F290$^{6.52}$N are present in both active and inactive X-ray structures. The propensities for the two former main chain H-bonds are 0.2–04 (active) and 0.8–1.0 (inactive), and 0.6–0.7 (active) and 1 (inactive) for the last three. Hence the H-bonds are somewhat destabilized in the ensemble of active structures. Interesting, the H-bonds flank P288$^{6.50}$, suggesting its hinge function. In NPxxY, Y326$^{7.53}$O-S329$^{7.56}$N and N318$^{7.45}$O-N322$^{7.59}$N are destabilized.

The observed destabilizations of the intracellular parts of TMs 5, 6, and 7 in the present study have been supported by HDX-MS experiments that show binding of an agonist to β_2AR results in enhanced deuterium exchange,[29] indicating less protected H-bonds compared to that of apo β_2AR for the intracellular region of TM 6. In contrast, binding of an inverse agonist resulted in reduced deuterium exchange, indicating more protected H-bonds of the intracellular parts of TMs 5, 6, and 7. Further, NMR experiments show that agonist binding increases the structural heterogeneity of β_2ARs cytoplasmic domains.[36] These observations are consistent with the results presented here and supports the notion that footprints of changed flexibility and allostery can be detected by ER analysis of X-ray diffraction data. It is particularly interesting that TMs 5, 6, 7 are destabilized for β_2AR because its active agonist structure apart from being fused to T4-lysozyme also contains a stabilizing nanobody. Furthermore, MD simulations of β_2AR in active and inactive conformations indicate that the agonist-bound structure shows a more dynamic state for TMs 5, 6, and 7 than that of an inverse agonist-bound state in agreement with the results of the present analysis.[13] An analysis of H-bond propensity differences of MD trajectories deposited in GPCRmd[75] for a comparison with the propensity differences calculated from the ER ensembles was performed (Figure S6). It is recognized that the simulated differences main chain/main chain H-bond propensities between agonist- and antagonist-bound trajectories (Figure S6a) are much smaller than those seen in the corresponding ER ensemble (Figure S3). This might be due to limited sampling (3 × 500 ns) or because of the tendency of MD force fields to over-structure secondary structure elements, such as helices.[71,72] Intriguingly, however, there appears to be changes commensurate with those observed in the NPxxY motif of TM 7 (Figure S6b).

2.3 | Agonist and antagonist binding affect dynamics differently

Encouraged by the observed differences in H-bond stability for agonist and antagonist-bound structures, we complemented the H-bond analysis with PCA on the ER structures obtained for the two receptor systems considered above. PCA is a useful technique for reducing the dimensionality of general applicability, case in point being analysis of MD simulations and ER data.[76–83] Since the H-bond analysis indicated different stability in the upper and lower parts of the receptors, PCA was performed separately on the upper (extracellular) and the lower half (intracellular) of the TM region of the receptors. The selection of upper and lower half was based on the TM definitions in GPCRdb[84] by dividing each TM in two equal-sized halves with only minor modifications (see Table S2 for the exact definitions). The analysis reveals that the active and inactive ER structures exhibit different amount of variation in their respective essential subspaces gauged by the two major principal components.

In Figure 5a,b (5c,d) the first two principal components are shown for CB1R (β_2AR), where Figure 5a (5c) are the results from the lower TM region and Figure 5b (5d) are the results from the upper TM region. Results from active (inactive) conformations are in green (red) dots. Qualitatively, the two receptors react in much the same manner upon agonist and antagonist binding, respectively. For the lower TM region, the active conformations exhibit a larger essential-subspace variability compared to those of inactive conformations, indicating larger plasticity. In contrast, the smaller subspace of the inactive conformations indicate antagonist induced rigidity in the lower TM region. The opposite is seen for the upper TM region. Here, the inactive conformations span a somewhat larger subspace than the active

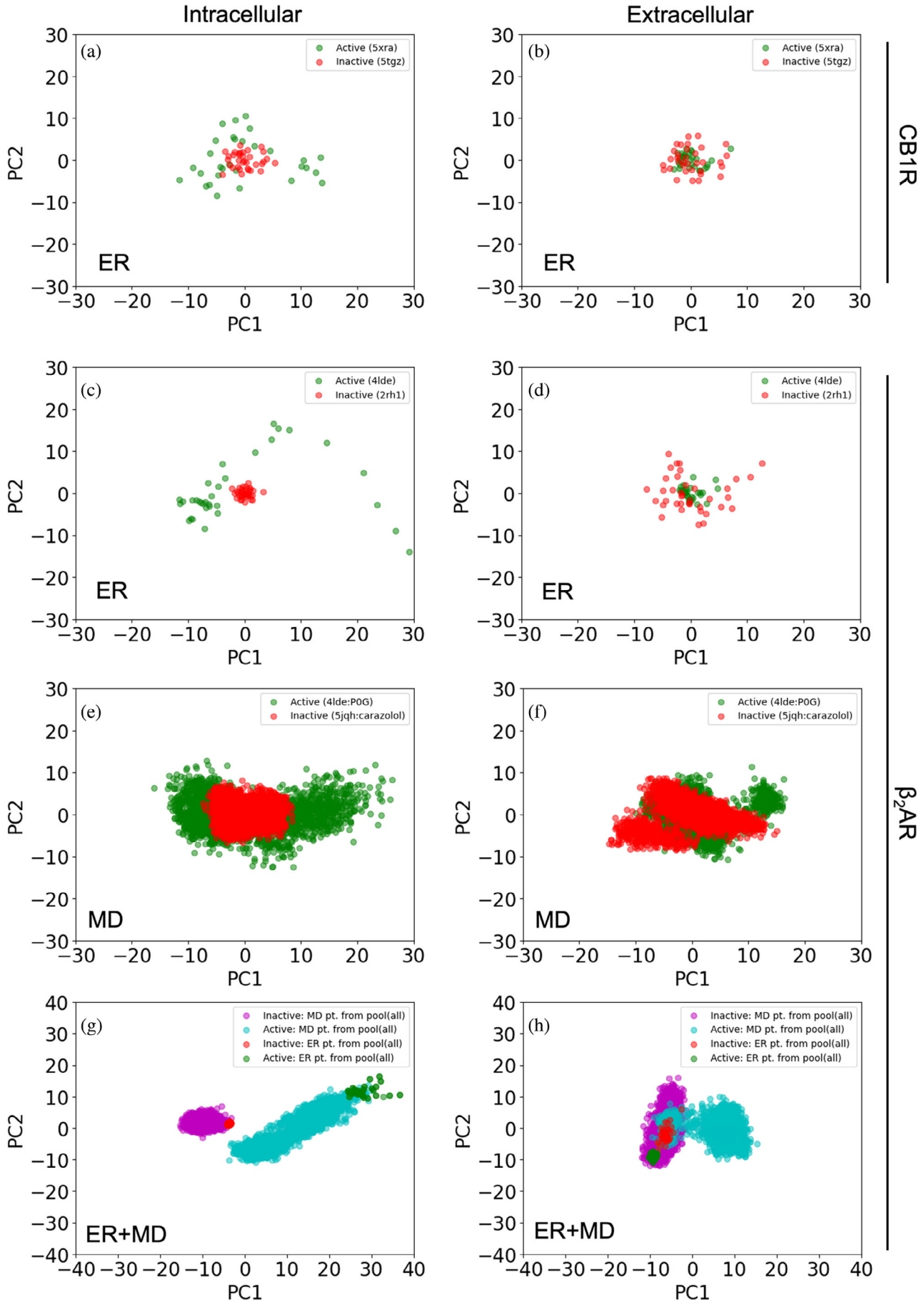

FIGURE 5 Legend on next page.

conformations. The observed difference in conformational flexibility in the lower TM region supports the notion of conformational selection of effector proteins.

Analysis of MD trajectories deposited in GPCRmd[75] strongly supports the results from the ER analyses in terms of the above phenomena. The relevant simulations offering points of comparison with ER of β_2AR were those of apo and agonist-bound (PDB ID: 4lde, simulated without and with agonist) as well as simulations with inverse agonist-bound (PDB ID: 5jqh). The analysis was carried out in a similar fashion to that of the ER ensembles (for the three available replicas of length 500 ns for each structure). In Figure 5e,f results from PCA performed on MD trajectories (first replica) are shown where data from active conformations are represented in green dots and those of inactive conformations are in red dots. For the TM lower (intracellular) region, the MD simulation data (Figure 5e) show the same qualitative behavior as obtained for the corresponding analysis of the ER structures (Figures 5a,c). Hence, the larger subspace spanned by the active MD conformations indicate agonist-induced plasticity, which is in contrast to the smaller subspace for the inactive conformations induced by the inverse agonist. The smaller subspace reflects the rigidity of the inactive conformations. In Figure 5f, PCA data for the upper (extracellular) region shows that the active conformations span a smaller subspace than that of the intracellular region supporting the observation obtained from the ER structures. However, in Figure 5f, the subspaces spanned by the active and inactive MD conformations, obtained from the first replica, are of same size, which is in contrast to what is seen for the ER structures in Figures 5b,d. In Figure S7, PCAs are shown for the remaining MD replicas. Results for the intracellular region resemble those of Figure 5e. However, for the extracellular region the subspaces spanned by the inactive conformations are smaller than those of active conformations. A likely explanation for this observation is that the ER structures are restrained by an extracellularly bound nanobody (see Figures S1a-d), which is lacking in the MD simulations.

While we have interpreted the relative trends of conformational flexibility between the intra- and extracellular segments of the receptors and made drawn parallels, we note that it would not be correct to directly compare PCAs performed on sets of structures gathered by different means (even of different receptors) and analyzed independently. Instead, for a direct comparison between the conformational spans observed in ER and MD simulations, we have performed PCA on the pooled dataset for β_2AR (Figure 5g,h). The configurations sampled by MD and resolved by ER exhibit overlap within each of the distinct groups (active vs. inactive; intracellular vs. extracellular), consistent with the results and interpretations alluded to in the above. Additionally, it can be seen that active conformations are separated from the inactive conformations only for the intracellular part of the receptor (Figure 5g). All data points exhibit overlap for the extracellular region (Figure 5h). In Figure S8, trajectories from the MD simulations of apo and agonist-bound β_2AR have been analyzed and compared. The agonist-bound conformations show larger conformational plasticity than the apo structures both for the extracellular and intracellular regions.

2.4 | Agonist binding in sodium-ion excluding A_{2A}AR mutants

Certain class A GPCRs are known to bind an allosteric sodium ion and the resulting allosteric effect is believed mediated largely by the negatively charged $D^{2.50}$.[12] While most receptors undergo a ligand-dependent activation/deactivation, some sodium can compete with agonist binding.[85–92] A_{2A}AR is a prime example of a GPCR that is allosterically regulated by sodium.[12]

Here, ER analysis is applied to structures of wildtype and two sodium excluding mutants of A_{2A}AR in complex with the same agonist. The structures investigated are unique in that the same molecular construct (apart from the single point mutations) are crystallized in the same space group, making this system particularly well-suited

FIGURE 5 Results from PCA of ER structures of CB1R (a and b) and β_2AR (c and d). (a and c) Intracellular—lower TM half; (b and d) Extracellular—upper TM half. The plots show the first (PC1) and the second (PC2) principal component against each other. Results from active structures (PDB ID: 5xra and 5tgz) in green dots and from inactive structures (PDB ID: 4lde and 2rh1) in red dots. The active structures exhibit larger subspace (plasticity) than the inactive structures for lower TM halves (a and c), while for the upper TM halves the inactive structures exhibit larger subspace. Panels e (intracellular) and f (extracellular) depict results of PCA of MD trajectories for β_2AR. Active conformations (starting structure PDB ID: 4lde) in green; inactive conformations (starting structure PDB ID: 5jqh) in red dots. Panels g (intracellular) and h (extracellular) show results for PCA analyses on pooled datasets of ER and MD for β_2AR. Active conformations in green and cyan dots for ER and MD, respectively; inactive conformations in red and purple dots for ER and MD, respectively. β_2AR, β_2 adrenergic receptor; CB1R, cannabinoid receptor 1; ER, ensemble refinement; MD, molecular dynamics; PCA, principal component analysis; TM, transmembrane

for studying the isolated impact of sodium-ion-excluding mutations in the ion-coordinating residues and shed light on the mechanisms of allostery in the receptor[93] (see Table S1). One mutant ($S91^{3.39}A$) exhibits normal activity while the other ($D52^{2.50}N$) inhibits activity measured by agonist (UK432097)-induced cAMP accumulation. The same ER analysis as above is applied on wildtype (PDB ID: 3qak, resolution 2.7 Å), and mutant $S91^{3.39}A$ (PDB ID: 5wf6, resolution 2.9 Å) and $D52^{2.50}N$ (PDB ID: 5wf5, resolution 2.6 Å).[12] All three structures have the full agonist UK432097 bound.

In Figure 6, the main chain H-bond propensity differences between the mutant structures and wildtype $A_{2A}R$ have been mapped to the structure of the $D52^{2.50}N$ mutant (PDB ID: 5wf5, Figure 6a) as well as to that of the $S91^{3.39}A$ mutant (PDB ID: 5wf6, Figure 6b). The red arrows point to the position of the mutations. As above, blue and red spheres depict donor and acceptor atoms with $|\Delta P| > 0.5$, blue >0, red <0. Notable ΔP differences are seen when the mutants are compared to wildtype. Red ovals encircle the intracellular part of TM 7 harboring the microswitch NPxxY motif, where the major differences are observed. In the $D52^{2.50}N$ mutant, the ΔP for main chain H-bonds $V283^{7.48}O$-$I287^{7.52}N$ and $N280^{7.45}O$-$N284^{7.49}N$ is 0.6, while that of $F286^{7.51}O$-$A289^{7.54}N$ is 0.84. Further, the side chain $N24^{1.50}ND2$ shifts H-bond partner from $V282^{7.47}O$ to its neighbor $S281^{7.46}O$, which establishes another H-bond to $N52^{2.50}ND2$ to stabilize the changed bulge conformation of TM7.

In Figure S9, the H-bond network of an antagonist-bound inactive structure of wildtype $A_{2A}AR$ (PDB ID: 4eiy, resolution 1.8 Å) has been compared to H-bond networks of $D52^{2.50}N$ and $S91^{3.39}A$ mutants. Note, in contrast to the three structures discussed above this structure contains an allosteric sodium ion. The ΔP differences are not as pronounced as in Figure 6 which may be explained by the presence of the sodium ion in 4eiy. However, in Figure S9b H-bonds flanking the NPxxY motif in TM 7 ($V283^{7.48}N$- $N280^{7.45}O$ and $Y290^{7.55}N$-$F286^{7.51}O$) are destabilized compared to those in Figure S9a. In TM 6 H-bonds in Figure S9b ($S234^{6.36}N$-$A231^{6.33}O$ and $K233^{6.35}N$-$H230^{6.33}O$) are also destabilized compared to those in Figure S9a. Hence, the active mutant, $S91^{3.39}A$, is more destabilized than the inactive structure, $D52^{2.50}N$, in accordance with the observations above. Hence, the $D52^{2.50}N$ mutant is primarily captured in an inactive conformation stabilized by an enhanced H-bond network. This finding is supported by NMR spectroscopy studies which showed that $D52^{2.50}N$ mutant has an altered conformational dynamics intercellularly without changing the receptors conformation extracellularly.[12] In $S91^{3.39}A$ (Figure 6), the corresponding ΔPs are less than 0.2, suggesting that this part of TM 7 is stabilized by the $D52^{2.50}N$ mutant. Interestingly, this observation as well as the observation of changed C_α positions fit with observation that the $D52^{2.50}N$ mutant has increased temperature stability.[12] Specifically, the change in melting temperature, ΔT_m, for $D52^{2.50}N$ is 8°C larger than the ΔT_m for wildtype and $S91^{3.39}A$, while it is 6°C larger when they are bound to UK432097. The stabilization of this vital region of the receptor indicates an induced rigidity conferred by the mutation, which prevents proper conformational

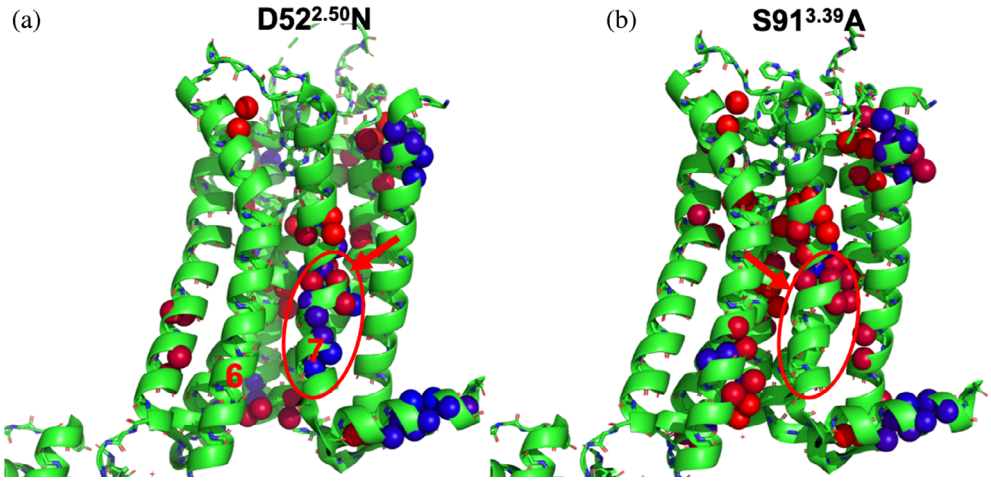

FIGURE 6 Comparison of agonist-bound structures of $A_{2A}AR$ mutations excluding the allosteric sodium to the structure of agonist-bound wt in terms of propensity differences (ΔP) mapped to mutant structure (a) $D52^{2.50}N$ (5wf5) and (b) $S91^{3.39}A$ (5wf6). The red arrows point to the position of the mutations and red ovals depict the part of TM7 which contains the NPxxY microswitch motif. In (a) ($D52^{2.50}N$) the ΔP for mainchain H-bonds $V283^{7.48}O$-$I287^{7.52}N$ and $N280^{7.45}O$-$N284^{7.49}N$ is 0.6, while that of $F286^{7.51}OA289^{7.54}N$ is 0.84. In (b) corresponding $|\Delta P|$s are less than 0.2. There are no observed changes in TM 7. So, this part of TM 7 is not stabilized in contrast to what is observed for the mutant D52N. TM, transmembrane

selection of G proteins. It is of further note that analysis of structures of thermophilic and mesophilic proteins show consistent increase in the number of H-bonds with increased thermostability.[94]

2.5 | Conclusion and outlook

ER analysis has been applied on available X-ray crystallographic structures of CB1R, β_2AR, and A_{2A}AR co-crystallized with agonists and antagonist/inverse agonists representing active and inactive conformations, respectively. Our results support the notion of a conformational selection mechanism when GPCRs bind G proteins or molecular effectors and might reflect their preferences for different G proteins and consequently affect the capacity for biased signaling. However, the observed changed H-bond patterns and plasticities between active and inactive states do not reveal which specific effectors bind in the active state.

Analysis of ΔPs showed that the transition of CB1R and β_2AR from their antagonist-bound (inactive) conformation to the corresponding agonist-bound (active) conformation results in destabilization of H-bonds in the intracellular parts of TM 5, 6, and 7. In particular, main-chain ΔP shifts are observed in the vicinity of the micro-switches. In CB1R, the DRY motif has two changed main chain H-bonds, the CWxP and NPxxY motifs has three each while for β_2AR the corresponding changes are one, two and two, respectively. These shifts reflect the changes in TM orientations and probably support the enhanced mobility of the intracellular region. It is noteworthy that the overall destabilizations of the intracellular regions of the two examined systems are alike, even though some details differ.

PCA analysis performed on the ER ensembles as well as available MD simulation data support that the active and inactive structures exhibit different essential subspace variations. For the lower (intracellular) TM region, the active conformations exhibit a larger subspace compared to those of inactive conformations, indicating larger conformational plasticity. In contrast, the smaller subspace of the inactive conformations indicates antagonist-induced rigidity in the lower TM region. The opposite is seen for the upper (extracellular) TM region. Here, the inactive conformations span a somewhat larger subspace than the active conformations. The observations are supported by HDX-MS experiments, which revealed that binding of an agonist induced the largest degree of conformational plasticity in the intracellular region while an inverse agonist was shown to be the most stabilizing agent.

The impact of sodium-ion binding on A_{2A}AR was examined by applying ER analyses to structures of wild-type and two sodium excluding mutants of A_{2A}AR in complex with same agonist. One mutant, S91$^{3.39}$A, exhibits normal activity while the other, D52$^{2.50}$N inhibits activity. The main chain H-bond pattern of the latter is stabilized in the intracellular part of TM 7 containing the NPxxY motif, indicating an induced rigidity by the mutation prevents conformational selection of G proteins. These observations are supported by the increased temperature stability of D52$^{2.50}$N which are further supported by analysis of structures of thermophilic and mesophilic proteins, which show a consistent increase in the number of H-bonds with increased thermostability.

The importance of water molecules in stabilizing the active and inactive conformations of GPCRs has been discussed in some detail previously,[52–62] on the basis of structural and dynamic measurements, revealing the crucial role for internal water networks in receptor activation. There appears to be a part of the protein-water network that is maintained between active and inactive states, while another part is rearranged upon activation. As an example, osmotic stress studies underline the importance of bulk waters in the visual receptor rhodopsin,[61,62] which may suggest a model of GPCR activation where the receptor becomes solvent-swollen upon activation. These results encourage future H-bond analysis of water mediated networks in ER ensembles and MD simulations of GPCRs.

3 | EXPERIMENTAL PROCEDURES

3.1 | Time-averaged X-ray restrained ER

Structure factors and rebuild model coordinates for structures listed in Table S1 in combination with their resolution, space group, and co-crystallized proteins were downloaded from the PDB_REDO server.[95] Phenix.ready_set was used to generate ligand restraint and add explicit hydrogen atoms to the downloaded PDB files. Refinement was performed with the most recent implementation of Phenix.refine[96] using one TLS group per domain. ER was performed using the implementation in Phenix version 1.10.1-2155[97] according to refs. 46, 47. A grid search was performed for optimal values of p_{TLS} (0.6–0.9), T_{bath} (2.5, 5, and 10), and tx (0.5, 1, and 2x the resolution-dependent value) scoring for the lowest final R_{free}. After selection of optimal parameters, five random seed repeats were generated for each structure.

3.2 | Analysis of ensembles

The H-bond networks in the ensemble-refined structures were analyzed using the python script list_mc_hbonds.py (http://pldserver1.biochem.queensu.ca/~rlc/work/pymol/list_mc_hbonds.py, accessed on 08/08/2017) executed from PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC), with distance cutoff = 3.5 Å and angle cutoff = 45°. Hence, the propensity of a H-bond was determined as the number of structures in which it was present relative to the total number of structures in the ensemble. The propensity of a H-bond is consequently in the range from 0 to 1, the latter meaning presence in all structures. In order to compare ensembles obtained from different structures, the difference in propensities are mapped to the B-factor column in an appropriate structure and color-coded accordingly (from blue, $\Delta P = 1$, to red, $\Delta P = -1$), using the python script data2bfactor.py executed from PyMOL.

3.3 | MD simulations

MD simulations of the β_2AR in a bilayer environment were taken from the GPCRmd database[75] (https://gpcrmd.org) entries corresponding to PDB IDs 4lde[69] and 5jqh.[98] Details of how these simulations were set up and performed are available in the published work.[75] Three trajectories (each of length 500 ns) for each structure were analyzed. H-bond propensities were analyzed using MDAnalysis[99,100] using the identical geometric definition as for the ER ensembles (with distance cutoff = 3.5 Å and angle cutoff = 45°).

3.4 | PCA of ensembles and MD trajectories

PCA was used for reducing the dimensionality of MD simulation data or ER data in order to extract larger amplitude motions. The method works by computing the eigenvalues of the mass-weighted covariance matrix of the atomic positional fluctuations.[80] To perform the PCA, we used pytraj[101] and scikit.[102] The atomic positions for α-carbons in the protein selection were used in the PCA after aligning the individual frames of the ER ensemble, MD trajectory, or a pooled dataset of both onto the mean structure with pytraj.superpose.[101]

AUTHOR CONTRIBUTIONS

Jesper J. Madsen: Formal analysis (equal); investigation (equal); methodology (equal); visualization (equal); writing – original draft (equal). **Libin Ye:** Formal analysis (equal); writing – review and editing (equal). **Thomas M. Frimurer:** Formal analysis (equal); writing – review and editing (equal). **Ole H. Olsen:** Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); supervision (equal); visualization (equal); writing – original draft (equal).

ACKNOWLEDGMENT

The authors thank David Kouvchinov for proofreading the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

DATA AVAILABILITY STATEMENT

Structural ensembles are available for download at Zenodo (https://zenodo.org/record/4896421).[103]

ORCID

Jesper J. Madsen https://orcid.org/0000-0003-1411-9080

REFERENCES

1. Krishna Kumar K, Shalev-Benami M, Robertson MJ, et al. Structure of a signaling cannabinoid receptor 1-G protein complex. Cell. 2019;176:448–458.e12.
2. Schwartz TW, Frimurer TM, Holst B, Rosenkilde MM, Elling CE. Molecular mechanism of 7TM receptor activation—a global toggle switch model. Annu Rev Pharmacol Toxicol. 2006;46:481–519.
3. Elling CE, Frimurer TM, Gerlach LO, Jorgensen R, Holst B, Schwartz TW. Metal ion site engineering indicates a global toggle switch model for seven-transmembrane receptor activation. J Biol Chem. 2006;281:17337–17346. https://doi.org/10.1074/jbc.M512510200
4. Schwartz TW, Hubbel WL. Structural biology: a moving story of receptors. Nature. 2008;455:473–474.
5. Frimurer TM, Hogberg T. Drug design of GPCR ligands using physicogenetics and chemogenomics—principles and case studies. Curr Top Med Chem. 2011;11:1882–1901.
6. Zhou Q, Yang D, Wu M, et al. Common activation mechanism of class A GPCRs. Elife. 2019;8:1–31.
7. Weis WI, Kobilka BK. The molecular basis of G protein-coupled receptor activation. Annu Rev Biochem. 2018;87:897–919.
8. Wang W, Qiao Y, Li Z. New insights into modes of GPCR activation. Trends Pharmacol Sci. 2018;39:367–386. https://doi.org/10.1016/j.tips.2018.01.001
9. Bhattacharya S, Hall SE, Li H, Vaidehi N. Ligand-stabilized conformational states of human beta(2) adrenergic receptor: insight into G-protein-coupled receptor activation. Biophys J. 2008;94:2027–2042.
10. Tehan BG, Bortolato A, Blaney FE, Weir MP, Mason JS. Unifying family A GPCR theories of activation. Pharmacol Ther. 2014;143:51–60.

11. Venkatakrishnan AJ, Deupi X, Lebon G, et al. Diverse activation pathways in class a GPCRs converge near the G-protein-coupling region. Nature. 2016;536:484–487.
12. White KL, Eddy MT, Gao ZG, et al. Structural connection between activation microswitch and allosteric sodium site in GPCR signaling. Structure. 2018;26:259–269.e5.
13. Bhattacharya S, Vaidehi N. Differences in allosteric communication pipelines in the inactive and active states of a GPCR. Biophys J. 2014;107:422–434.
14. Katritch V, Fenalti G, Abola EE, Roth BL, Cherezov V, Stevens RC. Allosteric sodium in class A GPCR signaling. Trends Biochem Sci. 2014;39:233–244. https://doi.org/10.1016/j.tibs.2014.03.002
15. Ye L, Neale C, Sljoka A, et al. Mechanistic insights into allosteric regulation of the A2A adenosine G protein-coupled receptor by physiological cations. Nat Commun. 2018;9:1372. http://www.nature.com/articles/s41467-018-03314-9
16. Mattedi G, Acosta-Gutiérrez S, Clark T, Gervasio FL. A combined activation mechanism for the glucagon receptor. Proc Natl Acad Sci U S A. 2020;117:15414–15422.
17. Liu W, Chun E, Thompson AA, et al. Structural basis for allosteric regulation of GPCRS by sodium ions. Science. 2012;337:232–236.
18. Rosenbaum DM, Zhang C, Lyons JA, et al. Structure and function of an irreversible agonist-β2 adrenoceptor complex. Nature. 2011;469:236–242.
19. Valentin-Hansen L, Groenen M, Nygaard R, Frimurer TM, Holliday ND, Schwartz TW. The arginine of the DRY motif in transmembrane segment III functions as a balancing microswitch in the activation of the β2-adrenergic receptor. J Biol Chem. 2012;287:31973–31982. https://doi.org/10.1074/jbc.M112.348565
20. Ma N, Lee S, Vaidehi N. Activation microswitches in adenosine receptor A2a function as rheostats in the cell membrane. Biochemistry. 2020;59:4059–4071.
21. 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.
22. Knight KM, Ghosh S, Campbell SL, et al. A universal allosteric mechanism for G protein activation. Mol Cell. 2021;81:1384–1396.e6. https://doi.org/10.1016/j.molcel.2021.02.002
23. Vardy E, Roth BL. Conformational ensembles in GPCR activation. Cell. 2013;152:385–386. https://doi.org/10.1016/j.cell.2013.01.025
24. Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. Nature. 2014;508:331–339. http://www.ncbi.nlm.nih.gov/pubmed/24740064
25. Nygaard R, Zou Y, Dror RO, et al. The dynamic process of β2-adrenergic receptor activation. Cell. 2013;152:532–542.
26. Mafi A, Kim SK, Goddard WA. The mechanism for ligand activation of the GPCR-G protein complex. Proc Natl Acad Sci U S A. 2022;119:e2110085119.
27. Sušac L, Eddy MT, Didenko T, Stevens RC, Wüthrich K. A2A adenosine receptor functional states characterized by ^{19}F-NMR. Proc Natl Acad Sci U S A. 2018;115:12733–12738.
28. Ye L, Van Eps N, Zimmer M, Ernst OP, Scott Prosser R. Activation of the a 2A adenosine G-protein-coupled receptor by conformational selection. Nature. 2016;533:265–268.
29. West GM, Chien EYT, Katritch V, et al. Ligand-dependent perturbation of the conformational ensemble for the GPCR β 2 adrenergic receptor revealed by HDX. Structure. 2011;19:1424–1432.
30. Sandhu M, Touma AM, Dysthe M, Sadler F, Sivaramakrishnan S, Vaidehi N. Conformational plasticity of the intracellular cavity of GPCR−G-protein complexes leads to G-protein promiscuity and selectivity. Proc Natl Acad Sci U S A. 2019;116:11956–11965.
31. Huang SK, Pandey A, Tran DP, et al. Delineating the conformational landscape of the adenosine A2A receptor during G protein coupling. Cell. 2021;184:1884–1894.e14.
32. 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;10:1–46.
33. Lu S, He X, Yang Z, et al. Activation pathway of a G protein-coupled receptor uncovers conformational intermediates as targets for allosteric drug design. Nat Commun. 2021;12:1–15. https://doi.org/10.1038/s41467-021-25020-9
34. Koretz KS, McGraw CE, Stradley S, Elbaradei A, Malmstadt N, Robinson AS. Characterization of binding kinetics of A2AR to Gαs protein by surface plasmon resonance. Biophys J. 2021;120:1641–1649.
35. Forneris F, Burnley BT, Gros P. Ensemble refinement shows conformational flexibility in crystal structures of human complement factor D. Acta Crystallogr Sect D Biol Crystallogr. 2014;70:733–743.
36. Manglik A, Kim TH, Masureel M, et al. Structural insights into the dynamic process of β2-adrenergic receptor signaling. Cell. 2015;161:1101–1111. https://doi.org/10.1016/j.cell.2015.04.043
37. Shimada I, Ueda T, Kofuku Y, Eddy MT, Wüthrich K. GPCR drug discovery: integrating solution NMR data with crystal and cryo-EM structures. Nat Rev Drug Discov. 2018;18:59–82.
38. Wang X, McFarland A, Madsen JJ, Aalo E, Ye L. The potential of ^{19}F NMR application in GPCR biased drug discovery. Trends Pharmacol Sci. 2021;42:19–30.
39. Ma N, Nivedha AK, Vaidehi N. Allosteric communication regulates ligand-specific GPCR activity. FEBS J. 2021;288:2502–2512.
40. Wang X, Bushra N, Muschol M, Madsen JJ, Ye L. An in-membrane NMR spectroscopic approach probing native ligand-GPCR interaction. Int J Biol Macromol. 2022;206:911–916.
41. Wang X, Zhao W, Al-Abdul-Wahid S, et al. Trifluorinated keto–enol tautomeric switch in probing domain rotation of a G protein-coupled receptor. Bioconjug Chem. 2021;32:99–105.
42. Ye L, Wang X, McFarland A, Madsen JJ. ^{19}F NMR: a promising tool for dynamic conformational studies of g protein-coupled receptors. Structure. 2022;30:1–13.
43. Park PS-H. Ensemble of G protein-coupled receptor active states. Curr Med Chem. 2012;19:1146–1154.
44. Shukla AK, Westfield GH, Xiao K, et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature. 2014;512:218–222.
45. Agasid MT, Sørensen L, Urner LH, Yan J, Robinson CV. The effects of sodium ions on ligand binding and conformational states of G protein-coupled receptors—insights from mass spectrometry. J Am Chem Soc. 2021;143:4085–4089.
46. Burnley BT, Gros P, Scheikunde UUBCD. phenix.ensemble_refinement: a test study of apo and holo BACE1. Comput Crystallogr Newsl. 2013;4:51–58.

47. Burnley BT, Afonine PV, Adams PD, Gros P. Modelling dynamics in protein crystal structures by ensemble refinement. Elife. 2012;2012:1–29.
48. Ploscariu N, Burnley T, Gros P, Pearce NM. Improving sampling of crystallographic disorder in ensemble refinement research papers. Acta Crystallogr D Struct Biol. 2021;77:1357–1364.
49. Furnham N, Blundell TL, DePristo MA, Terwilliger TC. Is one solution good enough? Nat Struct Mol Biol. 2006;13:184–185.
50. Sorensen AB, Madsen JJ, Frimurer TM, et al. Allostery in coagulation factor VIIa revealed by ensemble refinement of crystallographic structures. Biophys J. 2019;116:1823–1835.
51. Fleming PJ, Rose GD. Do all backbone polar groups in proteins form hydrogen bonds? Protein Sci. 2005;14:1911–1917.
52. Angel TE, Gupta S, Jastrzebska B, Palczewski K, Chance MR. Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. Proc Natl Acad Sci U S A. 2009;106:14367–14372.
53. Nygaard R, Valentin-Hansen L, Mokrosinski J, Frimurer TM, Schwartz TW. Conserved water-mediated hydrogen bond network between TM-I, -II, -VI, and -VII in 7TM receptor activation. J Biol Chem. 2010;285:19625–19636. https://doi.org/10.1074/jbc.M110.106021
54. Tichy A-M, So WL, Gerrard EJ, Janovjak H. Structure-guided optimization of light-activated chimeric G-protein-coupled receptors. Structure. 2022;30:1075–1087.e4. https://doi.org/10.1016/j.str.2022.04.012
55. Standfuss J, Edwards PC, D'Antona A, et al. The structural basis of agonist-induced activation in constitutively active rhodopsin. Nature. 2011;471:656–660.
56. Yuan S, Filipek S, Palczewski K, Vogel H. Activation of G-protein-coupled receptors correlates with the formation of a continuous internal water pathway. Nat Commun. 2014;5:4733.
57. Lee Y, Kim S, Choi S, Hyeon C. Ultraslow water-mediated transmembrane interactions regulate the activation of A2A adenosine receptor. Biophys J. 2016;111:1180–1191. https://doi.org/10.1016/j.bpj.2016.08.002
58. Venkatakrishnan AJ, Ma AK, Fonseca R, et al. Diverse GPCRs exhibit conserved water networks for stabilization and activation. Proc Natl Acad Sci U S A. 2019;116:3288–3293.
59. Bertalan É, Lešnik S, Bren U, Bondar AN. Protein-water hydrogen-bond networks of G protein-coupled receptors: graph-based analyses of static structures and molecular dynamics. J Struct Biol. 2020;212:107634.
60. Louet M, Casiraghi M, Damian M, et al. Concerted conformational dynamics and water movements in the ghrelin G protein-coupled receptor. Elife. 2021;10:1–21.
61. Chawla U, Perera SMDC, Fried SDE, et al. Activation of the G-protein-coupled receptor rhodopsin by water. Angew Chem Int Ed Engl. 2021;60:2288–2295.
62. Fried SDE, Hewage KSK, Eitel AR, et al. Hydration-mediated G protein-coupled receptor activation. Proc Natl Acad Sci U S A. 2022;119:e2117349119.
63. Hua T, Vemuri K, Nikas SP, et al. Crystal structures of agonist-bound human cannabinoid receptor CB 1. Nature. 2017;547:468–471.
64. Hua T, Vemuri K, Pu M, et al. Crystal structure of the human cannabinoid receptor CB1. Cell. 2016;167:750–762.e14.
65. Schönegge AM, Gallion J, Picard LP, et al. Evolutionary action and structural basis of the allosteric switch controlling β2AR functional selectivity. Nat Commun. 2017;8:2169.
66. Liu X, Xu X, Hilger D, et al. Structural insights into the process of GPCR-G protein complex formation. Cell. 2019;177:1243–1251.e12.
67. Hanson MA, Cherezov V, Griffith MT, et al. A specific cholesterol binding site is established by the 2.8 Å structure of the human β2-adrenergic receptor. Structure. 2008;16:897–905.
68. Cherezov V, Rosenbaum DM, Hanson MA, et al. High-resolution crystal structure of an engineered human β2-adrenergic G protein-coupled receptor. Science. 2007;318:1258–1265.
69. Ring AM, Manglik A, Kruse AC, et al. Adrenaline-activated structure of β 2-adrenoceptor stabilized by an engineered nanobody. Nature. 2013;502:575–579.
70. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res. 2012;40:370–376.
71. Best RB, Buchete NV, Hummer G. Are current molecular dynamics force fields too helical? Biophys J. 2008;95:7–9.
72. Robustelli P, Piana S, Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. Proc Natl Acad Sci U S A. 2018;115:E4758–E4766.
73. Yang X, Wang X, Xu Z, et al. Molecular mechanism of allosteric modulation for the cannabinoid receptor CB1. Nat Chem Biol. 2022;18:831–840.
74. Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. Methods Neurosci. 1995;25:366–428.
75. Rodríguez-Espigares I, Torrens-Fontanals M, Tiemann JKS, et al. GPCRmd uncovers the dynamics of the 3D-GPCRome. Nat Methods. 2020;17:777–787.
76. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter. Phys Rev Lett. 1996;77:1905–1908.
77. Haliloglu T, Bahar I, Erman B. Gaussian dynamics of folded proteins. Phys Rev Lett. 1997;79:3090–3093.
78. Sinitskiy AV, Voth GA. Coarse-graining of proteins based on elastic network models. Chem Phys. 2013;422:165–174. https://doi.org/10.1016/j.chemphys.2013.01.024
79. Madsen JJ, Sinitskiy AV, Li J, Voth GA. Highly coarse-grained representations of transmembrane proteins. J Chem Theory Comput. 2017;13:935–944.
80. Amadei A, Linssen ABM, Berendsen HJC. Essential dynamics of proteins. Proteins. 1993;17:412–425.
81. Martín-García F, Papaleo E, Gomez-Puertas P, Boomsma W, Lindorff-Larsen K. Comparing molecular dynamics force fields in the essential subspace. PLoS One. 2015;10:e0121114. https://doi.org/10.1371/journal.pone.0121114
82. Csermely P. Plasticity-rigidity cycles: a general adaptation mechanism. arXiv:1511.01239v3 [q-bio.MN]
83. Madsen JJ, Olsen OH. Conformational plasticity-rigidity axis of the coagulation factor VII zymogen elucidated by atomistic simulations of the N-terminally truncated factor VIIa protease domain. Biomolecules. 2021;11:549.
84. Kooistra AJ, Mordalski S, Pándy-Szekeres G, et al. GPCRdb in 2021: integrating GPCR sequence, structure and function. Nucleic Acids Res. 2021;49:D335–D343.
85. Crozier PS, Stevens MJ, Woolf TB. How a small change in retinal leads to G-protein activation: initial events suggested by molecular dynamics calculations. Proteins Struct Funct Genet. 2007;66:559–574.

86. Vanni S, Neri M, Tavernelli I, Rothlisberger U. A conserved protonation-induced switch can trigger "ionic-lock" formation in adrenergic receptors. J Mol Biol. 2010;397:1339–1349.
87. Dror RO, Arlow DH, Maragakis P, et al. Activation mechanism of the β 2-adrenergic receptor. Proc Natl Acad Sci U S A. 2011;108:18684–18689.
88. Goetz A, Lanig H, Gmeiner P, Clark T. Molecular dynamics simulations of the effect of the G-protein and diffusible ligands on the β2-adrenergic receptor. J Mol Biol. 2011;414:611–623.
89. Li J, Jonsson AL, Beuming T, Shelley JC, Voth GA. Ligand-dependent activation and deactivation of the human adenosine A_{2A} receptor. J Am Chem Soc. 2013;135:8749–8759.
90. Prosser RS, Ye L, Pandey A, Orazietti A. Activation processes in ligand-activated G protein-coupled receptors: a case study of the adenosine A_{2A} receptor. BioEssays. 2017;39:1–10.
91. Alhadeff R, Vorobyov I, Yoon HW, Warshel A. Exploring the free-energy landscape of GPCR activation. Proc Natl Acad Sci U S A. 2018;115:10327–10332.
92. Ciancetta A, Gill AK, Ding T, et al. Probe confined dynamic mapping for G protein-coupled receptor allosteric site prediction. ACS Cent Sci. 2021;7:1847–1862.
93. Lee JY, Lyman E. Agonist dynamics and conformational selection during microsecond simulations of the A_{2A} adenosine receptor. Biophys J. 2012;102:2114–2120. https://doi.org/10.1016/j.bpj.2012.03.061
94. Vogt G, Argos P. Protein thermal stability: hydrogen bonds or internal packing? Fold Des. 1997;2:40–S46.
95. Joosten RP, Long F, Murshudov GN, Perrakis A. The PDB_REDO server for macromolecular structure model optimization. IUCrJ. 2014;1:213–220.
96. Afonine PV, Grosse-Kunstleve RW, Echols N, et al. Towards automated crystallographic structure refinement with phenix. refine. Acta Crystallogr D Biol Crystallogr. 2012;68:352–367.
97. Adams PD, Afonine PV, Bunkóczi G, et al. PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66:213–221.
98. Staus DP, Strachan RT, Manglik A, et al. Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. Nature. 2016;535:448–452.
99. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. J Comput Chem. 2011;32:2319–2327.
100. Gowers R, Linke M, Barnoud J, Reddy T, Melo M, Seyler S, Domański J, Dotson D, Buchoux S, Kenney I, et al. (2016) MDAnalysis: a python package for the rapid analysis of molecular dynamics simulations. Proc 15th Python Sci Conf. 98–105.
101. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput. 2013;9:3084–3095.
102. Pedregosa F, Varoquax G, Gramfort A, et al. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12:2825–2830.
103. Olsen OH (2021) Ensemble refinement pub files https://doi.org/10.5281/zenodo.4896421.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Madsen JJ, Ye L, Frimurer TM, Olsen OH. Mechanistic basis of GPCR activation explored by ensemble refinement of crystallographic structures. Protein Science. 2022;31(11):e4456. https://doi.org/10.1002/pro.4456