

RESEARCH ARTICLE SUMMARY

SIGNAL TRANSDUCTION

Molecular determinants of ligand efficacy and potency in GPCR signaling

Franziska M. Heydenreich*, Maria Marti-Solano†, Manbir Sandhu†, Brian K. Kobilka*, Michel Bouvier*, M. Madan Babu*

INTRODUCTION: Heterotrimeric guanine nucleotide-binding protein (G protein)-coupled receptors (GPCRs) bind extracellular ligands to modulate intracellular signaling responses. Two fundamental properties of ligand-receptor-signaling systems are efficacy (the maximum achievable response) and potency (the ligand concentration required to produce the half-maximal response). Although efficacy and potency have been measured for numerous ligand-receptor-signaling systems for several decades, the molecular determinants and principles governing these pharmacological properties remain a mystery.

RATIONALE: Understanding how each receptor residue contributes to efficacy and potency can aid in designing drugs to elicit desired signaling responses. Using the adrenaline- β_2 adrenergic receptor (β_2 AR)- G_s system, we perturbed the side chain of each of the receptor's 412 residues and determined the impact on efficacy and potency. By developing a data science framework that integrates pharmacological and structural data, we contextualized ligand-induced structural changes and revealed the principles of efficacy and potency.

RESULTS: Only 20% of the β_2 AR residues contribute to the receptor's pharmacological properties. One-third of those pharmacologically relevant residues map to ligand- or G protein-binding sites or to evolutionarily conserved motifs; the other two-thirds are distributed throughout the receptor. All ligand-binding residues are important for signaling but show differential contributions to efficacy and potency, which indicates that specific contacts between ligand and receptor could be modulated to fine-tune pharmacological properties. By contrast, only one-third of the receptor residues that contact the G protein contribute to pharmacology, which indicates that these positions tolerate mutations and provides an explanation for the evolvability of G protein selectivity.

We integrated structural data about active and inactive receptor conformations with pharmacological measurements. Not all residues that undergo structural change are pharmacologically important, and several pharmacologically important residues do not undergo structural change upon activation. We categorized all receptor residues into four classes: Driver and passenger residues mediate active

state-specific contacts, and, whereas drivers affect pharmacology when mutated, passengers do not. Modulator and bystander residues do not mediate active state-specific contacts, but modulators affect pharmacology when mutated, whereas bystanders do not.

We uncovered an allosteric network of active state-specific contacts mediated by driver residues from the ligand-binding pocket to the G protein-binding interface, thereby identifying structural changes that are pharmacologically relevant. Modulator residues are located near the allosteric network and functional sites. Surface-exposed driver, modulator, and passenger residues represent key allosteric sites and potential targets for the development of new allosteric ligands. Analysis of human polymorphisms and the conservation of residues across species revealed that passenger, modulator, and driver residues are under increasing selection pressure compared with bystander residues in this receptor.

CONCLUSION: Our work reveals how a GPCR decodes and translates the information encoded in a ligand to mediate a distinct signaling response. We anticipate that the application of this data science framework will enable the design of orthosteric and allosteric molecules that can elicit defined signaling responses. ■

The list of author affiliations is available in the full article online.

*Corresponding author. Email: franziska.heydenreich@pharmazie.uni-marburg.de (F.M.H.); kobilka@stanford.edu (B.K.K.); michel.bouvier@umontreal.ca (M.B.); madan.babu@stjude.org (M.M.B.)

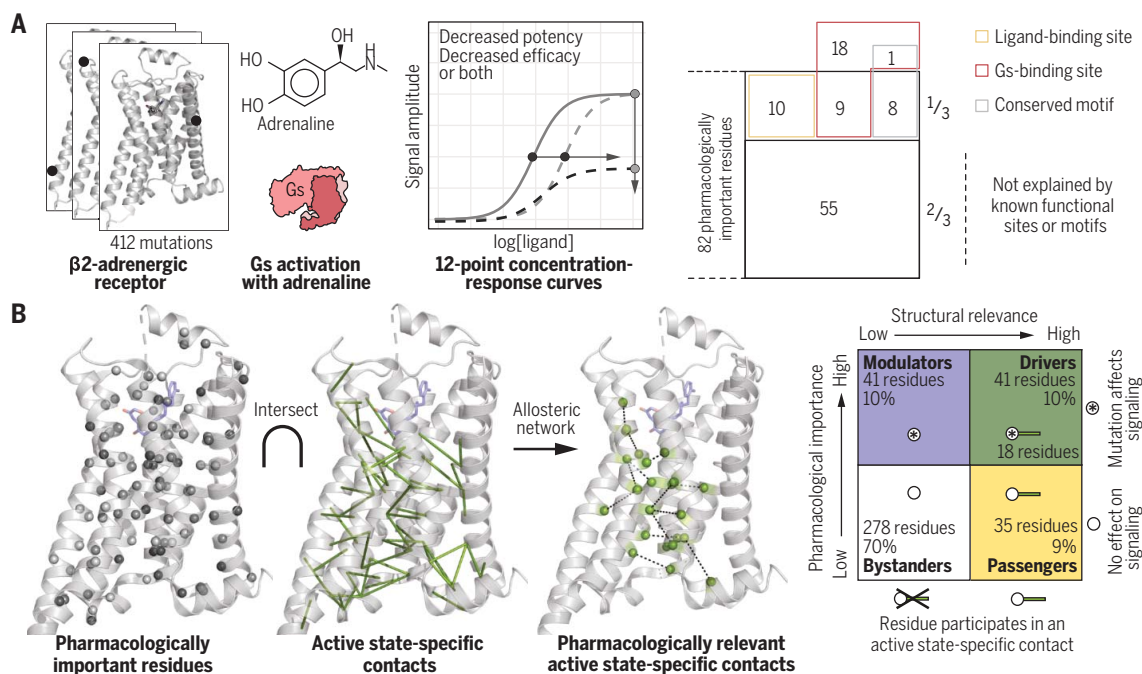
†These authors contributed equally to this work.

Cite this article as F. M. Heydenreich et al., *Science* 382, eadh1859 (2023). DOI: 10.1126/science.adh1859

READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.adh1859>

Determinants and molecular origins of efficacy and potency in a ligand-receptor-signaling system. (A and B)

By developing a data science framework that integrates experimental values of adrenaline-stimulated G_s signaling for 412 mutants of the β_2 AR (A) and data on structural changes upon ligand-induced receptor activation (B), we reveal the determinants and the allosteric network governing ligand efficacy and potency in this prototypical GPCR system.



RESEARCH ARTICLE

SIGNAL TRANSDUCTION

Molecular determinants of ligand efficacy and potency in GPCR signaling

Franziska M. Heydenreich^{1,2,3*}†, Maria Marti-Solano^{2,4}‡, Manbir Sandhu⁵‡, Brian K. Kobilka^{1*}, Michel Bouvier^{3*}, M. Madan Babu^{2,5*}

Heterotrimeric guanine nucleotide-binding protein (G protein)-coupled receptors (GPCRs) bind to extracellular ligands and drugs and modulate intracellular responses through conformational changes. Despite their importance as drug targets, the molecular origins of pharmacological properties such as efficacy (maximum signaling response) and potency (the ligand concentration at half-maximal response) remain poorly understood for any ligand-receptor-signaling system. We used the prototypical adrenaline- β_2 adrenergic receptor-G protein system to reveal how specific receptor residues decode and translate the information encoded in a ligand to mediate a signaling response. We present a data science framework to integrate pharmacological and structural data to uncover structural changes and allosteric networks relevant for ligand pharmacology. These methods can be tailored to study any ligand-receptor-signaling system, and the principles open possibilities for designing orthosteric and allosteric compounds with defined signaling properties.

Heterotrimeric guanine nucleotide-binding protein (G protein)-coupled receptors (GPCRs) constitute a major family of membrane proteins that respond to diverse extracellular ligands, including photons, small molecules, neurotransmitters, and hormones (1–5). In humans, more than 500 endogenous ligands (6) act on 800 GPCRs and regulate key aspects of human biology (7, 8). As a result of their capacity to modulate human physiology, GPCRs are targets of about one-third of all US Food and Drug Administration (FDA)-approved drugs for various diseases (9). Ligand binding to a GPCR induces conformational changes that enable modulation of downstream signaling responses (10). During that process, the ligand shifts the conformational equilibrium of the receptor toward a more active or inactive state, depending on whether it is an agonist, antagonist, or inverse agonist (11, 12). The conformational states associated with agonist binding strongly increase the likelihood of a signaling response (e.g., G protein activation). This agonist capac-

ity to promote signaling has been classically quantified by two fundamental pharmacological parameters: efficacy and potency (13), which can be determined by monitoring the receptor signaling response at various ligand concentrations (concentration-response curves). Efficacy relates to the maximum amplitude of a signaling response, whereas potency refers to the agonist concentration at which signaling reaches the half-maximal response. Endogenous agonist efficacy and potency are within a physiologically relevant range as a result of evolutionary selection pressure on the tertiary structure of GPCRs. Thus, understanding how the ligand interacts with the receptor to modulate the signaling response is paramount to rationalizing the link between endogenous ligand properties and receptor sequence and structure, which can aid in designing drugs that engage specific receptor residues to elicit distinct signaling responses in a disease context.

Numerous structural studies have revealed the importance of the movement of specific transmembrane helices as a key feature that is associated with receptor activation upon ligand binding (14–16). Furthermore, changes in the size of the ligand-binding cavity have been associated with high-affinity ligand binding and the strength of the signaling response (17). Extensive mutagenesis studies have identified residues along the receptor sequence that are important for modulating downstream signaling activity (18, 19). Despite the thorough pharmacological, structural, and mutational characterization of numerous ligand-GPCR pairs over the years, we do not yet fully understand how specific residues of the receptor decode and translate the information that is encoded in the ligand atoms to mediate a dis-

tinct intracellular signaling response (i.e., the molecular origins of efficacy and potency). For instance, in the well-studied adrenaline- β_2 adrenergic receptor (β_2 AR)-Gs signaling system, it is unclear whether all residues that contact the ligand or the G protein (fig. S1) are important for adrenaline's efficacy and potency toward Gs and whether all residues that are involved in structural changes during activation are important for signaling (fig. S1). If not, it would be helpful to identify the residues that help convert ligand binding into a signaling response and how they contribute to the efficacy and potency response of the ligand at the receptor.

To systematically infer how ligand atoms are decoded by the receptor sequence and structure to modulate downstream signaling response, one needs to perturb the side chain of every receptor residue, reliably measure the impact on efficacy and potency for a signaling response, and contextualize the structural changes of the receptor residue with its pharmacological importance. We used β_2 AR, a well-studied, prototypical family A GPCR that mediates the fight-or-flight response when activated by its endogenous agonist adrenaline (epinephrine) and signals through the G protein Gs. By developing a data science framework that integrates (i) experimentally measured values of adrenaline-stimulated Gs signaling upon systematically perturbing the side chain of every residue in β_2 AR and (ii) data on structural changes associated with agonist binding, receptor activation, and Gs coupling (Fig. 1), we reveal the molecular origins and principles that govern ligand efficacy and potency in this prototypical GPCR signaling system.

BRET-based biosensors allow in-depth pharmacological profiling of receptor mutants

We perturbed the side chain of each of the 412 residues in the receptor by mutating the residue to alanine, or glycine if the native amino acid was alanine, because both types of substitutions are well tolerated in the long, membrane-spanning α helices (20) of a GPCR. For each mutant, we then evaluated its adrenaline-stimulated Gs signaling profile in a live-cell assay with a bioluminescence resonance energy transfer (BRET)-based biosensor (21, 22) (fig. S2 and Materials and methods). The biosensor reports on the distance between the $G\alpha$ and $G\gamma$ subunits of the heterotrimeric G protein; this distance increases upon receptor activation as a result of conformational changes in the α subunit and its dissociation from the $\beta\gamma$ subunits, leading to a decrease of resonance energy transfer between the donor and acceptor that are fused to the α and γ subunit, respectively. Each assay was performed as 12-point concentration-response curves in biological triplicates (Fig. 1,

¹Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA, USA. ²MRC Laboratory of Molecular Biology, Cambridge, UK.

³Department of Biochemistry and Molecular Medicine, Institute for Research in Immunology and Cancer, Université de Montréal, Montreal, QC, Canada. ⁴Department of Pharmacology, University of Cambridge, Cambridge, UK.

⁵Department of Structural Biology and Center of Excellence for Data-Driven Discovery, St. Jude Children's Research Hospital, Memphis, TN, USA.

*Corresponding author. Email: franziska.heydenreich@pharmazie.uni-marburg.de (F.M.H.); kobilka@stanford.edu (B.K.K.); michel.bouvier@umontreal.ca (M.B.); madan.babu@stjude.org (M.M.B.)

†Present address: Department of Pharmaceutical Chemistry, University of Marburg, Marburg, Germany.

‡These authors contributed equally to this work.

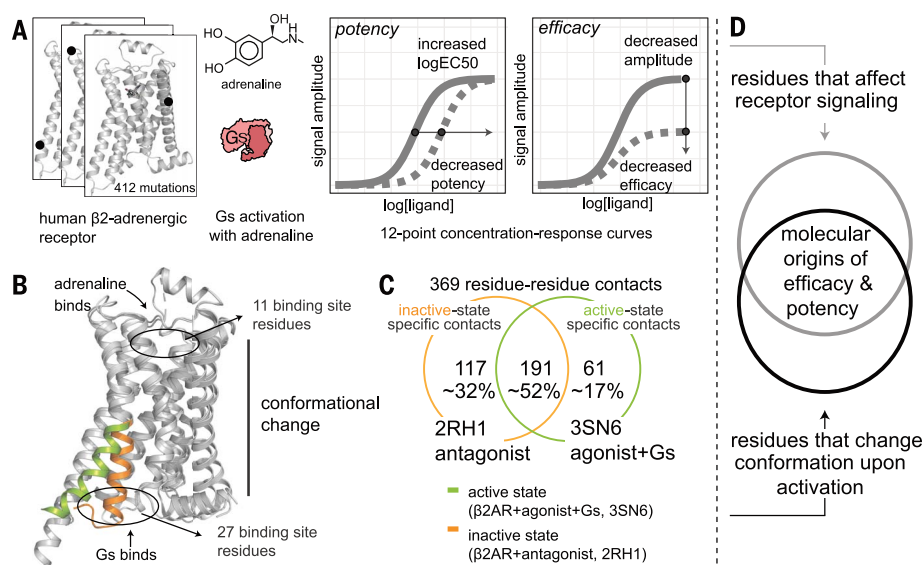


Fig. 1. An approach to integrate pharmacological and structural data to reveal the molecular determinants of efficacy and potency. (A) Every residue of a receptor is mutated, and key pharmacological properties of the adrenaline-β2AR-Gs signaling system are determined experimentally. (B and C) Active- and inactive-state structures of receptors are analyzed as networks of noncovalent contacts between residues to infer contacts that are specific for the active and/or inactive state. (D) These data are integrated using a data science approach to discover the molecular determinants and the underlying allosteric network governing efficacy and potency.

top panel, and Materials and methods) to quantify the signal amplitude (efficacy) and the log mean effective concentration ($\log EC_{50}$) (potency) of the adrenaline-stimulated G protein activation (fig. S2 and Materials and methods). Because a mutation may affect receptor abundance and low cell-surface abundance could negatively influence the signaling, we also evaluated the abundance of all mutants by cell-surface enzyme-linked immunosorbent assay (ELISA) (fig. S3, A to C, and Materials and methods). We performed measurements on cells expressing wild-type β2AR or no receptor as controls. In total, we obtained more than 16,000 signaling data points; the pharmacological properties for agonist-induced Gs activation for each mutant are reported in table S1. For subsequent analyses of these data, we used the GPCRdb numbering scheme (23), a system based on Ballesteros-Weinstein numbering (24), wherein the first number in the superscript refers to the helix or loop, and the second number refers to the position relative to the most conserved residue (numbered 50) in that helix [e.g., L124^{3x43}, which is 7 residues before the most conserved residue in transmembrane helix 3 (TM3)].

One-fifth of the positions in β2AR are important for adrenaline efficacy, potency, or both

For 16 receptor positions, side chain perturbation severely affected cell-surface abundance (<25% of wild-type level; Materials and methods). These mutants were excluded from fur-

ther analysis because we could not reliably determine whether the observed reduction in signaling resulted from altered function, reduced abundance, or both (fig. S3, A to C). Many of these positions are highly conserved and are likely important for protein folding, receptor biogenesis, and transport to the cell membrane (fig. S3, D to H). The remaining 396 mutants had cell-surface abundance that allowed reliable estimation of pharmacological parameters and were therefore considered for further analysis (abundance >25% of the wild-type β2AR; Materials and methods). Most of the single-point mutations (~80%; at 314 positions) did not negatively affect Gs signaling, and ~20% (at 82 positions) had an impact on signaling. Of these 82, 21 primarily reduced efficacy, 37 primarily reduced potency, 21 reduced both potency and efficacy, and 3 resulted in no measurable signal (Fig. 2, A and B).

Mapping these positions on the secondary structure elements of the receptor showed that TM5 was enriched in mutations that affected potency only, whereas TM3 was enriched in mutations that affected efficacy only. The mutations that affected efficacy and potency simultaneously were distributed across the different transmembrane helices (fig. S4, A to D). Although some of the identified positions overlap with functional residues in the ligand-binding pocket and the G protein-binding interface, several positions mapped to other parts of the receptor, including residues on the receptor surface, with the highest

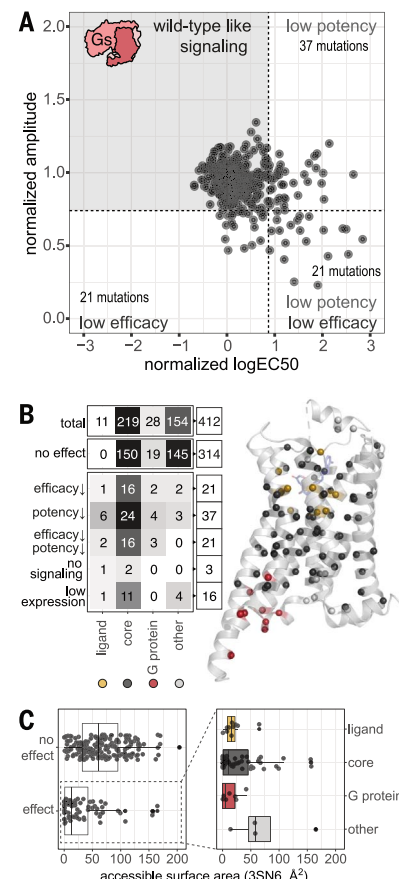


Fig. 2. Receptor positions that affect efficacy or potency upon mutation. (A) Overview of potency (normalized $\log EC_{50}$) and efficacy (normalized signal amplitude) of all β2AR mutations. Cutoffs are shown as dashed lines (Materials and methods). (B) Distribution of residues by functional relevance and their effect on efficacy and potency (Materials and methods). Ligand-binding pocket and G protein-binding site include residues within 4 Å of adrenaline (PDB ID: 4LDO) or Gs (PDB ID: 3SN6), respectively. (C) Accessible surface area of residues, grouped by whether a mutation affects signaling (effect) or not (no effect). For residues where signaling was affected by mutation, the location of mutations in the ligand-binding site, in the core, in the G protein-binding site, or in none of the above (other) is indicated.

density mapping to the receptor core (Fig. 2, B and C). These findings collectively indicate that the efficacy and potency of this endogenous agonist at this receptor are likely governed by a subset of receptor positions (~20%) that are not just restricted to known functional sites but are also distributed in different parts of the structure. These receptor positions thus decode and translate ligand binding into a signaling response and are, hence, pharmacological determinants for this ligand-receptor-signaling system.

Receptor residue–ligand atom contacts contribute differentially to efficacy and potency

We investigated how each ligand-contacting residue contributes to efficacy and potency. For this, we first identified all receptor residues that contact adrenaline using a distance cutoff [Protein Data Bank (PDB) ID: 4LDO (25)] (Fig. 3A and Materials and methods) and

the ligand atoms that they contact by constructing a receptor residue–ligand atom contact matrix (Fig. 3B). Although mutants of all 10 of the 11 adrenaline-contacting residues that expressed well negatively affected the signaling response, they had distinct effects on efficacy and potency (Y316^{7x42} markedly decreased receptor abundance; fig. S3). Mutations in around half of the positions (6 of 10)

affected potency only, one affected efficacy only, two affected both efficacy and potency when mutated, and D113^{3x32} completely abolished signaling. D113^{3x32} is critical for agonist binding in β 2AR (26, 27) and other aminergic receptors because it contacts the positively charged amine group, which is a common feature of endogenous agonists for monoamine receptors. When this information was mapped onto the ligand-binding pocket (Fig. 3C), it was obvious that the ligand-contacting residues that affect potency only upon mutation group together on one side of the receptor involving TM3, TM5, and TM6, whereas those that affect both efficacy and potency upon mutation group together on the other side of the receptor, largely involving TM6 and TM7. The ligand-contacting residue that affected efficacy only is in TM3 and is deeply buried in the binding pocket. These results indicate that each receptor residue that contacts the ligand contributes differentially to the overall pharmacological response. Of the 159 Å² of receptor surface area occupied by the ligand, 20% (31.5 Å²) appears to be important for both efficacy and potency, 6% (9 Å²) for efficacy only, and 59% (93 Å²) for potency only. Thus, the pharmacological parameters of a ligand at a receptor seem to emerge from the differential contribution of each receptor residue to the overall efficacy and potency in response to a ligand.

From the ligand perspective, the distance-based contact map shows that some adrenaline atoms are in close proximity to residues that only affect potency when mutated (the oxygen atoms in the hydroxyl groups and the aromatic carbon C3; Fig. 3B). The map also shows that the aromatic atoms C1, C2, C5, and C6 are in close proximity to residues that affect both efficacy and potency when mutated, including the N312^{7x38} and the key D113^{3x32} (O3, C7, C8, N1, and C9). These findings reveal the existence of spatially clustered groups of adrenaline atoms that mediate contact with specific receptor residues to preferentially affect efficacy and/or potency responses in receptor signaling. An implication of this observation is that chemical modifications of these atoms that can mediate contacts with specific receptor residues might allow a more targeted ligand modification to alter the efficacy or potency of Gs activation. This can be seen from the receptor residue–ligand atom contact matrix for the β 2-adrenergic ligand salmeterol (fig. S5). Together, these results indicate that, although all ligand-contacting residues are important for receptor signaling, distinct receptor residues and specific chemical groups on the ligand contribute differentially to efficacy and potency responses. Thus, although efficacy

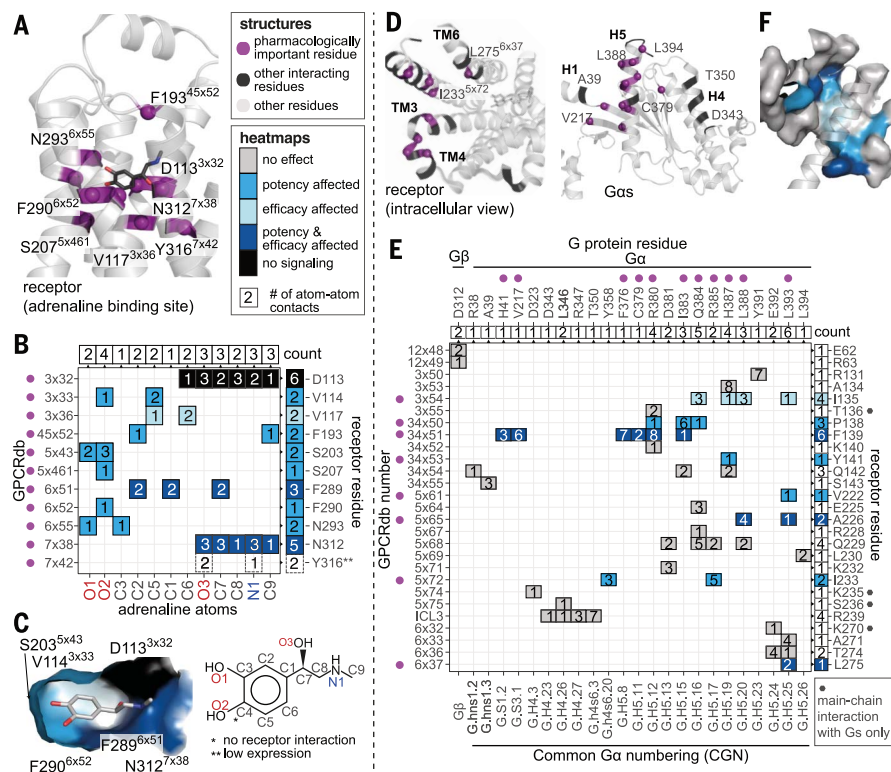


Fig. 3. Effect of mutations in the ligand- and G protein-binding sites. (A) View of the ligand-binding site; positions are labeled with residue number and GPCRdb number in superscript. Pharmacologically important residues (according to cutoff values; see Materials and methods) are indicated in violet. The following are not labeled for clarity: V114^{3x33}, S203^{5x43}, and F289^{6x51}. (B) Receptor residue–ligand atom contact plot with the ligand atoms on the x axis and the receptor's GPCRdb and residue numbers on the y axis. The number of noncovalent contacts between a receptor residue and ligand atom is shown in each square of the heatmap. The chemical structure of adrenaline below the heatmap indicates the labeling of adrenaline atoms used for the x axis. Box colors in the heatmap refer to the pharmacological effect of the mutation [efficacy affected (light blue), potency affected (ocean blue), both efficacy and potency affected (dark blue), or no measurable signaling (black)]. The number of receptor residues contacted by each ligand atom and the number of ligand atoms contacted by each receptor residue are indicated in boxes at the top and right-hand side of the heatmap, respectively. (C) Surface view of the ligand-binding site, where the surface of each receptor residue is colored by the effect that its mutation had on signaling. (D) Views of the G protein-binding site on the receptor and the Gs protein. Positions are labeled with residue number and with GPCRdb numbers. Pharmacologically important residues are indicated in violet, and other interacting residues are marked in black. (E) Receptor–G protein residue contact plot based on the structure of β 2AR in complex with heterotrimeric Gs (PDB ID: 3SN6). G protein residues are shown on the x axis and the receptor residues on the y axis. The receptor residues that contact the G protein through a receptor main chain contact only are marked with a dark gray hexagon on the right side of the matrix. The number of noncovalent contacts between a receptor residue and a G protein residue are mentioned in each square of the heatmap. Boxes are colored as in (B). The CGN number is provided on the x axis. The number of contacting residues by the receptor and the G protein are shown on the top and right-hand side of the heatmap. Residue T68 is not included because it is a contact that is observed only in the structure of β 2AR with Gs peptide (PDB ID: 6E67). (F) Surface view of the G protein-binding site colored by mutational effect, with helix 5 shown as cartoon. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

and potency are measured as a general ligand property at a receptor, they arise from a collection of receptor residue–ligand atom noncovalent contacts, each of which contributes differentially to pharmacological properties in downstream signaling response.

Only one-third of the G protein–contacting residues are important for efficacy and potency

From the structures of the $\beta 2\text{AR}$ -Gs complex (PDB IDs: 3SN6 and 6E67), we observed that 27 receptor residues make 49 noncovalent contacts with 24 G protein residues (Fig. 3, D to F, and Materials and methods). Only one-third (9 of 27) of the receptor interface residues showed an effect on G protein activation when mutated. This is in stark contrast to the ligand-binding residues, where all positions were important for receptor signaling (Fig. 3B). Of the positions in the G protein–binding site, four affected potency only, two affected efficacy only, and three affected both when mutated. From the G protein perspective, key conserved residues that undergo conformational changes during receptor binding [e.g., H5.12,15,16,19,20,25, S3.1, and S1.2 in the common G protein numbering (CGN) scheme (28)] mediate extensive contacts with receptor residues that affected efficacy, potency, or both when mutated (28, 29). Collectively, these results reveal that there is a subset of receptor–interface residues that translate ligand potency and efficacy into G protein activation at the receptor–G protein interface. A large fraction of receptor residues that contact the G protein can be mutated to alanine with little effect on efficacy or potency and are thus evolvable. The latter findings imply that upon gene duplication and divergence, receptors can tolerate certain types of mutations at the G protein–binding interface without losing their capacity for signal transduction. Such a structure-to-function architecture may provide an additional explanation as to why GPCRs could evolve by gene duplication and acquire new G protein selectivity (30, 31).

Conserved motifs only explain a fraction of residues important for signaling

Of all the positions that had an influence on ligand efficacy and potency, only a fraction (19 of 82 residues; 24%) were in the ligand-binding and G protein–binding sites. A detailed analysis of the key motif residues of GPCRs, as well as the most highly conserved positions (x50 positions in each helix) (Materials and methods), accounted for an additional eight residues (10%) (fig. S6). Key sequence and structural motifs with known functional roles include CWxP, PIF, NPxxY, and DRY (where each uppercase letter is an amino acid residue, and each “x” represents any amino acid). However, two-thirds of the pharmacologically important residues (55 of

82 residues) do not map to known functional sites or motifs and are hence not readily interpretable. These positions most likely represent receptor- or receptor family-specific switches or allosteric conduits that contribute to translating ligand binding into signaling.

Structural changes alone do not specify the pharmacological importance of a residue

Because key structural changes have been extensively described as a hallmark of receptor activation (14, 16, 32, 33), we investigated how those structural changes are associated with pharmacologically important residues. We calculated how much each residue moved during activation by computing the distance between the α -carbon atoms of each residue in the inactive (highest-resolution structure, 2.4 Å; PDB ID: 2RH1) and active, G protein–bound state (highest-resolution structure, 3.2 Å; PDB ID: 3SN6). The distribution of the distance moved by each residue is skewed to the right, where in ~8% of the cases, the α -carbon atom

moved more than 4 Å (22 residues; fig. S1B). Of these, only seven were pharmacologically important and affected efficacy or potency when mutated, whereas the rest had no impact on downstream signaling (Fig. 4, A and B). Furthermore, comparing the set of residues that were pharmacologically important with those that were not revealed that an approximately equal proportion between the two sets appeared to undergo a structural change (i.e., move >4 Å). Finally, an analysis of dihedral angles in the active and inactive states also revealed no association between the extent of change in the dihedral angle and pharmacological importance of a residue (fig. S7). These findings collectively suggest that it is not possible to infer the pharmacological importance of a residue just from the extent of structural change (translation or rotation) during receptor activation.

Because the formation or breaking of noncovalent contacts during activation comprises both translational (distance) and rotational

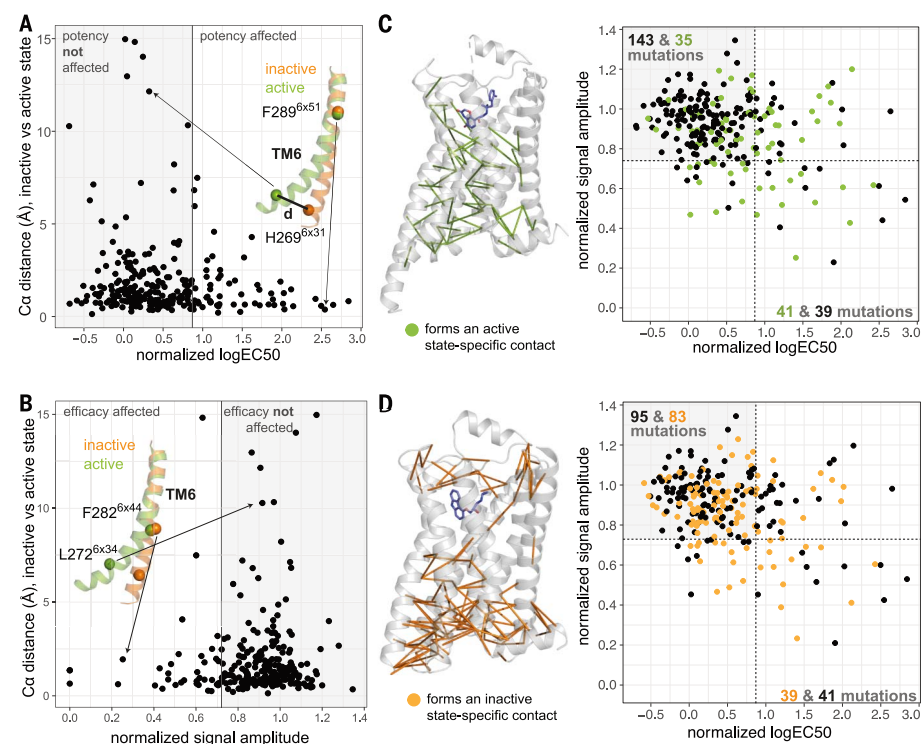


Fig. 4. Structural changes and pharmacological importance of residues. (A) Distance between C α atoms in the inactive state (PDB ID: 2RH1) versus the active, G protein–bound state (PDB ID: 3SN6) plotted against normalized Gs logEC₅₀ of the alanine or glycine mutant in that position. White and gray areas indicate residues with potency affected and not affected, respectively, upon mutation. (B) Distance between C α atoms plotted against normalized Gs signal amplitude. White and gray areas indicate residues with efficacy affected and not affected, respectively, upon mutation. (C) Active state–specific contacts (Materials and methods) are shown on the active, G protein–bound receptor structure. Potency and efficacy values for all mutants are plotted (except those with low abundance and those for which no signaling was detected). Mutants are colored green if the mutated residue is involved in an active state–specific contact and black if it is not. White areas indicate residues with efficacy, potency, or both affected, and the gray area indicates those that were not affected upon mutation. Numbers in the plots denote the number of residues in each category (black or green). (D) Inactive state–specific contacts on the inactive-state structure. A mutant (data point in the plot) is colored orange if the residue is involved in an inactive state–specific contact and black if it is not. White and gray areas of the graph are defined as in (C).

(dihedral angle) changes to different extents, we focused on residues that make active state-specific contacts or break inactive state-specific contacts to assess whether they can inform pharmacological importance (Fig. 4, C and D; Materials and methods; and table S2). Considering those two categories, the formation of an active state-specific contact by a residue showed a higher association with pharmacological importance. For instance, ~50% (41 of 80 residues with structural data) of residues that are pharmacologically important participated in an active state-specific contact, whereas only 20% (35 of 178) of the

pharmacologically unimportant residues did. This effect was not observed for inactive state-specific contacts, for which ~50% (39 of 80) of pharmacologically important residues with structural data and 47% (83 of 178) of unimportant residues participated in an inactive state-specific contact. These findings indicate that systematic integration of mutational effects and structural analyses considering active state-specific residue contacts may help identify and characterize the subset of pharmacologically important residues that undergo a structural change in this ligand-receptor system. Notably, it can explain which of the struc-

tural changes that happen during activation are relevant for ligand efficacy and potency at the receptor.

Linking structural changes with pharmacological importance reveals new roles for residues

To identify the structural changes relevant for efficacy and potency, we devised an approach to integrate the experimentally determined pharmacological parameters with the active state-specific contacts (Fig. 5A and Materials and methods). We first classified every residue as pharmacologically important (if it affects efficacy, potency, or both when mutated) or

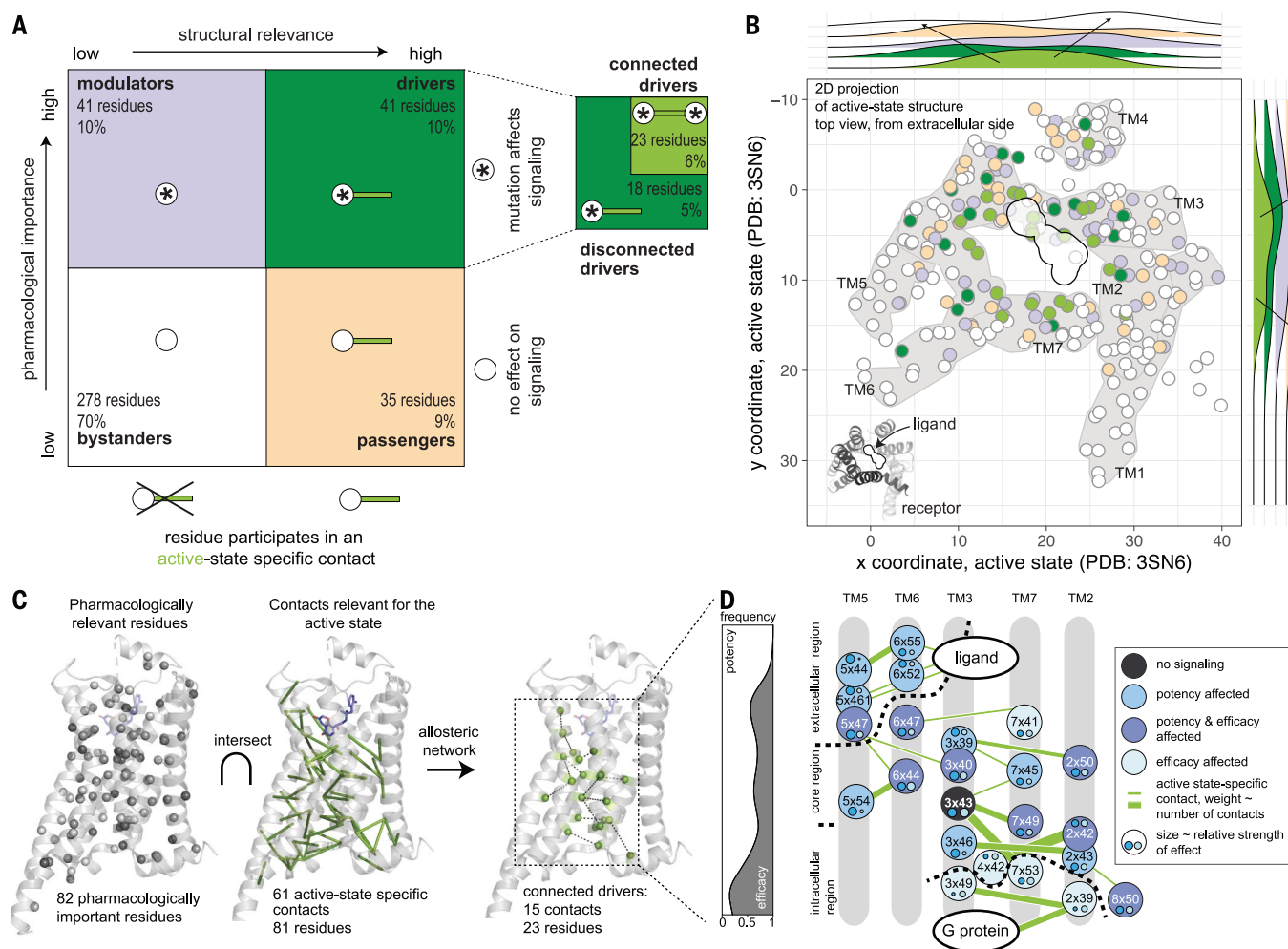


Fig. 5. The allosteric contact network for efficacy and potency as determined by structure and pharmacology. (A) Classification of residues according to their pharmacological importance (whether a mutation affected signaling or not) and structural relevance (whether a residue forms an active state-specific contact or not) into four main classes: bystanders (white), passengers (wheat), modulators (slate), and drivers (green), the latter of which is subdivided into connected drivers (driver residues connected to other driver residues) and unconnected driver residues (Materials and methods).

(B) Position of residues classified by their participation in active state-specific contacts and their importance for pharmacology on a two-dimensional (2D) top-view projection of the active, G protein-bound structure of the β 2AR (PDB ID: 3SN6). Frequency of residues along the x and y axes of the receptor (plane parallel to the membrane) is shown on the distributions

outside. Arrows denote the direction of the peaks of the distribution for the residue classes. Driver residues tend to be in the center, whereas bystander residues tend to be at the periphery. (C) Integration of pharmacologically important residues and the residues contributing to newly established contacts in the active-state structure yields the allosteric network for efficacy and potency (see main text). (D) Residues forming the allosteric networks for efficacy and potency are represented as a cartoon, colored by the effect of the mutation. The graph (left) depicts the frequency of potency versus efficacy effects for residues in the network, projected on the axis running along the extra- to intracellular side. Small circles and their sizes (scaled to the relative strength of effect) within each residue indicate the magnitude and effect on potency and efficacy. Edge thickness in the network corresponds to the number of contacts, ranging from 1 to 5.

not pharmacologically important (if it has no effect on either efficacy or potency). We then stratified according to whether a residue forms an active state-specific contact or not (i.e., whether it is structurally important or not). On the basis of these two properties, we defined four classes of residues: (i) driver residues that mediate an active state-specific contact and affect pharmacology (41 residues), (ii) modulator residues that do not form an active state-specific contact but are important for pharmacology (41 residues), (iii) passenger residues that mediate an active state-specific contact but are not important for pharmacology (35 residues), and (iv) bystander residues that neither mediate active state-specific contacts nor affect pharmacology when mutated (278 residues) (Materials and methods and Fig. 5A). Thus, the pharmacologically important residues are either drivers or modulators (82 mutations that affect pharmacology), and the structurally important residues are either drivers or passengers (76 residues that mediate an active state-specific contact). As a general pattern, driver residues tend to be close to the central axis going through the receptor and perpendicular to the membrane, whereas modulator and passenger residues are located further away from this axis, in proximity with solvent or membrane-exposed positions (Fig. 5B). These observations prompted us to investigate the existence of a network of pharmacologically relevant active state-specific residue contacts mediated by driver residues and further understand the role of modulator residues for signal transduction.

A subset of residues driving efficacy and potency form an allosteric contact network

Of the 41 driver residues, 23 form 15 active state-specific noncovalent contacts to other driver residues with a shared pharmacological effect (connected drivers) (Materials and methods). This allosteric network of active state-specific contacts mediated by driver residues allowed us to describe how potency and efficacy signals are transmitted across the receptor structure at the residue level through coordinated structural rearrangements (Fig. 5C and fig. S8). The network involving connected drivers started extracellularly in TM5 and TM6 and formed several contacts between TM5 and TM6, including with F282^{6x44} of the PIF motif (34, 35) (Fig. 5D), before reaching TM3 in the receptor core. TM3 further connected to TM7 and the most conserved residue in TM2, D79^{2x50} (26), which is part of the allosteric sodium-binding motif (36) (Fig. 5D). Finally, the network reached residues in the intracellular region, including D130^{3x49} and Y326^{7x53} of the conserved DRY and NPxxY motifs in TM3 and TM7, respectively (Fig. 5D). Potency-affecting driver residues in the network were enriched at the extracellular side of

the receptor involving TM5 and TM6 (total of 9 residues; Fig. 5D), and efficacy-affecting driver residues were enriched at the intracellular side of the receptor in TM7, TM2, and TM3 (total of 5 residues; Fig. 5D). Driver residues in the allosteric network that affected both efficacy and potency were enriched in the center of the receptor, connecting to potency-only and efficacy-only residues (9 residues; Fig. 5D). This raises the question as to how ligand binding to the receptor initiates and transmits information through the allosteric network to modulate downstream signaling.

Key contact-rewiring events initiate and transmit information to mediate signaling

To understand how information is transmitted in the adrenaline- β 2AR-Gs signaling system, we analyzed the relation between the ligand- and G protein-binding residues and the allosteric network. Focusing on the ligand-receptor interface, the allosteric contact network included three residues directly located in the ligand-binding site (N293^{6x55}, F290^{6x52}, and S207^{5x461}, all affected potency only when mutated); the other eight ligand-binding residues do not form active state-specific contacts. This highlights the existence of residues in the ligand-binding pocket that directly participate in the allosteric network. It also highlights the distinct roles mediated by the ligand-contacting residues on the receptor: although several residues stabilize ligand-receptor interactions, a subset of three driver residues that also contact the ligand likely initiate the allosteric changes during receptor activation (Fig. 5 and fig. S9). The latter set of three residues makes more extensive contact with other receptor residues compared with adrenaline. By contrast, ligand-binding residues that do not mediate active state-specific contacts (e.g., modulator residues D113^{3x32} and N312^{7x38}) make substantial contacts with adrenaline and fewer contacts with other receptor residues (fig. S9). Thus, some receptor and ligand atom contacts might contribute to the affinity of ligand and binding, whereas others might have a more direct role in triggering the allosteric network to transmit the signaling response, perhaps through initiating a structural change and stabilizing a particular receptor conformation.

Focusing on the G protein-binding residues, of the nine that affected efficacy or potency when mutated (Materials and methods), one residue (driver residue T68^{2x39}) is part of the allosteric network. Although T68 does not mediate a contact in the receptor-heterotrimeric G protein complex structure, it does form a contact with Gs in the receptor-Gs peptide complex structure (37), which may represent an intermediate state in complex formation. Furthermore, L275^{6x37} (a disconnected driver that affects efficacy when mutated) mediates contact with a conserved residue on the G pro-

tein [H5.25; CGN scheme (28)]. L275^{6x37} is held in place through an inactive state-specific contact with I127^{3x46}. Upon receptor activation, this inactive state-specific contact is broken, allowing 6x37 to contact the G protein and 3x46 to mediate an active state-specific contact with Y326^{7x53} (38). Notably, this noncovalent contact is between a residue that affects potency only (3x46) and one that affects efficacy only (7x53). Thus, the G protein-binding residues that affect efficacy and potency when mutated can be classified into those that directly interact with the G protein and those that are part of or are connected to the allosteric network (fig. S10). Together, these observations highlight the importance of key rewiring events in the allosteric network involving the switch from the inactive state-specific contacts to active state-specific contacts upon ligand binding and G protein engagement for signal transduction (Fig. 5D and fig. S10). It also highlights that as functionally relevant intermediate state conformations are structurally characterized, the roles of some of the residues identified can be interpreted in more mechanistic detail.

Modulator residues near the allosteric network and functional sites modulate pharmacology

Of the 82 residues that were important for efficacy or potency, the role of two-thirds (54 of 82 residues; 41 are driver and 13 are modulators) could be explained as either involved in ligand or G protein binding, conserved motif residues, or driver residues forming the allosteric network or that mediate active state-specific contacts (fig. S11). We then investigated the role of the remaining one-third of the residues that are pharmacologically important (28 of 82 residues) and categorized them according to their likely function (fig. S11; these are the remaining 28 of the 41 modulator residues). Of those, the function of 10 residues may be explained by their ability to directly contact and perhaps stabilize a residue in the ligand-binding pocket or by being in the putative ligand entry pathway. Another six residues mediate noncovalent contacts with a driver residue (network modifier), and the function of seven modulators may be explained by lipid interaction. The roles for the remaining five residues were less clear (table S1, explanation column). We analyzed conformations of the active and inactive states, but a deeper understanding of the role of modulator residues can be obtained through the investigation of intermediate conformational states, by considering the role of structured water molecules in stabilizing or mediating conformational transitions, or through the analyses of data describing receptor dynamics (39, 40). Taken together, these results identify and describe how residues around the allosteric network

and functional sites may modulate ligand efficacy and potency in β 2AR.

Surface-exposed driver, modulator, and passenger residues represent key allosteric sites

Some of the pharmacologically important positions mapped to surface-exposed sites that were not part of the orthosteric ligand-binding site or the G protein-binding interface (Fig. 2C). The identification of pharmacologically important residues that lie on the receptor surface suggested possible structural sites for modulating receptor signaling that could be targeted by natural allosteric molecules (41, 42). To test whether lipids, such as cholesterol, exert effects allosterically by binding to such sites and altering signaling, we analyzed structures of receptors in complex with cholesterol (Materials and methods and fig. S12). Cholesterol bound to a site containing three modulator and three passenger residues on the receptor surface. Thus, surface-exposed pharmacologically important residues might be targeted by synthetic allosteric ligands. In line

with this idea, an analysis of the structure of the receptor in complex with the negative allosteric modulator (NAM) AS408 revealed that its binding interface contained two amino acids, I214^{5x53} (driver) and C125^{3x44} (modulator), that are important for receptor signaling and three passenger residues, all of which are surface exposed (Fig. 6, A and B, and Materials and methods; 7 residues contact the NAM). Similarly, an analysis of the structure of the receptor in complex with the positive allosteric modulator (PAM) 6FA showed that it bound to a different exposed surface containing the driver residue, D130^{3x49}, which is important for receptor signaling (Fig. 6, A to C). Five of the nine residues in the PAM binding site were passengers—i.e., residues that mediate active state-specific contacts but that do not affect efficacy or potency when mutated (Fig. 6, A to C). Passenger residues were significantly overrepresented (compared with random expectation) in both PAM and NAM binding sites, even when the N and C termini were excluded from the test ($P = 1.5 \times 10^{-3}$; hypergeometric test).

These results show that allosteric modulators, which were discovered through screening efforts, tend to bind surface-exposed sites that contain driver, modulator, and passenger residues rather than the more prevalent bystander residues (i.e., residues that do not mediate active state-specific contacts and are not important for pharmacology). Although the active state-specific contacts mediated by the passenger residues might minimize entropic cost by stabilizing a specific receptor conformation, thereby energetically favoring allosteric ligand binding, the perturbation of the driver or modulator residues at such sites might affect the allosteric network and hence mediate the positive or negative cooperativity at the orthosteric ligand-binding site. In other words, a combination of (i) conformational stabilization of the active or inactive state by the allosteric ligand through interactions with passenger residues and (ii) perturbation of the allosteric network through interactions with surface-exposed modulator or driver residues might contribute to the property of the allosteric ligand. An important implication

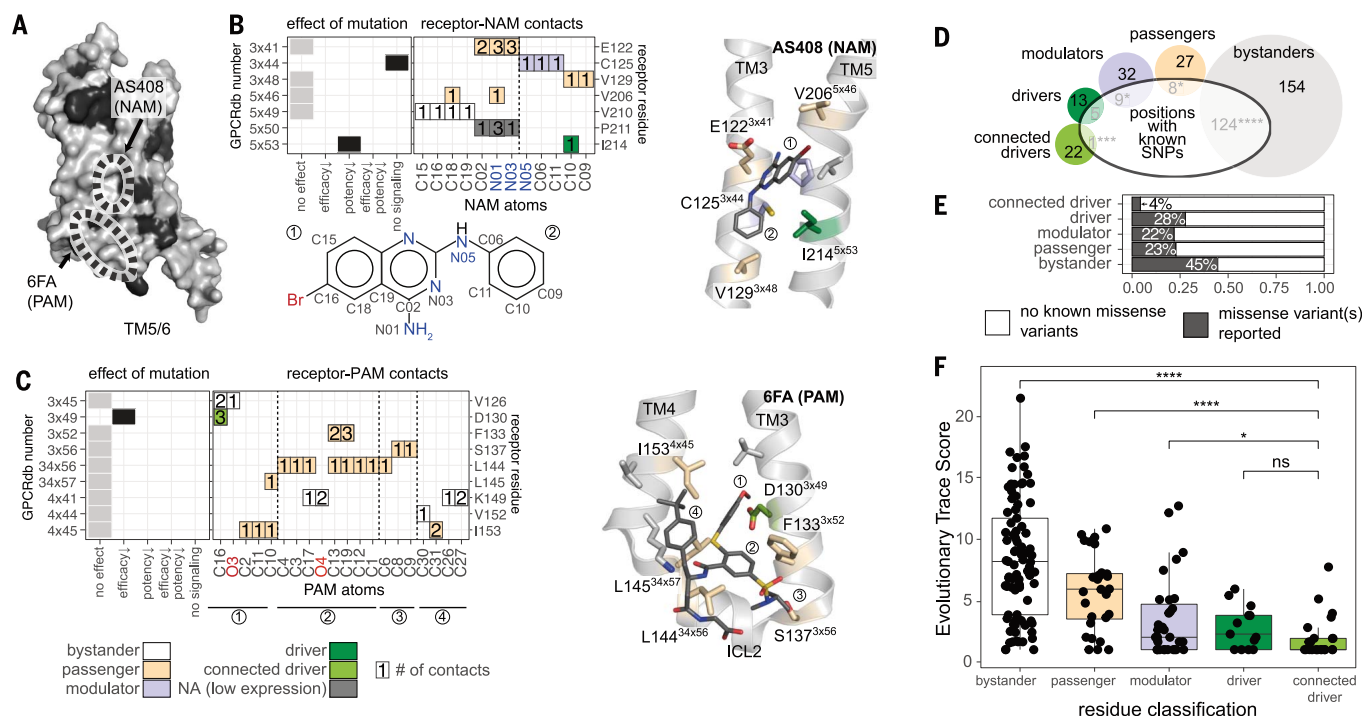


Fig. 6. Allosteric sites, evolutionary conservation, and natural variation.

(A) Surface-exposed, pharmacologically important residues are shown in black on the β 2AR (PDB ID: 3SN6). (B) Effect of mutations in the binding site for AS408, a NAM, receptor residue–ligand atom contact plot, and structural view. The contact plot shows the number of contacts between receptor residues (y axis) and NAM atoms (x axis). Colors for the receptor–NAM contacts in the plot refer to the classification of the residues as in Fig. 5. The chemical formula of the allosteric modulator below the heatmap indicates the labeling of NAM atoms used for the x axis of the contact plot. (C) Effect of mutations in the binding site for 6FA, a PAM, receptor residue–ligand atom contact plot, and structural view. The contact plot shows the number of contacts between receptor residues

(y axis) and PAM atoms (x axis). Colors for the receptor–PAM contacts in the plot refer to the classification of the residues as in Fig. 5. The circled numbers below the x axis refer to different parts of the modulator and are mapped on the structure view for reference. (D) The number of residues with known SNPs for each class. (E) Percentage of residues for which at least one SNP has been reported by class. Significant differences from the expected frequency by chance are indicated according to a hypergeometric test (* $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$; ns, not significant). (F) Adrenergic receptor ET score for the different residue classes. Groups were compared using a Wilcoxon test (* $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$; ns, not significant). Driver refers to the subset that is in the disconnected driver group.

of this observation is that the discovery of such surface-exposed driver, modulator, and passenger residues through the integration of pharmacological and structural data (in the presence of orthosteric ligands and the absence of allosteric compounds) may reveal sites that could be targeted for development of new allosteric ligands.

Passenger, modulator, and driver residues are under higher evolutionary selection pressure

To assess the importance of the residues relevant for pharmacology and/or active-state conformation, we analyzed data on natural genetic variation from the human population. An analysis of the genome sequence data from more than 140,000 individuals available through the gnomAD database (43) revealed a notable pattern in the distribution of single-nucleotide polymorphisms (SNPs) according to the different residue categories (Fig. 6, D and E). Bystander residues harbored a larger number of positions with SNPs compared with what is expected by chance. Passenger and modulator residues contained fewer positions with SNPs than expected by chance. The allosteric network mediated by the connected driver residues had the lowest number of positions harboring SNPs in the human population, which indicates that these residues were under the strongest selection pressure (Fig. 6, D and E, and Materials and methods). Thus, driver, modulator, and passenger residues appear to be under differential selection pressure compared with bystander residues of this receptor in the human population, which is consistent with the importance of the residue categories.

To assess the importance of the residues across orthologs of adrenergic receptors from different species, we performed an analysis of evolutionary trace (ET) scores in the context of our residue classification (i.e., driver, modulator, passenger, and bystander). ET is a phylogenetic method that identifies functionally important positions in protein families. Low ET scores indicate positions that are highly conserved across homologs. Depending on how the alignments are constructed, residues with low ET scores can also represent positions that are conserved in a subfamily-specific manner (44–46). Driver residues were the most conserved followed by modulator, passenger, and bystander residues (Fig. 6F). Not all highly conserved residues were pharmacologically or structurally important; hence, the role of a residue cannot be predicted from evolutionary conservation alone. These findings reveal that categorizing residues on the basis of structural and pharmacological importance can reveal category-specific associations of positions under purifying selection within the adrenergic receptors, aminergic receptors, as well as the entire family of class A GPCRs (Fig. 6F and fig. S13).

Mapping of the ET scores (adrenergic GPCRs) onto the allosteric networks revealed that driver residues with the lowest ET scores were at the center of the network in the receptor core, including all of the network residues in TM3 and TM7 (fig. S13). By contrast, the network positions with the lowest conservation (i.e., higher variability; ET scores > 3rd quartile, higher than 75% of the values) were either close to the ligand-binding site or in the intracellular region (fig. S13). These positions are more tolerant to mutations and may evolve or have evolved to recognize new ligands or different G proteins (fig. S13). This suggests that although a major part of the residues constituting the allosteric network is conserved within adrenergic GPCRs, key positions that interface with the ligand or G protein that are necessarily receptor-specific tend to be more variable. The uncoupling of ligand- and G protein-binding sites might have facilitated the evolution of GPCRs to respond to chemically diverse ligands and to couple to different G proteins. Collectively, the observations on human genetics and evolutionary data, which are independent of the approaches that we used to identify the residue classes, represent further evidence to support the model of the residues that govern the molecular origins of efficacy and potency in this signaling system.

Discussion

Ligand efficacy and potency at a receptor are the key pharmacological parameters used to describe ligand-mediated receptor signaling. Although it is widely established that agonist binding activates the receptor and mediates structural rearrangements, ultimately resulting in a signaling response, we do not fully understand the molecular origins of this process and how each amino acid in the receptor decodes the information encoded in ligand atoms and collectively translates this into the emergence of efficacy and potency. We designed an integrative approach, combining pharmacological, structural, computational, and evolutionary analyses at a per-residue level, that revealed how receptors decode the information on the ligand atoms, mediate a series of structural changes through an allosteric network, and ultimately influence G protein binding and the downstream signaling response.

One-fifth of all residues in β 2AR were important for adrenaline efficacy and potency for Gs signaling. Although some residues mapped to functionally important sites, such as ligand-binding, G protein-binding, and conserved motif residues, more than two-thirds of the pharmacologically important residues mapped to other sites on the receptor, including those on the receptor surface. A conceptual framework that combined structural and pharmacological data helped us to propose potential

roles for the receptor residues and allowed us to distinguish key, pharmacologically important structural changes from those that are less pharmacologically relevant. This framework resulted in the identification of an allosteric network of noncovalent contacts (based on active state-specific structural changes) that are mediated by pharmacologically important residues from the extracellular to the intracellular side. Our classification of residues into drivers, modulators, passengers, and bystanders provides a rational framework to explain the molecular and structural origins of potency and efficacy.

The exact nature of the pharmacologically and structurally important residues, as well as the allosteric network, likely depends on the choice of ligand, receptor, and effector. We expect considerable overlap for closely related systems based on the evolutionary analyses, but it is likely that the residues and the network will consist of nonoverlapping receptor positions that depend on the specific ligand and the effector, including the G protein subtype.

In addition to revealing such networks, the detailed analysis of ligand atom contacts with receptor residues presented in this work can help us to better understand the role of individual atoms in eliciting a pharmacological response. It can also aid in ligand design by enabling the identification of chemical groups that can be modified to achieve a desired signaling response. Our observation that surface-exposed driver, modulator, and passenger residues identified using the endogenous ligand and are contacted by NAMs and PAMs suggests that such sites should be prioritized in high-throughput virtual screening efforts for the discovery of potential allosteric modulators.

Apart from revealing general principles, the approach and the findings presented here can help us understand how disrupting and engineering parts of the network could alter receptor output, thereby facilitating the design of receptors with desired signaling properties (47). It could also help us understand how changes in receptor residues, such as SNPs or disease-related mutations, could influence signaling responses to endogenous agonists or GPCR drugs. The framework developed in this work should enable the investigation of such questions for diverse ligand-receptor-signaling systems by leveraging high-throughput approaches, such as deep mutational scanning, to pharmacologically characterize mutations and the increasing availability of new structures of ligand-receptor-signaling protein complexes. Furthermore, the framework to classify residues on the basis of pharmacological and structural relevance can be applied to any functional readout and property describing conformational change. In this manner, the approach can be extended to biological systems beyond GPCRs.

Materials and methods

Experimental methods

DNA constructs and mutagenesis

The design of the human β 2AR construct was based on the commercially available pSNAPf-ADRB2 Control Plasmid (NEB) coding sequence, which consist of a N-terminal signal sequence, the SNAP-tag, and the receptor sequence. A BamHI restriction site was added between the SNAP tag and the receptor sequence, the receptor sequence was codon optimized. The wild-type receptor amino acid at position 16 was arginine. The two most common alleles at position 16 are arginine and glycine (as defined by the 1000 Genomes Project). The SNAP- β 2AR was cloned into pcDNA3.1. BRET-based biosensor constructs for Gas, G β 1, and G γ 1 subunits were in pcDNA3.1. The Gs biosensor consisted of RlucII-Gas, where Renilla luciferase (RlucII) was inserted after amino acid 67, unmodified human G β 1, and GFP10-G γ 1.

β 2AR mutants for positions 2 to 412 were generated as described (48). Amino acids other than alanine were replaced by alanine, and native alanines were replaced with glycines. Briefly, primers were designed using the custom-made software AAScan (49) (available at <https://github.com/dmitryveprintsev/AAScan>) and ordered from Integrated DNA Technologies (IDT) as 500 pmol DNA oligos in 96-well plates. The primer sequences are available in table S3. Forward and reverse mutagenesis primers were used in separate polymerase chain reaction (PCR) reactions together with one primer each, which annealed in the ampicillin resistance gene (CTCTTACTGTGCATGCCATCCGTAA-GATGC and GCATCTTACGGATGGCATGACAG-TAAGAG). PCRs were run as touchdown PCRs in 20 μ l final volume. Resulting half-vector fragments were combined, digested with 0.5 μ l DpnI (NEB) for 1 hour, cleaned up using ZR-96 DNA Clean & Concentrator-5 (Zymo Research), assembled by Gibson assembly using HiFi DNA assembly Master Mix (NEB) at 45°C for 1 hour (50) and transformed using Mix & Go competent cells (Zymo Research). The resulting colonies were cultured in 5 ml LB medium in 24-well plates overnight. The DNA was purified using Qiaprep 96 Plus kits (Qiagen) and sent for sequencing (Bio Basic Inc. or in-house). Sequences were analyzed using custom software MutantChecker (available at <https://github.com/dmitryveprintsev/AAScan>). Mutants that were not successfully generated as described were cloned using an alternative method: pcDNA3.1 was digested with NheI and XhoI (NEB), the β 2AR coding sequence was amplified in two PCRs, each using one of the mutagenesis primers and either T7long or BGH primer (CGAAATTAATACGACTCACTA-TAGGGAGACCCAAGCTGG and TAGAAGG-CACAGTCGAGG), which anneal upstream and downstream of the coding sequence, respec-

tively. The two fragments were combined, digested using DpnI, and cleaned up as described above. The clean fragments were assembled with digested, cleaned pcDNA3.1 using HiFi DNA assembly Master Mix (NEB).

Cell-culture and BRET-based signaling assays

Adherent HEK-293 SL cells (a gift from S. Laporte) were cultured in Dulbecco's minimum essential medium (DMEM) with 4.5 g/l glucose, L-glutamine, supplemented with 10% newborn calf serum (NCS) (Wisent BioProducts, Canada) and 1x penicillin-streptomycin (PS) (100X, Wisent BioProducts, Canada) at 37°C with 5% CO₂. For transfections, DMEM without phenol red was used. Cells were transiently transfected using linear 25 kDa polyethylenimine (PEI) (Polysciences Inc., Canada, no. 23966) in a 3:1 ratio with DNA. Per condition, a total of 1 μ g DNA was used to transfect 240,000 cells in 1.2 ml in suspension. The DNA mix consisted of 100 ng receptor DNA, combined with 75 ng Gas-RlucII, 200 ng GFP10-G γ 1, and 100 ng G β 1 plasmid DNAs; finally, 525 ng salmon sperm DNA (ssDNA) was added to bring the total amount of DNA to 1000 ng total in a volume of 100 μ l. DNA was then combined with diluted PEI in phosphate-buffered saline (PBS) and added to the cells. Cells were seeded at 20,000 cells per well into Cellstar PS 96-well cell culture plates (Greiner Bio-One, Germany) and incubated for 2 days at 37°C, 5% CO₂. A Fluent780 liquid handler equipped with an MCA 96-well head (Tecan) and a Multidrop Combi (Thermo), each installed in a class II biosafety cabinet were used to aid with transfections.

Before measurement, the medium was removed from the 96-well plates followed by the addition of Tyrode's buffer (137 mM NaCl, 0.9 mM KCl, 1 mM MgCl₂, 11.9 mM NaHCO₃, 3.6 mM NaH₂PO₄, 25 mM Hepes, 5.5 mM glucose, 1 mM CaCl₂, pH 7.4) and incubated for at least 30 min at 37°C. Plates were treated with ligand or vehicle (buffer control) for 10 min before measurement and with coelenterazine 400a (Nanolight Technology) at 5 μ M final 5 min before measurement. Ligand concentrations used were: 31.6 nM (10^{-8.5} M) to 3.16 mM (10^{-3.5} M) in half-log steps, diluted in vehicle, plus vehicle itself. Coelenterazine 400 a was prepared with 1% Pluronic F-127 in Tyrode's buffer to increase solubility. Adrenaline was prepared in 0.01 M HCl to increase stability, the low pH was buffered by Tyrode's buffer upon addition of ligand to the cells. The stability of adrenaline in the above aqueous solution was confirmed by comparing the signaling responses obtained immediately after ligand preparation, after 3 hours, 6 hours 20 min, 8 hours 40 min, and 11 hours. The BRET signal was read using a Synergy Neo (Biotek) equipped with dual photomultiplier tubes (PMTs) (emission: 410 nm

and 515 nm, gain 150 for each PMT and 1.2 s integration time). BRET was measured on an integrated platform with a Biomek NX liquid handler equipped with a 96-well head (Beckman, used for ligand addition), a multi-drop combi (Thermo, luciferase substrate addition), an automated CO₂ incubator Cytomat 6001 (Thermo, plate storage before experiment and during incubation), a barcode reader and Synergy Neo plate reader (Biotek). Ligand dilutions were kept at 4°C with the aid of a Peltier element. All signaling experiments were run in three biological replicates by default. In those cases where very low luciferase counts for multiple plates were observed, indicative of a failing transfection, all transfections and measurements of the day were repeated. All transfections and measurements for the days where an instrument failure led to a premature stop of the robotic system were repeated. BRET data for transfections with very low luciferase counts were completely removed during the data analysis (see manual curation).

All signaling experiments were carried out in at least biological triplicates, wild-type and mock-transfection controls were repeated six times per measurement day.

Cell-surface ELISA

HEK-293 SL cells were transfected using PEI as above and seeded into poly-L-lysine coated 96-well plates (Greiner BioOne, Germany) and incubated for 2 days at 37°C 5% CO₂. Each well was washed with 200 μ l PBS, the cells were then treated with 50 μ l 3% paraformaldehyde per well for 10 min. Each well was washed as follows: Three washes with wash buffer total (PBS + 0.5% BSA); for the last wash step, the cells were incubated with wash buffer for 10 min. Primary rabbit anti-SNAP antibody (GenScript, USA) was added at 0.25 μ g/ml in 50 μ l and incubated for 1 hour at room temperature (RT), followed by three wash steps as described above. The cells were incubated with secondary anti-rabbit HRP antibody (GE Healthcare) diluted 1:1000 for 1 hour. Again, the cells were washed three times with wash buffer and then three times with PBS. Per well, 100 μ l SigmaFast solution was added, the plates were incubated at RT in the dark. Reactions were stopped by the addition of 25 μ l 3M HCl. 100 μ l of the resulting solution was transferred to a new transparent 96-well plate, and the optical density was read at 492 nm using a Tecan GENios Plus microplate reader. Cell-surface ELISAs were carried out in three biological replicates with internal quadruplicates. Each plate contained wild-type and mock-transfected cells as controls.

Analysis of biosensor signal dependence on cell-surface expression

Twelve different amounts of wild-type receptor DNA were combined with constant amounts

of biosensor DNA and transfected into HEK293 cells as described above. The amounts of wild-type receptor DNA were 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 150 and 200 ng. Two days after transfection, the cell-surface expression level and Gs biosensor BRET signal were determined for each transfection. Cell-surface expression in percent of the level obtained for a 100 ng wild-type receptor DNA transfection (the standard used for signaling experiments as described above) was then plotted against Gs biosensor signal amplitude and Gs biosensor $\log EC_{50}$ (fig. S3, A and B). All experiments were repeated in three to four biological replicates.

Computational methods and data analysis

Processing of signaling data

The BRET data were processed using BRET2DTF and DataFitter custom software (Dmitry Vepriyev; <https://github.com/dbv123w/DataFitter>). All biological replicates of concentration-response curves were separately fitted to a Hill equation using a Hill slope of 1 to determine the agonist concentration at the half-maximal signal response ($\log EC_{50}$) as well as the pretransition and posttransition baselines (may also be referred to as top and bottom). All concentration-response curves were visually inspected. We removed complete repeats (i.e., all 12 data points) in those cases where the experiment had obviously failed due to a transfection that did not work (see explanation in “Cell-culture and BRET-based signaling assays” section above) or where data could not be fitted reliably due to noise. In the whole mutant dataset of 1324 concentration-response curves (number excludes controls), 42 single replicate curves were excluded. In addition, replicates for seven mutations showed no measurable signaling (see table S1, curation column set to 2). Data could be fitted starting at ~20% of wild-type amplitude, corresponding to ~0.015 difference between the pre- and posttransition baselines in raw BRET (where raw BRET = GFP/RlucII signal). In total, after these curation steps, quadruplicates for 77 mutants, triplicates for 306 mutants, and duplicates for 29 mutants were retained. No single data points were removed in any case.

The results from data fitting were read into R (RStudio 1.3.959 to 2022.07.1 and R 4.0.0 to 4.2.1) and processed using custom scripts. The following packages were used: tidyverse (especially dplyr, ggplot2, purrr, tibble, tidyr, forcats, and stringr), plotly, MASS, reshape, reshape2, ggrepel, patchwork, ggpubr, bio3d (51), and openxlsx.

The data were read in, reformatted, averaged, annotated (addition of GPCRdb number, expression level, amino acid number, mutation), filtered based on the results of manual curation, normalized, and visualized in R. We checked the dataset for outliers, day-to-day

variation and trends, the effect of expression level on signaling response, and baseline signal variation. We compared the BRET signal for mock transfections where ssDNA was used instead of receptor DNA with mutant pretransition baselines to assess possible changes in constitutive activity. Based on the variability of wild-type pretransition baselines we decided not to interpret changes in mutant pretransition baselines. Although we were not able to measure changes in constitutive activity reliably, this does not mean that none of the mutants affect constitutive activity. Measured mutant amplitudes were normalized to the wild type using the most recently measured wild-type dataset to correct for day-to-day variation and variation during the day, whereas mutant $\log EC_{50}$ values were normalized to the mean $\log EC_{50}$ obtained for the wild type. The normalization did not correct for expression level because the dependence of the signal was minor for receptor expressing to at least 25% of wild-type level (fig. S3, A and B). The measured $\log EC_{50}$ and amplitude values were also used to interpret 22 receptor positions in Hauser *et al.* (15) and Hedderich *et al.* (42). The typical error of an ELISA was much larger than the error of a BRET experiment, at least in our experience; such a correction would therefore have rather lowered the confidence in the results. Cutoffs were applied to the normalized, and curated data, resulting in discretization. Once normalized, the wild-type data were centered around a normalized amplitude of 1 and a normalized $\log EC_{50}$ of 0. Cutoffs were chosen so that none of the 57 wild-type concentration-response curves (measured on different days and batches) were classified as not wild-type-like. For the normalized amplitude, the cutoff was <0.74 (this corresponds to two wild-type standard deviations). For the normalized $\log EC_{50}$, the cutoff was >0.87, this corresponds to a 7.4-fold change.

Our cutoff for a gain of function would be >1.28 for the normalized amplitude and <−0.71 for the normalized $\log EC_{50}$ (cutoffs were applied to nonrounded values; table S1 values are displayed to two decimal precision). Because we were interested in identifying those positions whose mutation impairs signaling, the one mutation that increased efficacy was not explicitly considered in our analyses.

Graphs and illustrations were generated in R (ggplot2), Adobe Illustrator CS6 and CC and Affinity Designer 2. For display in figures, concentration-response curves were fitted in R using the package drc. The raw BRET values were fitted to obtain the BRET value for the pretransition baseline. Subsequently, the data were normalized using the pretransition baseline and the wild-type signal amplitude, resulting in a normalized BRET response where 0% and 100% correspond to the wild-type re-

sponse. For fitting of the normalized data, the LL.4 function was used, the slope was fixed to 1, and the pretransition baseline was fixed to 0.

Residue-residue contacts

Residue-residue contacts were obtained from the Protein Contacts Atlas (<http://pca.mbggroup.bio/index.html>) (52). Two residues are listed as a contact if the distance between the two atoms minus their van der Waals radii equals 0.5 Å or less, corresponding to a maximum distance of ~4.2 Å. Text files downloaded from Protein Contacts Atlas containing residue-residue contacts were read into R and cleaned. We then fortified the tables with additional information [secondary structure elements (SSEs) and GPCRdb numbers for each amino acid], filtered the contact tables to exclude main chain–main chain contacts, and include contacts between two different SSEs only. The contacts were filtered to include only those residues resolved in both crystal structures used for the comparison, 2RH1 and 3SN6. In the inactive and in the active G protein-bound structures of the $\beta 2AR$, 282 and 285 residues form 1275 and 1233 residue-residue contacts, respectively. To focus on residue contacts that contribute to large domain motions instead of local rearrangements during activation, we excluded contacts formed exclusively between backbone atoms and contacts within the same secondary structure element (i.e., contacts formed within a single helix). This reduced the number of residue-residue contacts to 117 contacts that are only present in the inactive state; 61 contacts that are only present in the active, G protein-bound state; and 191 contacts that are present in both the active and inactive states. Contacts are listed in table S2.

Definition of the ligand- and G protein-binding sites

Definitions of the ligand-binding site and the G protein-binding site were based on the structures of $\beta 2AR$ with adrenaline and an engineered nanobody (PDB ID: 4LDO) (25), in complex with heterotrimeric G protein and Gs peptide (3SN6 and 6E67) (14, 37). All residues within 4 Å of adrenaline were classified as part of the ligand-binding site, the residues were verified using LigPlot+ v.2.2 and literature data (26, 53–56). Using PyMOL 2.5.2, all residues within 4 Å of G protein or Gs peptide were determined and classified as residues in the G protein-binding site. Additionally, this residue selection was confirmed by atom-atom contacts from the Protein Contact Atlas. Residues in the ligand and G protein-binding site, respectively, are marked in table S1.

Residue classification

We defined driver and modulator residues as those residues that affect Gs signaling upon

mutation (see above for cutoffs). Additionally, disconnected drivers participate in an active state-specific contact, and connected driver residues form an active state-specific contact to another driver residue. In addition, both connected driver residues need to negatively affect potency, or both need to affect efficacy. That is, if the mutation at one driver residue affected potency and the mutation at the interacting driver residue affected efficacy they were classified as disconnected drivers due to a mismatch in effects. Passenger and bystander residues were defined as those residues that do not affect Gs signaling upon mutation. Additionally, passenger residues must form at least one active state-specific contact. Residues not forming such contacts and not affecting Gs signaling upon mutation were classified as bystanders. Hierarchically, we defined (i) connected drivers as the structurally and pharmacologically most important residues, followed by (ii) disconnected drivers and (iii) modulators that are pharmacologically important but do not mediate active state-specific contact, (iv) passenger residues which mediate active state-specific contact(s) but do not affect pharmacology, and finally (v) bystanders, which are not involved in forming an active state-specific contact and do not affect pharmacology. The allosteric network consisted of 15 residue-residue contacts (33 atom-atom contacts) between 23 residues. Of the 33 atom-atom contacts, 26 were side chain-side chain contacts and 7 were main chain-side chain contacts.

ET analysis

We used ET analysis to identify β 2AR residues with evolutionarily encoded functional roles at the cross section of the network of pharmacologically important residue positions. ET scores were obtained from the Lichtarge Lab ET webserver (<http://evolution.lichtargelab.org/gpqr>). We compared the conservation of amino acids at the level of the adrenergic receptor subfamily ["Adrenoceptor (118)"], amine receptors ["All_amine (547)"], and class A ("ClassA_5105") GPCRs. For each phylogenetic comparison, the lower ET scores at a residue position represent high conservation of specific amino acids at the position within the phylogenetic class. We compared the enrichment of low ET score positions among the residues found in the driver, modulator, passenger, and bystander categories. Groups were compared using a Wilcoxon test (* $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$; ns, not significant).

Extraction of structure features

To calculate the distance between C α atoms of each residue in the inactive versus the active state, we used the highest-resolution inactive-state structure (PDB ID: 2RH1) and the only

available G protein complex structure at the time of performing this analysis (PDB ID: 3SN6) of the β 2AR, aligned the two structures in PyMOL and exported the new coordinates. The PDB files were read into R using the package bio3d, followed by extraction of the x, y, and z coordinates for each structure and the calculation of distance as

distance =

$$\sqrt{(x_{\text{active}} - x_{\text{inactive}})^2 + (y_{\text{active}} - y_{\text{inactive}})^2 + (z_{\text{active}} - z_{\text{inactive}})^2}$$

Likewise, we extracted the dihedral angles for each residue in the above structures using bio3d. The difference in the angles ϕ , ψ , and ω between the active and inactive state was calculated as

$$\Delta\phi = (\phi_{\text{active}} - \phi_{\text{inactive}} + 180) \% 360 - 180$$

Calculation of accessible surface area

Accessible surface area was calculated using dssp (<https://swift.cmbi.umcn.nl/gv/dssp/>) (57, 58), and values are provided in table S1. In addition, we calculated the accessible surface area of the adrenaline-binding site with and without the ligand present using MDTraj (<https://www.mdtraj.org>) and GROMACS (<https://www.gromacs.org/>).

Analysis of SNP data

SNP data for ADRB2 was retrieved from the Genome Aggregation Database (gnomAD, <https://gnomad.broadinstitute.org/>, version gnomAD v2.1.1). The data were analyzed using custom scripts with functions from the tidyverse (especially readr, dplyr, and ggplot2), and the packages janitor, rebus, and stats. Statistical over- and underrepresentation of SNPs in driver, modulator, passenger, and bystander positions were calculated using a hypergeometric test.

REFERENCES AND NOTES

1. M. Audet, M. Bouvier, Restructuring G-protein-coupled receptor activation. *Cell* **151**, 14–23 (2012). doi: [10.1016/j.cell.2012.09.003](https://doi.org/10.1016/j.cell.2012.09.003); pmid: [23021212](https://pubmed.ncbi.nlm.nih.gov/23021212/)
2. C. G. Tate, G. F. X. Schertler, Engineering G protein-coupled receptors to facilitate their structure determination. *Curr. Opin. Struct. Biol.* **19**, 386–395 (2009). doi: [10.1016/j.sbi.2009.07.004](https://doi.org/10.1016/j.sbi.2009.07.004); pmid: [19682887](https://pubmed.ncbi.nlm.nih.gov/19682887/)
3. D. M. Rosenbaum, S. G. F. Rasmussen, B. K. Kobilka, The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363 (2009). doi: [10.1038/nature08144](https://doi.org/10.1038/nature08144); pmid: [19458711](https://pubmed.ncbi.nlm.nih.gov/19458711/)
4. A. J. Venkatakrishnan et al., Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013). doi: [10.1038/nature11896](https://doi.org/10.1038/nature11896); pmid: [23407534](https://pubmed.ncbi.nlm.nih.gov/23407534/)
5. G. Thiel, A. Kaufmann, O. G. Rössler, G-protein-coupled designer receptors – New chemical-genetic tools for signal transduction research. *Biol. Chem.* **394**, 1615–1622 (2013). doi: [10.1515/hsz-2013-0164](https://doi.org/10.1515/hsz-2013-0164); pmid: [23893685](https://pubmed.ncbi.nlm.nih.gov/23893685/)
6. J. F. Armstrong et al., The IUPHAR/BPS Guide to PHARMACOLOGY in 2020: Extending immunopharmacology content and introducing the IUPHAR/MMV Guide to MALARIA PHARMACOLOGY. *Nucleic Acids Res.* **48**, D1006–D1021 (2020). doi: [10.1093/nar/gkz951](https://doi.org/10.1093/nar/gkz951); pmid: [31691834](https://pubmed.ncbi.nlm.nih.gov/31691834/)

7. M. J. Marinissen, J. S. Gutkind, G-protein-coupled receptors and signaling networks: Emerging paradigms. *Trends Pharmacol. Sci.* **22**, 368–376 (2001). doi: [10.1016/S0165-6147\(00\)01678-3](https://doi.org/10.1016/S0165-6147(00)01678-3); pmid: [11431032](https://pubmed.ncbi.nlm.nih.gov/11431032/)
8. T. Schöneberg, I. Liebscher, Mutations in G Protein-Coupled Receptors: Mechanisms, Pathophysiology and Potential Therapeutic Approaches. *Pharmacol. Rev.* **73**, 89–119 (2021). doi: [10.1124/pharmrev.120.000011](https://doi.org/10.1124/pharmrev.120.000011); pmid: [33219147](https://pubmed.ncbi.nlm.nih.gov/33219147/)
9. 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). doi: [10.1038/nrd.2017.178](https://doi.org/10.1038/nrd.2017.178); pmid: [29075003](https://pubmed.ncbi.nlm.nih.gov/29075003/)
10. L. M. Winkler et al., Angiotensin Analogs with Divergent Bias Stabilize Distinct Receptor Conformations. *Cell* **176**, 468–478.e11 (2019). doi: [10.1016/j.cell.2018.12.005](https://doi.org/10.1016/j.cell.2018.12.005); pmid: [30639099](https://pubmed.ncbi.nlm.nih.gov/30639099/)
11. X. Deupi, B. Kobilka, Activation of G protein-coupled receptors. *Adv. Protein Chem.* **74**, 137–166 (2007). doi: [10.1016/S0065-3233\(07\)74004-4](https://doi.org/10.1016/S0065-3233(07)74004-4); pmid: [17854657](https://pubmed.ncbi.nlm.nih.gov/17854657/)
12. M. Audet, M. Bouvier, Insights into signaling from the β_2 -adrenergic receptor structure. *Nat. Chem. Biol.* **4**, 397–403 (2008). doi: [10.1038/nchembio.97](https://doi.org/10.1038/nchembio.97); pmid: [18560432](https://pubmed.ncbi.nlm.nih.gov/18560432/)
13. R. R. Neubig, M. Spedding, T. Kenakin, A. Christopoulos, International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. *Pharmacol. Rev.* **55**, 597–606 (2003). doi: [10.1124/pr.55.4.4](https://doi.org/10.1124/pr.55.4.4); pmid: [14657418](https://pubmed.ncbi.nlm.nih.gov/14657418/)
14. S. G. Rasmussen et al., Crystal structure of the β_2 adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011). doi: [10.1038/nature10361](https://doi.org/10.1038/nature10361); pmid: [21772288](https://pubmed.ncbi.nlm.nih.gov/21772288/)
15. A. S. Hauser et al., GPCR activation mechanisms across classes and macro/microstates. *Nat. Struct. Mol. Biol.* **28**, 879–888 (2021). doi: [10.1038/s41594-021-00674-7](https://doi.org/10.1038/s41594-021-00674-7); pmid: [34759375](https://pubmed.ncbi.nlm.nih.gov/34759375/)
16. G. G. Gregorio et al., Single-molecule analysis of ligand efficacy in β_2 AR–G-protein activation. *Nature* **547**, 68–73 (2017). doi: [10.1038/nature22354](https://doi.org/10.1038/nature22354); pmid: [28607487](https://pubmed.ncbi.nlm.nih.gov/28607487/)
17. T. Warne, P. C. Edwards, A. S. Doré, A. G. W. Leslie, C. G. Tate, Molecular basis for high-affinity agonist binding in GPCRs. *Science* **364**, 775–778 (2019). doi: [10.1126/science.aau5595](https://doi.org/10.1126/science.aau5595); pmid: [31072904](https://pubmed.ncbi.nlm.nih.gov/31072904/)
18. E. M. Jones et al., Structural and functional characterization of G protein-coupled receptors with deep mutational scanning. *eLife* **9**, e54895 (2020). doi: [10.7554/eLife.54895](https://doi.org/10.7554/eLife.54895); pmid: [33084570](https://pubmed.ncbi.nlm.nih.gov/33084570/)
19. M. P. Wescott et al., Signal transmission through the CXCR4 chemokine receptor 4 (CXCR4) transmembrane helices. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9928–9933 (2016). doi: [10.1073/pnas.1601278113](https://doi.org/10.1073/pnas.1601278113); pmid: [27543332](https://pubmed.ncbi.nlm.nih.gov/27543332/)
20. M. Eilers, A. B. Patel, W. Liu, S. O. Smith, Comparison of helix interactions in membrane and soluble α -bundle proteins. *Biophys. J.* **82**, 2720–2736 (2002). doi: [10.1016/S0006-3495\(02\)75613-0](https://doi.org/10.1016/S0006-3495(02)75613-0); pmid: [11964258](https://pubmed.ncbi.nlm.nih.gov/11964258/)
21. C. Galés et al., Probing the activation-promoted structural rearrangements in preassembled receptor–G protein complexes. *Nat. Struct. Mol. Biol.* **13**, 778–786 (2006). doi: [10.1038/nsmb1134](https://doi.org/10.1038/nsmb1134); pmid: [16906158](https://pubmed.ncbi.nlm.nih.gov/16906158/)
22. C. Galés et al., Real-time monitoring of receptor and G-protein interactions in living cells. *Met. Methods* **2**, 177–184 (2005). doi: [10.1038/nmeth743](https://doi.org/10.1038/nmeth743); pmid: [15782186](https://pubmed.ncbi.nlm.nih.gov/15782186/)
23. V. Isberg et al., Generic GPCR residue numbers – Aligning topology maps while minding the gaps. *Trends Pharmacol. Sci.* **36**, 22–31 (2015). doi: [10.1016/j.tips.2014.11.001](https://doi.org/10.1016/j.tips.2014.11.001); pmid: [25541108](https://pubmed.ncbi.nlm.nih.gov/25541108/)
24. J. A. Ballesteros, H. Weinstein, in *Methods in Neurosciences*, S. C. Sealfon, Ed. (Academic Press, 1995), pp. 366–428.
25. A. M. Ring et al., Adrenaline-activated structure of β_2 -adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575–579 (2013). doi: [10.1038/nature12572](https://doi.org/10.1038/nature12572); pmid: [24056936](https://pubmed.ncbi.nlm.nih.gov/24056936/)
26. C. D. Strader et al., Conserved aspartic acid residues 79 and 113 of the β -adrenergic receptor have different roles in receptor function. *J. Biol. Chem.* **263**, 10267–10271 (1988). doi: [10.1016/S0021-9258\(19\)81509-0](https://doi.org/10.1016/S0021-9258(19)81509-0); pmid: [2899076](https://pubmed.ncbi.nlm.nih.gov/2899076/)
27. C. D. Strader, I. S. Sigal, R. A. Dixon, Structural basis of β -adrenergic receptor function. *FASEB J.* **3**, 1825–1832 (1989). doi: [10.1096/fasebj.3.7.2541037](https://doi.org/10.1096/fasebj.3.7.2541037); pmid: [2541037](https://pubmed.ncbi.nlm.nih.gov/2541037/)
28. T. Flock et al., Universal allosteric mechanism for G_{α} activation by GPCRs. *Nature* **524**, 173–179 (2015). doi: [10.1038/nature14663](https://doi.org/10.1038/nature14663); pmid: [26147082](https://pubmed.ncbi.nlm.nih.gov/26147082/)
29. D. Sun et al., Probing G_{α_i} protein activation at single-amino acid resolution. *Nat. Struct. Mol. Biol.* **22**, 686–694 (2015). doi: [10.1038/nsmb.3070](https://doi.org/10.1038/nsmb.3070); pmid: [26258638](https://pubmed.ncbi.nlm.nih.gov/26258638/)

30. T. Flock *et al.*, Selectivity determinants of GPCR-G-protein binding. *Nature* **545**, 317–322 (2017). doi: [10.1038/nature22070](https://doi.org/10.1038/nature22070); pmid: [28489817](https://pubmed.ncbi.nlm.nih.gov/28489817/)
31. M. Sandhu *et al.*, Dynamic spatiotemporal determinants modulate GPCR:G protein coupling selectivity and promiscuity. *Nat. Commun.* **13**, 7428 (2022). doi: [10.1038/s41467-022-34055-5](https://doi.org/10.1038/s41467-022-34055-5); pmid: [36460632](https://pubmed.ncbi.nlm.nih.gov/36460632/)
32. A. Manglik *et al.*, Structural Insights into the Dynamic Process of β_2 -Adrenergic Receptor Signaling. *Cell* **161**, 1101–1111 (2015). doi: [10.1016/j.cell.2015.04.043](https://doi.org/10.1016/j.cell.2015.04.043); pmid: [25981665](https://pubmed.ncbi.nlm.nih.gov/25981665/)
33. G. Swaminath *et al.*, Sequential binding of agonists to the β_2 adrenoceptor: Kinetic evidence for intermediate conformational states. *J. Biol. Chem.* **279**, 686–691 (2004). doi: [10.1074/jbc.M310888200](https://doi.org/10.1074/jbc.M310888200); pmid: [14559905](https://pubmed.ncbi.nlm.nih.gov/14559905/)
34. S. Chen *et al.*, Mutation of a single TMVI residue, Phe²⁸², in the β_2 -adrenergic receptor results in structurally distinct activated receptor conformations. *Biochemistry* **41**, 6045–6053 (2002). doi: [10.1021/bi012189c](https://doi.org/10.1021/bi012189c); pmid: [11993999](https://pubmed.ncbi.nlm.nih.gov/11993999/)
35. A.-M. Schönege *et al.*, Evolutionary action and structural basis of the allosteric switch controlling β_2 AR functional selectivity. *Nat. Commun.* **8**, 2169 (2017). doi: [10.1038/s41467-017-02257-x](https://doi.org/10.1038/s41467-017-02257-x); pmid: [29255305](https://pubmed.ncbi.nlm.nih.gov/29255305/)
36. V. Katritch *et al.*, Allosteric sodium in class A GPCR signaling. *Trends Biochem. Sci.* **39**, 233–244 (2014). doi: [10.1016/j.tibs.2014.03.002](https://doi.org/10.1016/j.tibs.2014.03.002); pmid: [24767681](https://pubmed.ncbi.nlm.nih.gov/24767681/)
37. X. Liu *et al.*, Structural Insights into the Process of GPCR-G Protein Complex Formation. *Cell* **177**, 1243–1251.e12 (2019). doi: [10.1016/j.cell.2019.04.021](https://doi.org/10.1016/j.cell.2019.04.021); pmid: [31080070](https://pubmed.ncbi.nlm.nih.gov/31080070/)
38. A. J. Venkatakrishnan *et al.*, Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature* **536**, 484–487 (2016). doi: [10.1038/nature19107](https://doi.org/10.1038/nature19107); pmid: [27525504](https://pubmed.ncbi.nlm.nih.gov/27525504/)
39. A. J. Venkatakrishnan *et al.*, Diverse GPCRs exhibit conserved water networks for stabilization and activation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3288–3293 (2019). doi: [10.1073/pnas.1809251116](https://doi.org/10.1073/pnas.1809251116); pmid: [30728297](https://pubmed.ncbi.nlm.nih.gov/30728297/)
40. N. R. Latorraca, A. J. Venkatakrishnan, R. O. Dror, GPCR Dynamics: Structures in Motion. *Chem. Rev.* **117**, 139–155 (2017). doi: [10.1021/acs.chemrev.6b00177](https://doi.org/10.1021/acs.chemrev.6b00177); pmid: [27622975](https://pubmed.ncbi.nlm.nih.gov/27622975/)
41. D. M. Thal, A. Glukhova, P. M. Sexton, A. Christopoulos, Structural insights into G-protein-coupled receptor allostery. *Nature* **559**, 45–53 (2018). doi: [10.1038/s41586-018-0259-z](https://doi.org/10.1038/s41586-018-0259-z); pmid: [29973731](https://pubmed.ncbi.nlm.nih.gov/29973731/)
42. J. B. Hedderich *et al.*, The pocketome of G-protein-coupled receptors reveals previously untargated allosteric sites. *Nat. Commun.* **13**, 2567 (2022). doi: [10.1038/s41467-022-29609-6](https://doi.org/10.1038/s41467-022-29609-6); pmid: [35538063](https://pubmed.ncbi.nlm.nih.gov/35538063/)
43. K. J. Karczewski *et al.*, The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020). doi: [10.1038/s41586-020-2308-7](https://doi.org/10.1038/s41586-020-2308-7); pmid: [32461654](https://pubmed.ncbi.nlm.nih.gov/32461654/)
44. S. Madabushi *et al.*, Evolutionary trace of G protein-coupled receptors reveals clusters of residues that determine global and class-specific functions. *J. Biol. Chem.* **279**, 8126–8132 (2004). doi: [10.1074/jbc.M312671200](https://doi.org/10.1074/jbc.M312671200); pmid: [14660595](https://pubmed.ncbi.nlm.nih.gov/14660595/)
45. O. Lichtarge, H. R. Bourne, F. E. Cohen, An evolutionary trace method defines binding surfaces common to protein families. *J. Mol. Biol.* **257**, 342–358 (1996). doi: [10.1006/jmbi.1996.0167](https://doi.org/10.1006/jmbi.1996.0167); pmid: [8609628](https://pubmed.ncbi.nlm.nih.gov/8609628/)
46. B. Selçuk, I. Erol, S. Durdagi, O. Adebali, Evolutionary association of receptor-wide amino acids with G protein-coupling selectivity in aminergic GPCRs. *Life Sci. Alliance* **5**, e202201439 (2022). doi: [10.26508/lsa.202201439](https://doi.org/10.26508/lsa.202201439); pmid: [35613896](https://pubmed.ncbi.nlm.nih.gov/35613896/)
47. K. M. Chen, D. Keri, P. Barth, Computational design of G Protein-Coupled Receptor allosteric signal transductions. *Nat. Chem. Biol.* **16**, 77–86 (2020). doi: [10.1038/s41589-019-0407-2](https://doi.org/10.1038/s41589-019-0407-2); pmid: [31792443](https://pubmed.ncbi.nlm.nih.gov/31792443/)
48. F. M. Heydenreich *et al.*, High-throughput mutagenesis using a two-fragment PCR approach. *Sci. Rep.* **7**, 6787 (2017). doi: [10.1038/s41598-017-07010-4](https://doi.org/10.1038/s41598-017-07010-4); pmid: [28754896](https://pubmed.ncbi.nlm.nih.gov/28754896/)
49. D. Sun *et al.*, AAscan, PCRdesign and MutantChecker: A suite of programs for primer design and sequence analysis for high-throughput scanning mutagenesis. *PLOS ONE* **8**, e78878 (2013). doi: [10.1371/journal.pone.0078878](https://doi.org/10.1371/journal.pone.0078878); pmid: [24205336](https://pubmed.ncbi.nlm.nih.gov/24205336/)
50. B. Paetzold, C. Carolis, T. Ferrar, L. Serrano, M. Lluch-Senar, *In situ* overlap and sequence synthesis during DNA assembly. *ACS Synth. Biol.* **2**, 750–755 (2013). doi: [10.1021/sb400067v](https://doi.org/10.1021/sb400067v); pmid: [24161008](https://pubmed.ncbi.nlm.nih.gov/24161008/)
51. B. J. Grant, A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon, L. S. D. Caves, Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006). doi: [10.1093/bioinformatics/btl461](https://doi.org/10.1093/bioinformatics/btl461); pmid: [16940322](https://pubmed.ncbi.nlm.nih.gov/16940322/)
52. M. Kayikci *et al.*, Visualization and analysis of non-covalent contacts using the Protein Contacts Atlas. *Nat. Struct. Mol. Biol.* **25**, 185–194 (2018). doi: [10.1038/s41594-017-0019-z](https://doi.org/10.1038/s41594-017-0019-z); pmid: [29335563](https://pubmed.ncbi.nlm.nih.gov/29335563/)
53. C. D. Strader *et al.*, Identification of residues required for ligand binding to the beta-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4384–4388 (1987). doi: [10.1073/pnas.84.13.4384](https://doi.org/10.1073/pnas.84.13.4384); pmid: [2885836](https://pubmed.ncbi.nlm.nih.gov/2885836/)
54. C. Ambrosio, P. Molinari, S. Cotecchia, T. Costa, Catechol-binding serines of β_2 -adrenergic receptors control the equilibrium between active and inactive receptor states. *Mol. Pharmacol.* **57**, 198–210 (2000). pmid: [10617695](https://pubmed.ncbi.nlm.nih.gov/10617695/)
55. R. Del Carmine, P. Molinari, M. Sbraccia, C. Ambrosio, T. Costa, “Induced-fit” mechanism for catecholamine binding to the β_2 -adrenergic receptor. *Mol. Pharmacol.* **66**, 356–363 (2004). doi: [10.1124/mol.66.2.356](https://doi.org/10.1124/mol.66.2.356); pmid: [15266027](https://pubmed.ncbi.nlm.nih.gov/15266027/)
56. G. Liapakis *et al.*, The forgotten serine: A critical role for Ser-203^{5,42} in ligand binding to and activation of the β_2 -adrenergic receptor. *J. Biol. Chem.* **275**, 37779–37788 (2000). doi: [10.1074/jbc.M002092200](https://doi.org/10.1074/jbc.M002092200); pmid: [10964911](https://pubmed.ncbi.nlm.nih.gov/10964911/)
57. W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983). doi: [10.1002/bip.360221211](https://doi.org/10.1002/bip.360221211); pmid: [6667333](https://pubmed.ncbi.nlm.nih.gov/6667333/)
58. W. G. Touw *et al.*, A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* **43**, D364–D368 (2015). doi: [10.1093/nar/gku1028](https://doi.org/10.1093/nar/gku1028); pmid: [25352545](https://pubmed.ncbi.nlm.nih.gov/25352545/)
59. J. Patel, F. Heydenreich, mb-group/DeterminantsSignalingGPCR: DeterminantsSignalingGPCR, version v1.0, Zenodo (2023); <https://doi.org/10.5281/zenodo.8349361>

ACKNOWLEDGMENTS

We thank I. Chen, N. Gough, K. Nemec, E. Park, N. Casajuna-Martin, B. Plouffe, and J. Janetzko for their constructive comments on the manuscript as well as D. Malinverni for discussions. We thank B. Sejdiu for help with ASA calculations and J. Patel for help with the GitHub repository. We thank members of the Koblika and Babu groups for constructive feedback on this work. **Funding:** This study was supported by Marie Skłodowska-Curie Individual Fellowships from the European Union's Horizon 2020 research and innovation program under grant agreement nos. 844622 and 832620 (F.M.H. and M.M.-S.); the American Lebanese Syrian Associated Charities (ALSAC) (M.M.B. and M.S.); American Heart Association grant no. 19POST34380839 (F.M.H.); National Institutes of Health grant R01NS028471 (B.K.K.); Canadian Institute of Health Research grant FDN no. 148431 (M.B.); UK Medical Research Council grant no. MC_U105185859 (M.M.B. and M.M.-S.); Royal Society University Research Fellowship URF\R1\221205 (M.M.-S.); Isaac Newton Trust grants [22.23(d)] (M.M.-S.); and the Wellcome Trust Institutional Strategic Support Fund (M.M.-S.). M.B. holds a Canada Research Chair in Signal Transduction and Molecular Pharmacology, B.K.K. is a Chan Zuckerberg Biohub Investigator. M.M.B. holds the St. Jude Endowed Chair in Biological Data Science. **Author contributions:** Conceptualization: F.M.H., B.K.K., M.B., and M.M.B. Methodology: F.M.H., M.B., M.M.B., and B.K.K. Formal analysis: F.M.H. Investigation: F.M.H., M.M.-S., and M.S. Project administration: F.M.H. Visualization: F.M.H. Software: F.M.H. Funding acquisition: F.M.H., M.M.-S., B.K.K., M.B., and M.M.B. Supervision: B.K.K., M.B., and M.M.B. Writing – original draft: F.M.H., M.M.-S., M.S., and M.M.B. Writing – review & editing: F.M.H., M.M.-S., M.S., B.K.K., M.B., and M.M.B. **Competing interests:** M.B. is the chairman of the scientific advisory board of Domain Therapeutics, a biotech company to which the BRET-based sensor used in this study was licensed for commercial use. B.K.K. is a cofounder and consultant for ConfoMetRx. The authors declare no other competing interests. **Data and materials availability:** The datasets generated during and/or analyzed during the current study are available in table S1 and table S2. The sequences of the primers used for mutagenesis are listed in table S3. All data are available in the main text or the supplementary materials; code can be accessed at Zenodo (59) or Github (<https://github.com/mb-group/DeterminantsSignalingGPCR>). Constructs from the alanine scanning library are available upon request to F.M.H. The BRET-based biosensor is available upon request to M.B. for academic, noncommercial studies through regular material transfer agreements. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh1859

Figs. S1 to S13

Tables S1 to S3

MDAR Reproducibility Checklist

Submitted 12 April 2023; accepted 7 November 2023

10.1126/science.adh1859