



The mechanism for ligand activation of the GPCR–G protein complex

Amirhossein Mafi^a, Soo-Kyung Kim^a, and William A. Goddard III^{a,1}

Contributed by William A. Goddard III; received May 31, 2021; accepted March 24, 2022; reviewed by Marco De Vivo and Kenneth A. Jacobson

G protein–coupled receptors (GPCRs) activate cellular responses ranging from odorants to neurotransmitters. Binding an agonist leads to activation of a heterotrimeric G protein (GP) that stimulates external signaling. Unfortunately, the mechanism remains unknown. We show for 15 class A GPCRs, including opioids, adrenergics, adenosines, chemokines, muscarinics, cannabinoids, serotoninins, and dopamines, that interaction of an inactive GP, including Gs, Gi, Go, G11, and Gq, to the inactive GPCR, containing the intracellular ionic lock between transmembrane (TM) helices 3 and 6, evolves exothermically to form a precoupled GPCR–GP complex with an opened TM3–TM6 and the GP- $\alpha 5$ helix partially inserted into the GPCR but not activated. We show that binding of agonist to this precoupled GPCR–GP complex causes the G α protein to open into its active form, with the guanosine diphosphate exposed for signaling. This GP–first paradigm provides a strategy for developing selective agonists for GPCRs since it is the pharmacophore for the precoupled GPCR–GP complex that should be used to design drugs.

G protein activation | molecular metadynamics | biased agonists | opioids | adrenergic

G protein–coupled receptors (GPCRs) constitute the largest group of membrane receptors in eukaryotes. They are responsible for activating cellular responses to numerous bioactive molecules, including odorants, pheromones, hormones, and neurotransmitters (1). Binding these external signaling molecules activates a GPCR through two pathways: one involving heterotrimeric G proteins (GPs) and the other involving arrestins, often with quite different consequences. Since $\sim 34\%$ of all modern medicinal drugs (2, 3) act on one or more of the 800 human GPCRs (4), there is great interest in understanding how binding of a ligand elicits this signaling. The aim of this paper is to elucidate this mechanism.

Generally, it has been assumed that binding agonists to the inactive conformation of a GPCR shifts the equilibrium toward the activated conformation (5, 6), allowing the liganded GPCR to recruit the inactive GP with bound guanosine diphosphate (GDP) and then activate the GP to mediate specific cellular signaling (7). We refer to this paradigm as the ligand–first mechanism of GP activation. There is support for this mechanism from experiments that excluded GP to allow agonist to bind first, leading to subsequent full activation after adding GP. On the other hand, many experimental observations (8–13) and computational studies (10, 11, 14–16) have revealed that agonists alone often do not stabilize the active conformation of the GPCR, hindering the subsequent recruitment of GP. Thus, complex formation between class A GPCRs and their cognate GPs greatly relies on random collisions between the pair (17, 18). Given that different types of GPs exist in a cell, there is tight competition between different subtypes of GPs to possibly couple with a given liganded GPCR, which makes a complex formation between liganded GPCR and GP relatively slow. However, the cellular response through the activation of GPs was shown to be rather rapid (7, 19, 20). In addition, the intrinsic basal activity of GPCRs, which leads to constitutive activation of GPs, is another phenomenon that the ligand–first mechanism of GP activation cannot describe. Many of the GPCRs that display this intrinsic basal activity in the absence of a ligand (the apo state) are required for normal physiological functions (21). For example, serotonin receptors (5-HT_{2A} and 5-HT_{2c}) exhibit a high level of constitutive activity (22). Suppressing the constitutive activity of 5-HT_{2A} and 5-HT_{2c} triggers mechanisms underlying depression and anxiety (22).

In this paper, we propose an alternate paradigm (Fig. 1) in which prior to ligand binding, the inactive GP interacts with the inactive GPCR to open the intracellular region by breaking the transmembrane (TM) 3–TM6 tight coupling (an interaction from R^{3.50}, part of the DRY motif, to a conserved residue at the cytosolic end of TM6) to form a stable precoupled complex. This precoupled complex remains at this resting state until an agonist binds to the GPCR–GP complex to open the tightly

Significance

We report the detailed atomistic mechanism for how molecules such as morphine, dopamine, or epinephrine binding outside of a cell to a G protein–coupled receptor (GPCR) in the cell membrane cause a G protein (GP) bound at the inside of the cell to break apart and signal the cell to influence appetite, anxiety, memory, cognition, learning, and sleep. Most surprising is that the GP binds first to the GPCR to form a precoupled complex that remains at rest until the drug binds to induce the signaling. Most important, it is the precoupled GPCR–GP structure that provides the basis for the design of therapeutics to maximize activity and selectivity.

Author affiliations: ^aMaterials and Process Simulation Center, California Institute of Technology, Pasadena, CA 91125

Author contributions: A.M. and W.A.G. designed research; A.M. performed research; A.M., S.-K.K., and W.A.G. analyzed data; and A.M., S.-K.K., and W.A.G. wrote the paper.

Reviewers: M.D.V., Istituto Italiano di Tecnologia; and K.A.J., NIH.

The authors declare no competing interest.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: wag@caltech.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2110085119/-/DCSupplemental>.

Published April 22, 2022.

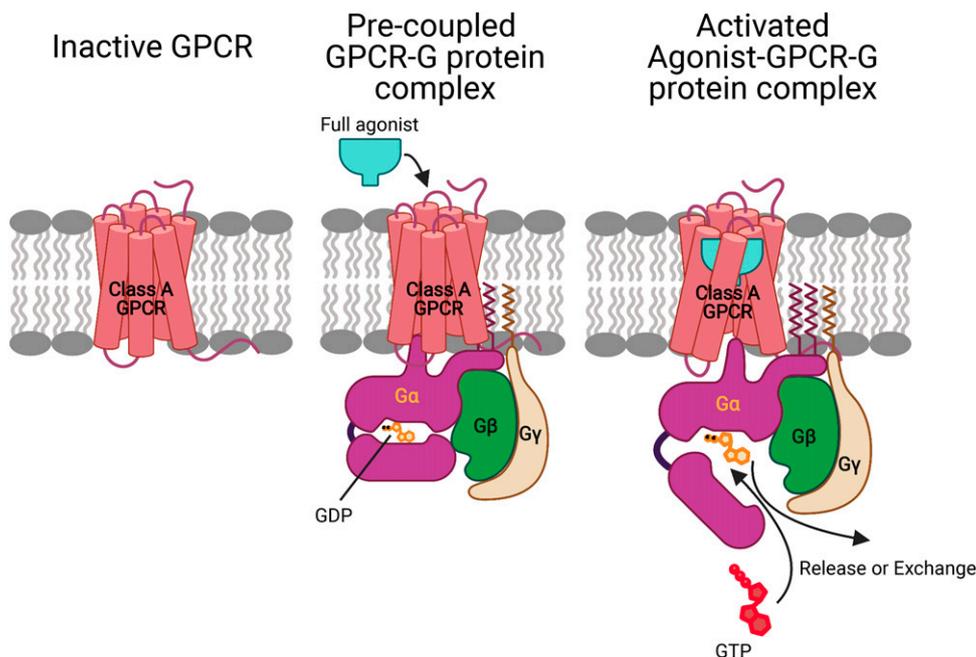


Fig. 1. GP-first mechanism of GP activation. Prior to ligand binding, the inactive GP interacts with the inactive GPCR to open the intracellular region by breaking the TM3-TM6 tight link to form a stable precoupled complex. This precoupled complex remains at this resting state until an agonist binds to the GPCR-GP complex to open the intracellular region of GPCR and the tightly coupled $G\alpha$ -GDP complex to form the fully activated agonist-GPCR-GP complex with the GDP available for exchange or release.

coupled $G\alpha$ -GDP complex while further opening the intracellular region of the GPCR to form the fully activated agonist-GPCR-GP complex with the GDP available for exchange. We refer to this as the GP-first mechanism of GP activation.

Previous observations have shown that a number of GPCRs form a precoupled complex with their cognate GPs (23–28), but the detailed molecular mechanism of subsequent activation by an agonist is not understood. In this study, to pursue the GP-first mechanism of GP activation, we examined coupling of 15 class A GPCRs to a total of six different GPs and show that in all 15 cases, the inactive GP couples with the inactive GPCR to form a precoupled-GPCR-GP complex in which the intracellular TM3-TM6 tight link is broken and the GP- $\alpha 5$ helix is partially inserted into the GPCR. We refer to this as the precoupled GPCR-GP complex. The combinations examined are as follows:

- 5-HT_{2A}-serotonin receptor–Gq protein
- A_{2A} adenosine receptor–Gs protein
- β_2 -adrenergic receptor–Gs protein
- μ -opioid receptor–Gi1 protein
- κ -opioid receptor–Gi1 protein
- δ -opioid receptor–Gi1 protein
- CCR5-chemokine receptor–Gi1 protein
- CB1-cannabinoid receptor–Gi1 protein
- A₁-adenosine receptor–Gi2 protein
- 5-HT_{1B}-serotonin receptor–Go protein
- D₂-dopamine receptor–Go protein
- M₁-muscarinic receptor–G11 protein
- M₃-muscarinic receptor–Gq protein
- α_{2A} -adrenergic receptor–Gq protein
- 5-HT_{2C}-serotonin receptor–Gq protein

The predicted precoupled complexes are discussed in detail below.

Moreover, for 5-HT_{2A}-serotonin receptor–Gq protein, A_{2A} adenosine receptor–Gs protein, and μ -opioid receptor–Gi1 protein, we predicted how binding a full agonist to the precoupled

tightly coupled $G\alpha$ -GDP complex causes the $G\alpha$ subunit to open and becomes activated with the GDP available for exchange.

In this paper, we investigate the GP-first mechanism of GPCR activation using long-scale molecular dynamics simulations (an aggregate of $\sim 20 \mu\text{s}$) and metadynamics (metaD) simulations to follow the sequence of structural and energetic steps involved in activation of class A GPCRs and their cognate GP. The activation process goes through several metastable states in which GPCR and GP undergo several structural changes on the sequence toward activation. Some metastable states may be separated by high energy barriers that could take microseconds or longer. Thus, we used metaD simulations (29) incorporating collective variables to describe the slow degrees of freedom but biased to encourage each GPCR and its cognate GP to explore large regions of conformational phase space to track the activation in much reduced time.

Results

We expect that prior to ligand binding, inactive GPCRs and their cognate inactive GPs have sufficient time to interact and form a precoupled complex. Thus, we considered the interactions between class A inactive apo-GPCRs (which generally have tight cytosolic TM3-TM6 coupling, such as an ionic lock) with their cognate inactive GP-bound GDP complex. The cryogenic electron microscopy (cryo-EM) structure of apo-5-HT_{1A}-Gi protein (30) indicates that GPs couple to unliganded GPCR in a similar fashion and geometry as they do with liganded GPCRs in the fully active state. Thus, we placed the inactive GPs close enough (31) to the inactive apo-GPCRs (*SI Appendix, Methods*), adopting a similar geometry and orientation revealed by cryo-EM and X-ray crystallography for the fully activated agonist-GPCR-GP complex, such that GPCRs and GPs could start interacting (Fig. 2*A* and *SI Appendix, Figs. S2–S4*). The inactive state of the GP does not have the $G\alpha 5$ helix fully extended, so there is no steric clash between GP and

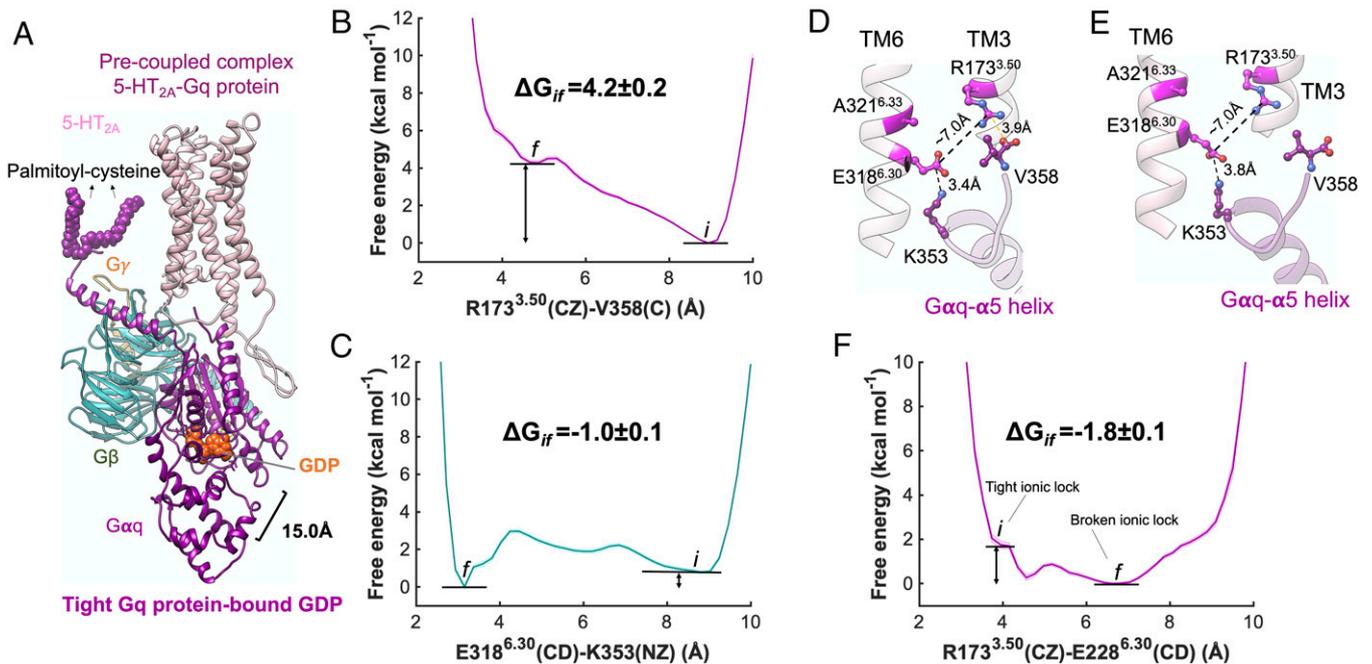


Fig. 2. The precoupled 5-HT_{2A}-Gq protein complex. (A) Our energetically optimized precoupled complex of 5-HT_{2A}-Gq protein-GDP. MetaD free energy of (B) R173^{3.50}(CZ)-V358(C) and (C) E318^{6.30}(CD)-K353(NZ). (D) Extensive engagement between the Gαq-α5 helix to the cytosolic end of TM3, and TM6 of 5-HT_{2A}, along the Gq precoupling that breaks open the ionic lock between TM3 and TM6. (E) Detailed atomic interactions between 5-HT_{2A} and Gαq-α5 helix in the precoupled state. (F) MetaD free energy of R173^{3.50}(CZ)-E318^{6.30}(CD) upon the formation of precoupling complex between 5-HT_{2A} and Gq protein, which was estimated by performing an independent ~600 ns metaD simulation. The metaD free energies were reweighted (51) for estimation of the free energy errors using the block averaging method.

GPCRs. Finally, we compare this optimized geometry between GP and GPCRs after activation directly to the structures resolved by experiments.

Gq Protein Precoupling to the Inactive 5-HT_{2A} Receptor. We used the serotonin 5-HT_{2A} receptor as the prototype for our proposed GP-first activation pathway, and then we examined a similar molecular mechanism of activation for 14 other class A GPCRs, including the A_{2A} adenosine receptor and β₂ adrenergic receptor, the archetypes for class A GPCRs.

The inactive conformation of 5-HT_{2A} receptor has a salt bridge R173^{3.50}-E318^{6.30} [the superscript is Ballesteros-Weinstein numbering for GPCRs (32), taken from Pándy-Szekeres et al. (33)], the ionic lock, that inhibits activation. Breaking and disruption of the ionic lock is believed to be an important step toward the activation of class A GPCRs (34–36). We hypothesize that during the formation of precoupled state between 5-HT_{2A} and Gq protein, the α5 helix of Gq protein partially penetrates to the intracellular region of 5-HT_{2A} and makes a salt bridge with R173^{3.50} and E318^{6.30}, opening the ionic lock. The rearrangements in ionic interactions between the cytoplasmic end of 5-HT_{2A} receptor and the α5 helix of Gq protein facilitates opening the ionic lock. To examine this hypothesis, we performed two independent metaD simulations for an aggregated ~1 μs, in which we evaluated the energetics of forming two salt bridges: V358 (terminal CO₂⁻)-R173^{3.50}(CZ) and K353(NZ)-E318^{6.30}(CD) and their consequences on the ionic lock between R173^{3.50}(CZ)-E318^{6.30}(CD).

We find that during formation of the precoupled complex (Fig. 2A), as the GP approaches the GPCR, the terminal CO₂⁻ at the end of the Gα5 helix forms a salt bridge with R173^{3.50} (Fig. 2D), initiating breaking of the ionic lock. However, our ~400-ns metaD simulation (Fig. 2B and C) reveals that this salt bridge V358-R173^{3.50} is endothermic by 4.2 ± 0.2 kcal/mol (Fig. 2B). Thus, after opening the ionic lock between TM3-

TM6, the V358 terminal carboxylate disengages from R173^{3.50} and penetrates deeper into the core of the receptor (Fig. 2E). Interestingly, K353 on the Gα5 helix forms a persistent salt bridge with E318^{6.30} (Fig. 2D and E) with a binding affinity of -1.0 ± 0.1 kcal/mol (Fig. 2C) that contributes substantially to breaking the R173^{3.50}-E318^{6.30} ionic lock while stabilizing the position of Gα5 inside the intracellular region of 5-HT_{2A}.

To determine whether disruption of the ionic lock R173^{3.50}-E318^{6.30} in the precoupled 5-HT_{2A}-Gq protein complex is statistically significant, we carried out an independent free energy metaD simulation for ~600 ns (Fig. 2F). In this study, we separately evaluated the energetics of opening the ionic lock, finding that precoupling of Gq protein to 5-HT_{2A} breaks open the ionic lock to ~7 Å while reducing the energy by -1.8 ± 0.1 kcal/mol, a thermodynamically favorable process. Indeed, this disruption with breaking of TM3-TM6 coupling is well known to be a critical step in activation of class A GPCRs (34–36). We find that a persistent charge-charge or salt bridge interaction K353-E318^{6.30} emerges in the precoupled complex concomitant with the Gα5 terminal carboxylate penetrating into the core of the GPCR after the ionic lock is broken (Fig. 3O), a result similar to our first simulation.

There remains a possibility that the rigid-body orientation of Gq protein could be different in the precoupled state from that in the fully active complex. To eliminate the possibility that the specific rigid-body orientation of Gq protein is solely responsible for opening the TM3-TM6 coupling, we carried out an independent third metaD free energy calculation for ~1.5 μs in which only the Gαq-α5 peptide (the last 26 residues: ³³³T-V³⁵⁸) is placed in close proximity (K353 and V358 10 Å away from R173^{3.50} and E318^{6.30}, respectively) to the inactive 5-HT_{2A} (SI Appendix, Fig. S1). The increased degrees of freedom for the Gαq-α5 peptide enabled the metaD to explore numerous positions and orientations that would emerge from various orientations of the whole Gq protein in complex with

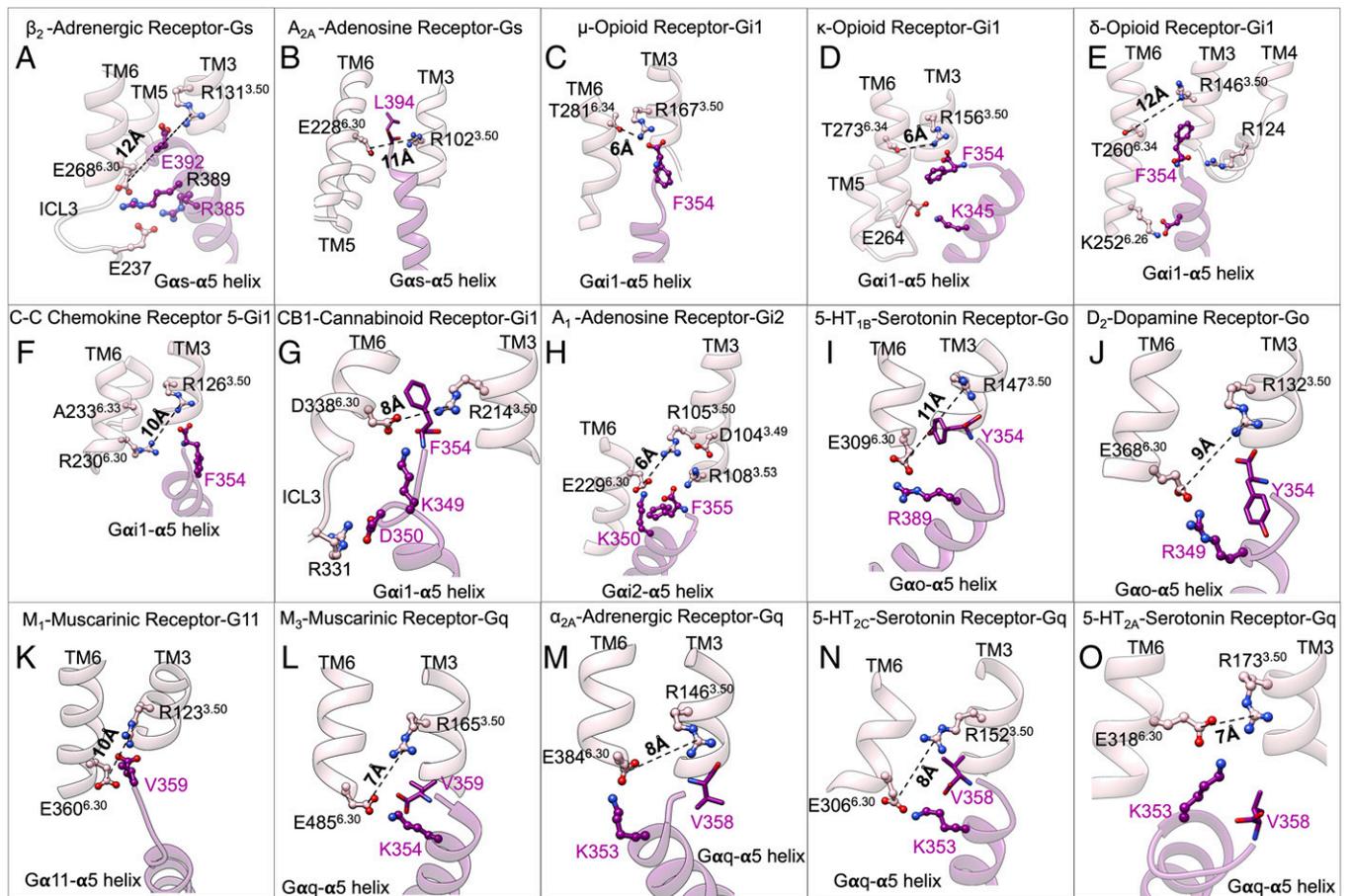


Fig. 3. Precoupled complexes of class A GPCR-GPs. Detailed atomic interactions between G α 5 helix and class A GPCRs in the precoupled state resulted from extensive metaD simulations for (A) β_2 -adrenergic receptor-Gs, (B) A $_{2A}$ adenosine-Gs, (C) μ -opioid-Gi1, (D) κ -opioid-Gi1, (E) δ -opioid-Gi1, (F) CCR5-chemokine-Gi1, (G) CB1-cannabinoid-Gi1, (H) A $_1$ -adenosine-Gi2 protein, (I) 5-HT $_{1B}$ -serotonin-Go, (J) D $_2$ -dopamine-Go, (K) M $_1$ -muscarinic-G11, (L) mouse M $_3$ -muscarinic-mouse Gq, (M) α_{2A} -adrenergic-Gq, (N) 5-HT $_{2C}$ -serotonin-Gq, and (O) 5-HT $_{2A}$ -serotonin-Gq. The details of the calculations are represented in *SI Appendix, Fig. S4 and Table S3*. (C–E) Adapted from figure 6A of ref. 14.

the 5-HT $_{2A}$. We find that prior to ligand binding, a salt bridge contact from the terminal carboxylate, V358, to R173 $^{3.50}$ (*SI Appendix, Fig. S1 A–C*), contributes to opening the 5-HT $_{2A}$ ionic lock to 7 Å, consistent with the precoupled state we found in the presence of the whole Gq protein (Figs. 2E and 3O). These calculations confirm that formation of the precoupled 5-HT $_{2A}$ -Gq protein complex is not an artifact resulting from a specific rigid-body orientation of the Gq protein.

Generalization of the GP-First Mechanism for Activation of Class A GPCRs. We expect that inactive GPs generally have sufficient time to couple to inactive GPCRs prior to drug binding and that formation of a precoupled state apply generally to all class A GPCRs and their cognate GP. In our proposed GP-first activation paradigm:

1) The apo-GPCR initially exhibits a tight cytoplasmic region due to the interaction of R $^{3.50}$ (part of the DRY motif) with TM6. This TM3-TM6 coupling for the inactive conformation of class A GPCRs can be either an ionic lock (Fig. 3 A, B, and G–O), a hydrogen bond (Fig. 3 C–E), or a hydrophobic interaction (Fig. 3 F). Thus, the TM3-TM6 coupling constitutes major slow degrees of freedom along the activation path that need to be disturbed to accommodate the G α 5 helix for the emergence of the precoupled state and later activation (34–36).

2) During the formation of the precoupled state between class A GPCRs and their cognate GPs, the α 5 helix of GPs partially penetrates the intracellular region of GPCRs and makes a salt bridge with R $^{3.50}$ (part of the DRY motif), which breaks the tight TM3-TM6 coupling, opening up the cytoplasmic region of GPCRs.

To further validate our GP-first paradigm, we examined whether Gs protein alone can open up the TM3-TM6 coupling of the apo- β_2 adrenergic receptor from its inactive conformation. We tested the β_2 adrenergic receptor because it is one of the best-characterized class A GPCRs. Although the X-ray crystallographic study of the inactive β_2 adrenergic receptor found a hydrophobic coupling between R131 $^{3.50}$ -L272 $^{6.34}$, a previous computational study (37) revealed that the inactive conformation of the β_2 adrenergic receptor with its native intracellular loop 3 forms an ionic lock between R131 $^{3.50}$ -E268 $^{6.30}$. To optimize the inactive state of the β_2 adrenergic receptor, we first inserted the native intracellular loop 3 into the crystallographic inactive state and then performed a \sim 2.2- μ s metaD simulation to find that R131 $^{3.50}$ makes an ionic lock with E268 $^{6.30}$. Subsequently, to examine if partial insertion of the G α 5 helix into the core of the β_2 adrenergic receptor perturbs the ionic lock, we performed a \sim 800-ns metaD simulation in which we evaluated the energetics of salt bridges involving the G α 5 helix: R389 to E268 $^{6.30}$ and E392 to R131 $^{3.50}$ (*SI Appendix, Fig. S2*). Our free energy calculations show that E392 forms a salt

bridge with R131^{3,50} (Fig. 3A) with a high binding affinity of $\sim -10.2 \pm 0.2$ kcal/mol. Simultaneously, E268^{6,30} (the partner in the ionic lock) establishes a high-affinity ($\sim -2.8 \pm 0.3$ kcal/mol) salt bridge with R389. As a result, the TM3-TM6 ionic lock opens fully to ~ 12 Å (Fig. 3A), leading to the precoupled state. The significant roles of R389 and E392 in the precoupled complex agree with a recent mutagenesis (38) study indicating that R389 and E392 are essential for efficient formation of a complex between Gs protein and β_2 adrenergic receptor. In fact, mutation of E392 to an Ala residue perturbed the initiation of GDP release (38). Additionally, a series of E392^{G_{as}} mutants to Ala, Arg, Gln, Val, Leu, and Ser exhibited impaired cyclic adenosine monophosphate accumulation (39), confirming that E392 serves a crucial role in the activation of Gs protein.

To examine the GP-first mechanism of activation for other class A GPCRs, we studied precoupling of 13 additional GPCRs to Gs (Fig. 3B), Gi/o (Fig. 3C–F), or Gq/11 (Fig. 3G–O). We also took all sorts of TM3-TM6 couplings—an ionic lock (Fig. 3B and G–O), a hydrogen bond (Fig. 3C–E), or a hydrophobic interaction (Fig. 3F)—into consideration. All sorts of TM3-TM6 couplings must be opened to accommodate the G α 5 helix for emergence of the precoupled state. Thus, we performed ~ 7.2 - μ s metaD simulations, with the main idea that the salt bridge interactions from the G α 5 helix to counterparts of the TM3-TM6 couplings, particularly R^{3,50} (part of the DRY motif), open up the cytoplasmic regions. To eliminate the probable impacts of the chosen collective variable from our results, we used various combinations of collective variables (*SI Appendix*, Figs. S3 and S4 and Table S3). For these calculations, we used either of two well-validated force fields [CHARMM36m (40) and AMBER14 (41); *SI Appendix*, Tables S2–S4] to eliminate the possibility that formation of the precoupled GPCR-GP complex results solely from the choice of a specific force field. Indeed, our free energy calculations find for 13 cases that the salt bridge interactions between the G α 5 helix and the intracellular region of the GPCR (particularly to the conserved R^{3,50}) break open the tight TM3-TM6 coupling to accommodate the G α 5 helix partially inside the core of receptors, leading to emergence of the precoupled GPCR-GP complex. Fig. 3 depicts the molecular interactions in the GPCR-GP precoupled complexes for all 15 cases. The GPCRs studied include the A_{2A} adenosine receptor (42, 43), D2 dopamine receptor (25, 44), α_{2A} adrenergic receptor (23, 25), M3 muscarinic receptor (26), and A1 adenosine receptor (25), which were previously shown to make a precoupled complex with their cognate GP.

Agonist Activation of GP. To determine the role of agonist in the GP-first activation paradigm, we inserted a full agonist, 25CN-NBOH, into the precoupled complex of 5-HT_{2A}-Gq protein such that a salt bridge from the conserved D155^{3,32} to the protonated N atom of the agonist locks the ligand into the orthosteric binding pocket of 5-HT_{2A} (*SI Appendix*, Fig. S5). Given that the cellular signaling through the GP activation arises from the exchange of a GDP for a guanosine triphosphate (GTP), we assessed the energetics of opening the G α q subunit from the cleft between α -helical (AH) (the center of mass of C α for the residues 154 to 161 and 175 to 182) and Ras-like (the center of mass of C α for the residues 51 to 62) domains, which defines the nucleotides' (GDP and GTP) binding pocket. Opening the tight G α q makes the GDP release or exchange facile (45, 46). Our free energy calculations show that the G α q subunit subsequently undergoes a remarkable opening, increasing the separation between AH and Ras-like domains from ~ 16 Å (tight conformation) to ~ 23 Å (open conformation) while opening the GDP binding site (Fig. 4A and B), a remarkable structural

rearrangement induced by ligand binding. This process is energetically favorable (-7.5 ± 0.6 kcal/mol; Fig. 4A) and leaves the GDP exposed to water, making it susceptible to dissociation or GTP exchange. This opening of G α q expedites GDP release, a critical event in activation of GP and GP signaling (45, 46). Our metaD simulation finds that although opening the G α q protein provides an exit path for GDP dissociation, the GDP remains bound to the Ras-like domain in our simulations, consistent with previous experimental (47, 48) and computational (49, 50) studies showing that GDP still remained only bound to the Ras-like domain even when the G α subunit opens up.

To examine whether the opening of GP from the GDP binding site prevails only in the presence of agonist, we also estimated the energetics for opening of the G α q subunit for

- a partial agonist, lysergic acid diethylamide (LSD);
- an inverse agonist, methiothepin; and
- the case with no ligand present in the orthosteric pocket of 5-HT_{2A} (*SI Appendix*, Fig. S5).

Our free energy calculations reveal that the presence of the partial agonist also induces the G α q subunit to open up from 16 to 23 Å (Fig. 4C and D) but in a less favorable process ($\Delta G = -3.8 \pm 0.6$ kcal/mol) compared to the full agonist binding ($\Delta G = -7.5 \pm 0.6$ kcal/mol). To compare the relative energy barrier of activation between these types of ligands in the same activation pathway, we reweighted (51) the metaD free energies to estimate the error. We find that the free energy barrier associated with activation for a full agonist ($\Delta G^\ddagger = +3.3 \pm 0.6$ kcal/mol) is far more favorable than for a partial agonist ($\Delta G^\ddagger = +6.8 \pm 0.5$ kcal/mol) (Fig. 4B and D).

On the other hand, in the absence of an agonist the free energy to activate the apo-5-HT_{2A}-Gq protein precoupled complex is endothermic by $\Delta G = +4.6 \pm 0.6$ kcal/mol (Fig. 4E and F), consistent with a previous study (44) showing a low level of constitutive activity for 5-HT_{2A} even with a broken ionic lock. Indeed, we find that insertion of an inverse agonist selectively disfavors the opening of Gq protein (Fig. 4G and H) even further, increasing the free energy by $\Delta G = +9.7 \pm 0.6$ kcal/mol (Fig. 4J) and producing a response opposite to the full agonist, which is consistent with the physiological role of an inverse agonist. Comparing the energy barrier of activation (Fig. 4F and H) for the apo-5-HT_{2A} ($\Delta G^\ddagger = +9.8 \pm 0.6$ kcal/mol) and inverse agonist (methiothepin-5-HT_{2A}, $\Delta G^\ddagger = +12.1 \pm 0.6$ kcal/mol) suggests that the inverse agonist suppresses constitutive activation of 5-HT_{2A} and Gq protein.

Opening the G α q subunit in the presence of a full agonist has a dramatic effect on the position of the G α 5 helix and consequently on the position of TM6. We find that in the fully activated 25CN-NBOH-5-HT_{2A}-Gq complex, the G α 5 helix undergoes a pronounced ~ 8 -Å upward movement along its axis into the receptor core (Fig. 5A and B), allowing the G α 5 helix to rotate 50° around its axis relative to the G α 5 helix of the inactive conformation (Fig. 5A) and leading to extensive interactions with the cytoplasmic region of 5-HT_{2A} receptor. This translation along and rotation of the G α 5 helix around its axis are known hallmarks of GP activation (49, 52, 53), playing a key role in nucleotide release. Indeed, this outward movement of the G α 5 helix is associated with ~ 9 -Å outward displacement of the cytosolic end of TM6 (Fig. 5C) to match closely the fully active state of 5-HT_{2A} resolved in the cryo-EM structure (54) with root mean-square displacement (rmsd) of 2.4 Å (Fig. 5D).

To independently determine if the presence of full agonists can activate the precoupled GPCR-GP complex for other GPCRs, we inserted

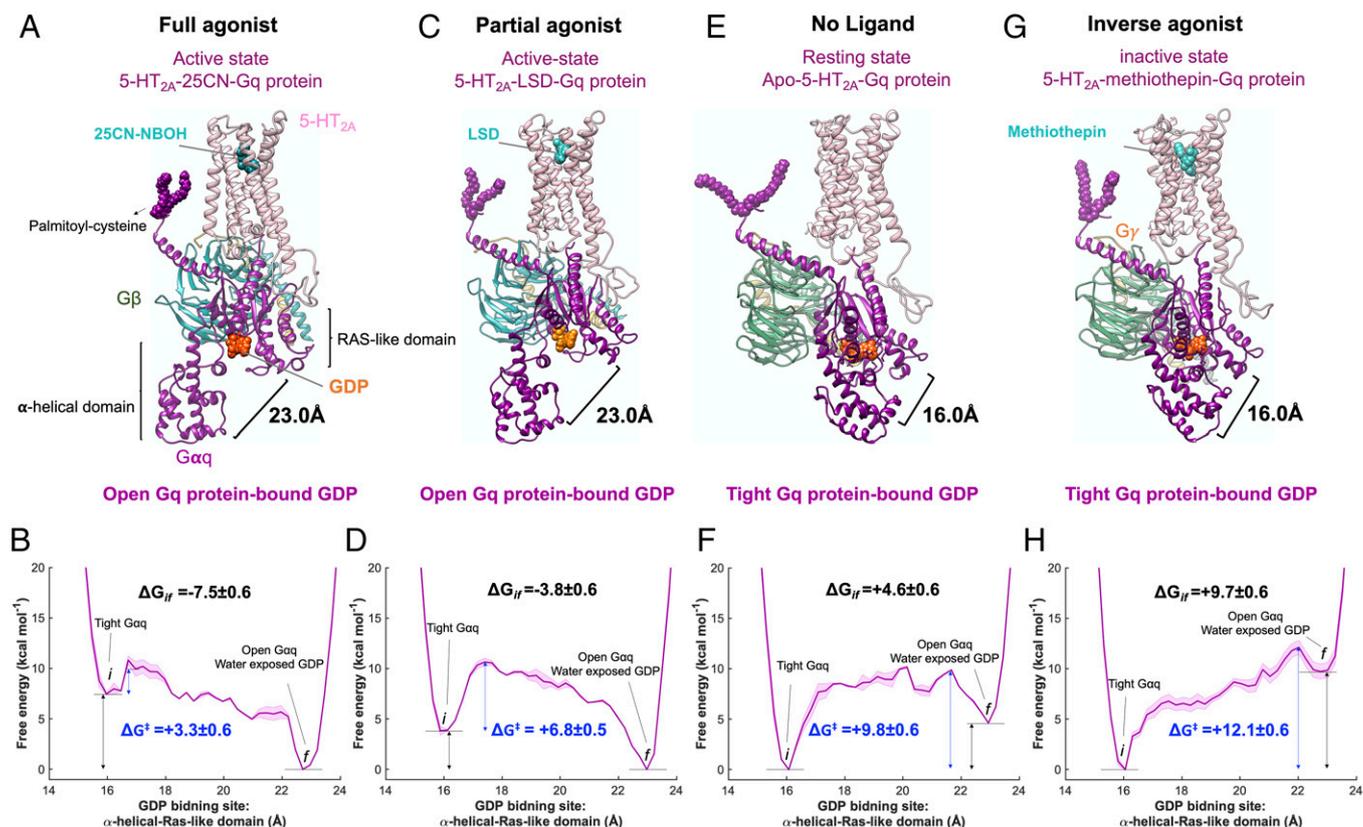


Fig. 4. Ligand activation of precoupled 5-HT_{2A}-Gq protein complex. Gq protein activation mediated by (A) 25CN-NBOH, a full agonist, and (C) LSD, a partial agonist, upon binding to the precoupled 5-HT_{2A}-Gq protein complex. Gq protein inhibited activation caused by (E) absence of an agonist (apo-5-HT_{2A}) and (G) methiothepin, an inverse agonist upon binding to the precoupled 5-HT_{2A}-Gq protein complex. (B, D, F, and H) MetaD free energy of Gαq-bound GDP opening from its GDP binding site. Here, the distance between the AH domain (the center of mass of Cα for the residues 154 to 161 and 175 to 182) and the Ras-like domain (the center of mass of Cα for the residues 51 to 62), was considered for the free energy calculations. The weighted averages and the SDs were calculated for ΔG_{if} for the converged period. The metaD free energies were reweighted (51) for estimation of the error, particularly energy barrier ΔG^\ddagger presented in B, D, F, and H, using the block averaging method.

- the full agonist 5'-N-ethylcarboxamide adenosine (NECA) into the A_{2A} adenosine receptor–Gs protein complex (Fig. 6A), and
- the full agonist morphine into the μ-opioid receptor–Gi protein complex (Fig. 6D and *SI Appendix*, Fig. S3).

Our free energy calculations show that the full agonists

- open the distance between the Gαs subunits for the A_{2A} adenosine receptor by ~10 Å in the cleft between AH and RAS-like domains (Fig. 6A) and
- open the distance between the Gαi subunits for the μ-opioid receptor by ~8 Å in the cleft between AH and RAS-like domains (Fig. 6C).

These structural rearrangements make GDP exchange and subsequent cellular signaling facile, confirming that agonists convert the precoupled to the fully active state in a thermodynamically favorable pathway.

An important implication of the GP-first mechanism of GP activation is that it is the structure the precoupled GPCR-GP complex that should be used to provide the pharmacophore for structure-based drug design of agonists to maximize activity and selectivity. Indeed, most important may be the structure at the transition state for opening the Gα, which probably dominates activity. Drugs must aim at quickly activating the precoupled complex of the target GPCR-GP complex while blocking activity of the precoupled complex for all other GPCR-GP complexes to reduce undesirable side effects so common in GPCR-targeted therapies (the target-antitarget strategy).

Discussion

We expect that prior to ligand binding, inactive GPCRs and their cognate inactive GP have sufficient time to associate, producing a precoupled complex. This coupling disrupts the intrinsic tight coupling of the cytoplasmic ends of TM3-TM6 that keeps the class A GPCRs inactive. Although this disruption of TM3-TM6 coupling is essential to activation, it need not necessarily result in a remarkable outward displacement of TM6 from TM3, the well-known structural rearrangement associated with activation of class A GPCRs (53). This precoupled GPCR-GP complex remains at rest until the agonist arrives to drive the precoupled complex to its final activated state, during which TM6 experiences the large outward movement necessary to fully accommodate the Gα5 helix. Concomitantly, ligand binding opens the Gα protein to expose the GDP and allows Gα5 to move outward while rotating about its axis to interact extensively with the intracellular GPCRs region.

On the other hand, NMR spectroscopy (8, 11, 12), double electron-electron resonance spectroscopy (12), crystallography (10), and computational (10, 11, 16) studies indicate that full agonists (high-efficacy ligands) alone do not stabilize the active conformation of GPCRs. Rather, they significantly increase the basal activity of GPCRs (imposing outward TM6 displacement), which was shown to be just sufficient for the recruitment and later activation of the GP (9). Indeed, our analysis shows that agonists alone cannot stabilize the active state conformation since they cannot break the tight TM3-TM6 coupling (14) in the inactive apo-GPCR (*SI Appendix*, Fig. S8).

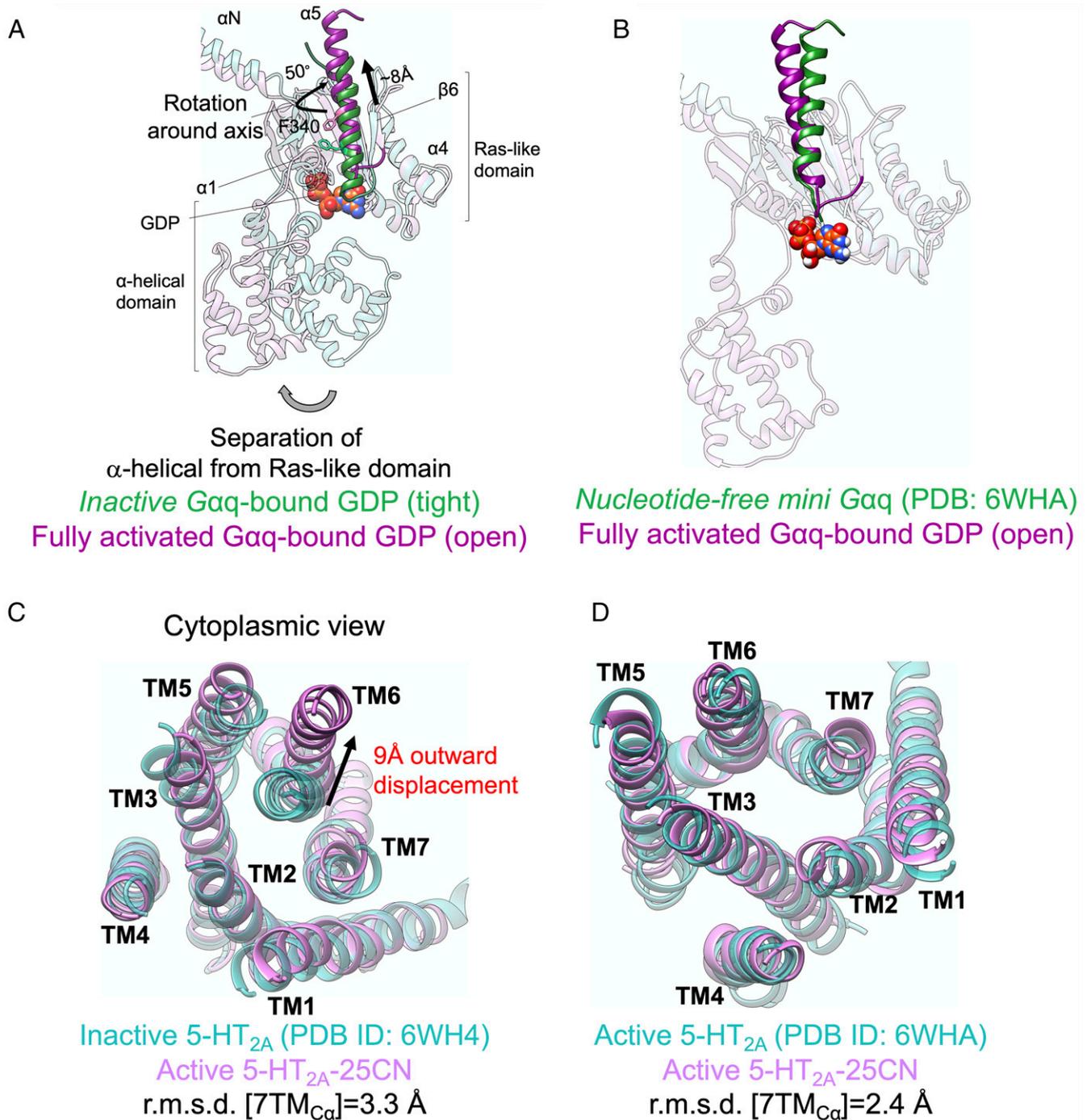


Fig. 5. Fully active state of 5-HT_{2A}-Gq protein. Comparison of Gαq in our optimized fully active 25CN-NBOH-5-HT_{2A}-Gq protein with (A) inactive Gαq protein-bound GDP and (B) fully active nucleotide free mini-Gαq subunit (54) resolved by cryo-EM. Comparison of the cytoplasmic region of 5-HT_{2A} in our optimized fully active 25CN-NBOH-5-HT_{2A}-Gq protein with (C) the inactive conformation (54) (Protein Data Bank [PDB]: 6WH4) resolved by X-ray crystallography and (D) the active conformation (54) (PDB: 6WHA) resolved by cryo-EM.

This is in stark contrast to our GP-first paradigm in which the strong affinity between GP and GPCRs leads to opening the tight TM3-TM6 coupling. Thus, prior to ligand binding, GP can interact directly with GPCRs to stabilize a precoupled complex (23–28). Indeed, it was shown experimentally for the M3 muscarinic receptor that the precoupled complex between inactive GP and inactive GPCR eventually leads to rapid GP activation once the agonist binds the receptor (26), consistent with our results. Moreover, the basal activity of many GPCRs in the absence of an agonist leads to constitutive activation of the GP (21), showing that GP activation can proceed without agonists, further evidence supporting our results that GP precoupling to

GPCRs is a viable activation pathway. Interestingly, the cryo-EM structure of apo-5-HT_{1A}-Gi protein (30) indicates that GP precoupling to the GPCR results in a complex similar to the one resolved in the presence of a full agonist.

Thus, agonists must bind sufficiently strongly to the precoupled GPCR-GP complex to force opening of the Gα subdomain. Indeed, for 5-HT_{2A}, we found that the full agonist leads to a smaller barrier for Gα opening, compared to a partial agonist or inverse agonist, to induce release of GDP while progressing toward the activated structure.

Unfortunately, the only knowledge about this ligand bound precoupled structure is from our simulations. The structure for

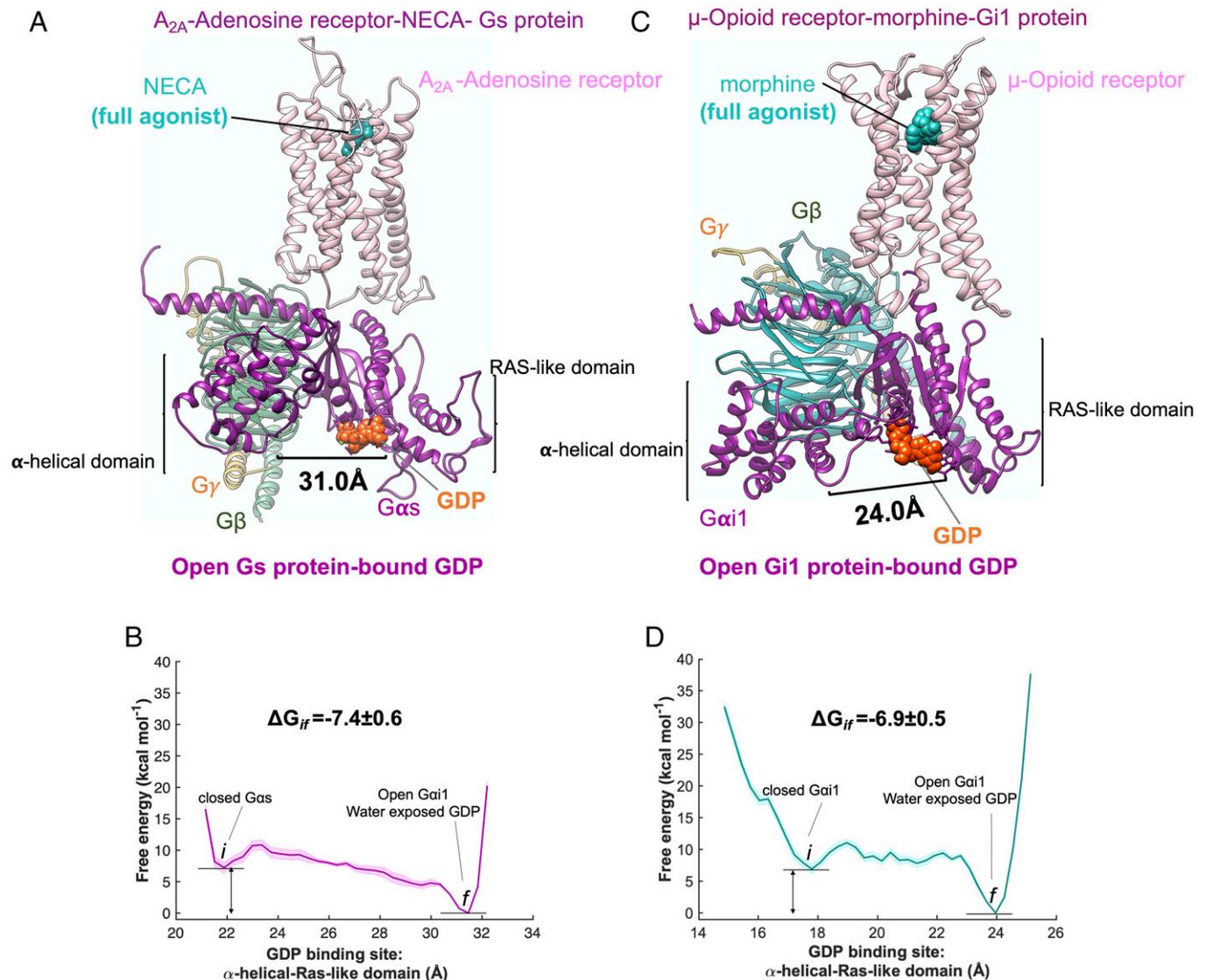


Fig. 6. Ligand activation of precoupled GPCR-GP complex. Binding of a full agonist (A) NECA to the precoupled A_{2A} adenosine receptor-Gs protein complex and (C) morphine to the precoupled μ -opioid receptor-Gi protein complex. MetaD free energy of (B) opening $G\alpha_s$ -bound GDP from its GDP binding site, distance between AH domain (center of mass of $C\alpha$ for residues 69 to 204) and the Ras-like domain (center of mass of $C\alpha$ for residues 49 to 65, 223 to 241, 294 to 303, and 369 to 374), and (D) opening $G\alpha_i$ -bound GDP from its GDP binding site (adapted from figure 6A from ref. 14), distance between AH domain (center of mass of $C\alpha$ for residues 147 to 181) and Ras-like domain (center of mass of $C\alpha$ for residues 42 to 59). The weighted averages and the SDs were calculated for ΔG_{if}^{\ddagger} for the converged period. The errors associated with the free energy were calculated after reweighting (51) the free energies and using the block averaging method.

the agonist bound to the precoupled complex has not yet been observed experimentally.

These insights on the mechanism of activation provide strategies for designing agonists. Thus, the key to agonist design is for the ligand to bind strongly to the target precoupled GPCR-GP complex in such a way as to reduce the barrier for opening the $G\alpha$ -bound GDP to release or exchange with GTP. Concomitantly, we need to examine the precoupled GPCR-GP complexes for all off-target GPCRs to ensure that the agonist does not activate them. Thus, armed with the precoupled GPCR-GP complexes for all relevant GPCRs, we should be able to design drugs optimally for the target signaling while

suppressing all off-target signaling (the target-antitarget strategy for drug design).

Our calculations predict at least 10 to 20 key interactions for each GPCR-GP combination (Fig. 3) to motivate mutation experiments to validate (or not) our predictions.

Data Availability. All study data are included in the article and/or *SI Appendix*. Model structures have been deposited in GitHub (<https://github.com/amafi-gpcr/G-protein-first-mechanism-of-activation-for-class-A-GPCRs-PNAS-2022>), (31).

ACKNOWLEDGMENTS. Partial support was provided by NIH (R35HL150807) and gifts to the Materials and Process Simulation Center (MSC).

1. R. J. Lefkowitz, Seven transmembrane receptors: Something old, something new. *Acta Physiol. (Oxf.)* **190**, 9–19 (2007).
2. A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, D. E. Gloriam, Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).
3. J. P. Overington, B. Al-Lazikani, A. L. Hopkins, How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).

4. R. Fredriksson, M. C. Lagerström, L.-G. Lundin, H. B. Schiöth, The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–1272 (2003).
5. A. J. Clark, The reaction between acetyl choline and muscle cells. *J. Physiol.* **61**, 530–546 (1926).
6. A. Karlin, On the application of "a plausible model" of allosteric proteins to the receptor for acetylcholine. *J. Theor. Biol.* **16**, 306–320 (1967).
7. H. R. Bourne, How receptors talk to trimeric G proteins. *Curr. Opin. Cell Biol.* **9**, 134–142 (1997).

8. R. Sounier *et al.*, Propagation of conformational changes during μ -opioid receptor activation. *Nature* **524**, 375–378 (2015).
9. G. G. Gregorio *et al.*, Single-molecule analysis of ligand efficacy in β_2 AR-G-protein activation. *Nature* **547**, 68–73 (2017).
10. D. M. Rosenbaum *et al.*, Structure and function of an irreversible agonist- $\beta(2)$ adrenoceptor complex. *Nature* **469**, 236–240 (2011).
11. R. Nygaard *et al.*, The dynamic process of $\beta(2)$ -adrenergic receptor activation. *Cell* **152**, 532–542 (2013).
12. A. Manglik *et al.*, Structural insights into the dynamic process of β_2 -adrenergic receptor signaling. *Cell* **161**, 1101–1111 (2015).
13. M. T. Lerch *et al.*, Viewing rare conformations of the β_2 adrenergic receptor with pressure-resolved DEER spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 31824–31831 (2020).
14. A. Mafi, S.-K. Kim, W. A. Goddard, The G protein-first activation mechanism of opioid receptors by Gi protein and agonists. *QRB Discov.* **2**, E9 (2021).
15. H. E. Kato *et al.*, Conformational transitions of a neurotensin receptor 1-G₁ complex. *Nature* **572**, 80–85 (2019).
16. R. O. Dror *et al.*, Activation mechanism of the β_2 -adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18684–18689 (2011).
17. J. Orly, M. Schramm, Coupling of catecholamine receptor from one cell with adenylate cyclase from another cell by cell fusion. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 4410–4414 (1976).
18. A. M. Tolkovsky, A. Levitzki, Mode of coupling between the β -adrenergic receptor and adenylate cyclase in turkey erythrocytes. *Biochemistry* **17**, 3795–3810 (1978).
19. A. G. Gilman, G proteins: Transducers of receptor-generated signals. *Annu. Rev. Biochem.* **56**, 615–649 (1987).
20. T. M. Cabrera-Vera *et al.*, Insights into G protein structure, function, and regulation. *Endocr. Rev.* **24**, 765–781 (2003).
21. R. Leurs, M. J. Smit, A. E. Alewijnse, H. Timmerman, Agonist-independent regulation of constitutively active G-protein-coupled receptors. *Trends Biochem. Sci.* **23**, 418–422 (1998).
22. K. A. Berg, J. A. Harvey, U. Spampinato, W. P. Clarke, Physiological relevance of constitutive activity of 5-HT_{2A} and 5-HT_{2C} receptors. *Trends Pharmacol. Sci.* **26**, 625–630 (2005).
23. C. Galés *et al.*, Probing the activation-promoted structural rearrangements in preassembled receptor-G protein complexes. *Nat. Struct. Mol. Biol.* **13**, 778–786 (2006).
24. M. A. Ayoub *et al.*, Real-time analysis of agonist-induced activation of protease-activated receptor 1/G α 1 protein complex measured by bioluminescence resonance energy transfer in living cells. *Mol. Pharmacol.* **71**, 1329–1340 (2007).
25. M. Nobles, A. Benians, A. Tinker, Heterotrimeric G proteins precouple with G protein-coupled receptors in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18706–18711 (2005).
26. K. Qin, C. Dong, G. Wu, N. A. Lambert, Inactive-state preassembly of G(q)-coupled receptors and G(q) heterotrimers. *Nat. Chem. Biol.* **7**, 740–747 (2011).
27. M. B. Kilander *et al.*, Disheveled regulates precoupling of heterotrimeric G proteins to Frizzled 6. *FASEB J.* **28**, 2293–2305 (2014).
28. K. W. Andressen *et al.*, Related GPCRs couple differently to G_s: Preassociation between G protein and 5-HT₁ serotonin receptor reveals movement of G α_s upon receptor activation. *FASEB J.* **32**, 1059–1069 (2018).
29. A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603 (2008).
30. P. Xu *et al.*, Structural insights into the lipid and ligand regulation of serotonin receptors. *Nature* **592**, 469–473 (2021).
31. A. Mafi, S.-K. Kim, W. A. Goddard, G-protein-first-mechanism-of-activation-for-class-A-GPCRs-PNAS-2022. GitHub. <https://github.com/amafi-gpcr/G-protein-first-mechanism-of-activation-for-class-A-GPCRs-PNAS-2022>. Deposited 5 April 2022.
32. J. A. Ballesteros, H. Weinstein, *Methods in Neurosciences* (Elsevier, 1995), vol. **25**, pp. 366–428.
33. G. Pándy-Szekeres *et al.*, GPCRdb in 2018: Adding GPCR structure models and ligands. *Nucleic Acids Res.* **46**, D440–D446 (2018).
34. B. K. Kobilka, G protein coupled receptor structure and activation. *Biochim. Biophys. Acta* **1768**, 794–807 (2007).
35. J. A. Ballesteros *et al.*, Activation of the β_2 -adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.* **276**, 29171–29177 (2001).
36. X. Yao *et al.*, Coupling ligand structure to specific conformational switches in the β_2 -adrenoceptor. *Nat. Chem. Biol.* **2**, 417–422 (2006).
37. R. O. Dror *et al.*, Identification of two distinct inactive conformations of the β_2 -adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4689–4694 (2009).
38. X. Liu *et al.*, Structural insights into the process of GPCR-G protein complex formation. *Cell* **177**, 1243–1251.e12 (2019).
39. G. DeMars, F. Fanelli, D. Puett, The extreme C-terminal region of Gas differentially couples to the luteinizing hormone and beta2-adrenergic receptors. *Mol. Endocrinol.* **25**, 1416–1430 (2011).
40. J. Huang *et al.*, CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
41. C. J. Dickson *et al.*, Lipid14: The amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
42. G. Navarro *et al.*, Evidence for functional pre-coupled complexes of receptor heteromers and adenyllyl cyclase. *Nat. Commun.* **9**, 1242 (2018).
43. S. K. Huang *et al.*, Delineating the conformational landscape of the adenosine A_{2A} receptor during G protein coupling. *Cell* **184**, 1884–1894.e14 (2021).
44. D. A. Shapiro, K. Kristiansen, D. M. Weiner, W. K. Kroeze, B. L. Roth, Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. *J. Biol. Chem.* **277**, 11441–11449 (2002).
45. S. R. Sprang, G protein mechanisms: Insights from structural analysis. *Annu. Rev. Biochem.* **66**, 639–678 (1997).
46. W. M. Oldham, H. E. Hamm, Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat. Rev. Mol. Cell Biol.* **9**, 60–71 (2008).
47. D. W. Markby, R. Onrust, H. R. Bourne, Separate GTP binding and GTPase activating domains of a G alpha subunit. *Science* **262**, 1895–1901 (1993).
48. B. Carpenter, R. Nehmé, T. Warne, A. G. Leslie, C. G. Tate, Structure of the adenosine A(2A) receptor bound to an engineered G protein. *Nature* **536**, 104–107 (2016).
49. R. O. Dror *et al.*, Signal transduction. Structural basis for nucleotide exchange in heterotrimeric G proteins. *Science* **348**, 1361–1365 (2015).
50. A. Mafi, S.-K. Kim, K. C. Chou, B. Güthrie, W. A. Goddard III, Predicted structure of fully activated Tas1R3/1R3' homodimer bound to G protein and natural sugars: Structural insights into G protein activation by a class C sweet taste homodimer with natural sugars. *J. Am. Chem. Soc.* **143**, 16824–16838 (2021).
51. D. Branduardi, G. Bussi, M. Parrinello, Metadynamics with adaptive Gaussians. *J. Chem. Theory Comput.* **8**, 2247–2254 (2012).
52. S. G. Rasmussen *et al.*, Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–555 (2011).
53. D. Hilger *et al.*, Structural insights into differences in G protein activation by family A and family B GPCRs. *Science* **369**, eaba3373 (2020).
54. K. Kim *et al.*, Structure of a hallucinogen-activated Gq-coupled 5-HT_{2A} serotonin receptor. *Cell* **182**, 1574–1588.e19 (2020).