LETTER

doi:10.1038/nature09665

Structure and function of an irreversible agonist–β_2 adrenoceptor complex

Daniel M. Rosenbaum[1]†, Cheng Zhang[1], Joseph A. Lyons[2,3], Ralph Holl[4], David Aragao[3], Daniel H. Arlow[5], Søren G. F. Rasmussen[1], Hee-Jung Choi[1,6], Brian T. DeVree[7], Roger K. Sunahara[7], Pil Seok Chae[8], Samuel H. Gellman[8], Ron O. Dror[5], David E. Shaw[5], William I. Weis[1,6], Martin Caffrey[3], Peter Gmeiner[4] & Brian K. Kobilka[1]

G-protein-coupled receptors (GPCRs) are eukaryotic integral membrane proteins that modulate biological function by initiating cellular signalling in response to chemically diverse agonists. Despite recent progress in the structural biology of GPCRs[1], the molecular basis for agonist binding and allosteric modulation of these proteins is poorly understood. Structural knowledge of agonist-bound states is essential for deciphering the mechanism of receptor activation, and for structure-guided design and optimization of ligands. However, the crystallization of agonist-bound GPCRs has been hampered by modest affinities and rapid off-rates of available agonists. Using the inactive structure of the human β_2 adrenergic receptor (β_2AR) as a guide, we designed a β_2AR agonist that can be covalently tethered to a specific site on the receptor through a disulphide bond. The covalent β_2AR-agonist complex forms efficiently, and is capable of activating a heterotrimeric G protein. We crystallized a covalent agonist-bound β_2AR–T4L fusion protein in lipid bilayers through the use of the lipidic mesophase method[2], and determined its structure at 3.5 Å resolution. A comparison to the inactive structure and an antibody-stabilized active structure (companion paper[3]) shows how binding events at both the extracellular and intracellular surfaces are required to stabilize an active conformation of the receptor. The structures are in agreement with long-timescale (up to 30 μs) molecular dynamics simulations showing that an agonist-bound active conformation spontaneously relaxes to an inactive-like conformation in the absence of a G protein or stabilizing antibody.

The relationship between agonist binding to GPCRs and the conformational changes that facilitate G protein binding and activation remains largely unknown. Characterization of GPCR activation at a molecular level has been driven by a combination of X-ray crystal structure analysis and spectroscopic approaches. Rhodopsin has served as a prototype GPCR for biophysical studies, given its ready availability and superior stability. Crystal structures have been obtained for the inactive dark state with inverse agonist 11-*cis*-retinal covalently bound[4,5], as well as the active state mimetic low-pH opsin lacking the retinal chromophore[6,7]. Opsin differs from rhodopsin by outward rigid-body movements of transmembrane helices (TMs) 5 and 6 at the cytoplasmic G protein binding surface. For other GPCRs, which respond to diffusible ligands, structural information has proven more difficult to obtain. Fluorescence spectroscopy studies show that activation of the β_2AR by diffusible ligands can follow multiple pathways, with a complex energy landscape of receptor conformations[8]. This multitude of accessible conformations has probably contributed to the difficulty in obtaining crystal structures of non-rhodopsin GPCRs.

Inactive state crystal structures of the β_2AR, β_1AR, and A_{2A} adenosine receptor have been solved over the past several years with the aid of protein engineering techniques[1], but agonist-bound structures for these proteins have yet to be reported. Like many GPCRs, the β_2AR has two agonist affinity states: a low affinity state in the absence of cognate G protein, and a high affinity state in the presence of G protein. This observation indicates that agonists can bind to two distinct receptor conformations. A more complete understanding of the processes of agonist binding and activation requires structures of both high and low affinity states. The high affinity state is challenging because a receptor G protein complex is unstable in detergent solutions required for purification of both GPCRs and heterotrimeric G proteins. In a companion paper[3], we describe the use of a conformationally selective camelid antibody (nanobody, Nb80) with G-protein-like properties to obtain a structure of a high-affinity agonist-bound conformation. Obtaining a structure of the low affinity state is also challenging because of the relatively rapid association and dissociation rates of commercial β_2AR agonists. Inspired by the covalent retinal-rhodopsin system, we proposed that the ability to crystallize an agonist-bound GPCR would be enhanced by chemically crosslinking ligand and receptor, preferably in a manner that would not inhibit conformational freedom and the capacity to activate a G protein.

Our design strategy for a covalent β_2AR agonist was to combine a β-adrenergic agonist core (procaterol) and a reactive chemical group that could be targeted to a specific residue on the receptor. Using the structure of the carazolol-bound β_2AR as a template (Fig. 1a), a flexible linker was added to bridge these two components such that the covalent attachment would not inhibit binding of the agonist core or conformational flexibility of the transmembrane helices. Biochemical precedent for this strategy came from the covalent labelling reagent BABC (Fig. 1b), in which an electrophilic group appended to the carazolol ligand core was determined to react with $His\,93^{2.64}$ at the extracellular end of TM2 (ref. 9; Ballesteros–Weinstein numbering[10] used in superscript). For crosslinking, we chose the reaction between a free cysteine on the receptor (introduced at position 93) and a ligand disulphide moiety, based on the mild and proximity-dependent 'tethering' approach that has proven broadly applicable to different protein targets including GPCRs[11]. The designed covalent β agonist FAUC50 (Fig. 1b) was synthesized in enantiomerically pure form[12,13], along with the noncovalent analogue FAUC72 (Fig. 1b, Supplementary Fig. 1 and Supplementary Information) for use in control experiments.

Incubation of compound FAUC50 with mutant H93C receptor led to efficient and irreversible blocking of radioligand binding (Supplementary Fig. 2, Supplementary Information). We sought to determine whether the tethered FAUC50–β_2AR^{H93C} complex is capable of activating a G protein. Wild-type and mutant receptor were

[1]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, 279 Campus Drive, Stanford, California 94305, USA. [2]Department of Chemical and Environmental Sciences, University of Limerick, Limerick, Ireland. [3]Membrane Structural and Functional Biology Group, School of Biochemistry and Immunology, Trinity College, Dublin 2, Ireland. [4]Department of Chemistry and Pharmacy, Friedrich Alexander University, Schuhstrasse 19, 91052 Erlangan, Germany. [5]D. E. Shaw Research, New York, New York 10036, USA. [6]Department of Structural Biology, Stanford University School of Medicine, 299 Campus Drive, Stanford, California 94305, USA. [7]Department of Pharmacology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. [8]Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706, USA. †Present address: Department of Biochemistry, University of Texas Southwestern, 5223 Harry Hines Blvd, Dallas, Texas 75390-8816, USA.

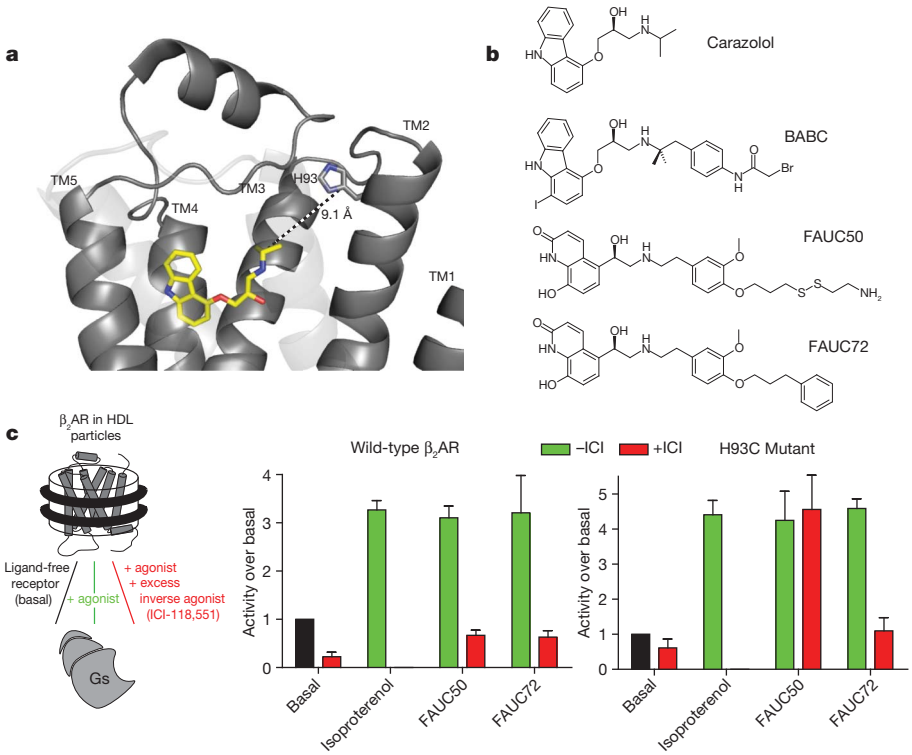

Figure 1 | Design and function of a covalent agonist. **a**, Structure of the carazolol-bound β_2AR, receptor in grey cartoon and ligand in yellow sticks, showing distance between isopropyl group and His $93^{2.64}$ imidazole. **b**, Structures of carazolol and the related BABC ligand, covalent ligand FAUC50 and noncovalent analogue FAUC72. **c**, G protein activation assay demonstrating that covalently bound FAUC50 activates the β_2AR. Error bars represent the standard deviation from three independent experiments.

reconstituted into high density lipoprotein (HDL) particles[14], and then incubated with ligand alone or ligand followed by an excess of the high-affinity inverse agonist ICI-118,551. Heterotrimeric Gs protein was then added to the particles, and activation was observed by measuring GTPγ^{35}S binding to the Gs α subunit. Figure 1c shows that the inverse agonist treatment prevented agonist-induced G protein activation by the wild-type receptor. However, excess ICI-118,551 is unable to reverse FAUC50-induced coupling in the case of β_2AR^{H93C}. The noncovalent analogue FAUC72 is displaced by the inverse agonist for both wild-type and H93C mutant receptors. From these experiments we conclude that FAUC50 not only reacts efficiently with Cys 93 to form a covalent complex, but this complex is also capable of activating a G protein.

We were motivated to develop a covalent agonist by our inability to produce diffraction-quality crystals of the previously described β_2AR–T4 lysozyme chimaera (β_2AR–T4L) in complex with available noncovalent agonists. In contrast, the purified FAUC50–β_2AR^{H93C}T4L complex readily yielded diffraction-quality crystals in lipidic mesophases[15]. We overexpressed the mutant receptor in Sf9 insect cells, purified the protein to homogeneity using immunoaffinity and ligand-affinity chromatography, and introduced the FAUC50 ligand during a subsequent chromatography step. Using a cholesterol-doped monoolein cubic phase and robotic *in meso* technology[16], we obtained $50 \times 15 \times 5$ μm blade-shaped crystals (Supplementary Fig. 3) that diffracted to 3.5 Å. Combining diffraction data from 19 microcrystals (Supplementary Table 1), we solved the agonist-bound structure by molecular replacement using the coordinates of carazolol-bound β_2AR–T4L[17] as a search model. The packing of protomers in the orthorhombic crystals is distinct from previous crystals of β_2AR–T4L fusion proteins. However, the orientation of T4L relative to the receptor closely resembles the original carazolol-bound structure (PDB ID 2RH1). An omit map at the ligand binding site reveals clear electron density for the agonist (Supplementary Fig. 4), and the refined structure shows prominent features expected from adrenergic receptor pharmacology[18]: the amine and β-hydroxyl of FAUC50 contact Asp $113^{3.32}$ and Asn $312^{7.39}$ in a manner similar to inverse agonists[17], whereas the aromatic moiety of FAUC50 that replaces the catechol ring forms hydrogen bonds with Ser $203^{5.42}$ and Ser $207^{5.46}$ (Fig. 2a, left).

The crystal structure of the covalent agonist-bound receptor in a lipidic mesophase does not show the conformational changes near the cytoplasmic surface associated with G protein binding, as observed in the nanobody-stabilized agonist complex (Fig. 2b). This was surprising given that the covalent agonist-bound receptor activates Gs (Fig. 1c) and the T4L fusion does not interfere with agonist-induced conformational changes detected with a fluorescent probe at the end of TM6 (ref. 17). This result implies that even an agonist with zero dissociation inefficiently stabilizes an active conformation of the β_2AR in a lipid bilayer environment, consistent with a higher stability of the inactive carazolol-bound conformation. Nevertheless, allosteric communication between the ligand-binding pocket and the cytoplasmic G-protein-binding surface is a fundamental feature of GPCRs, exemplified by agonist-induced G protein stimulation and G-protein-induced high-affinity agonist binding[14]. For the β_2AR, we can now compare an agonist-bound structure to that of a nanobody-stabilized active state mimetic (see companion paper[3]), to better understand conformational changes associated with activation. The superposition in Fig. 2b shows that agonist binding alone is insufficient to stabilize an active conformation at the cytoplasmic surface, with the requisite outward movement of TM5 and TM6. The largest differences proximal to the ligand-binding pocket in the nanobody-stabilized active conformation involve movement of Ile $121^{3.40}$ away from Pro $211^{5.50}$ and into space occupied by Phe $282^{6.44}$ in the inactive state (Fig. 2c, middle). The concomitant outward movement of Phe $282^{6.44}$ is accompanied by subtle backbone torsion changes distributed through TM6 that cause the cytoplasmic half of TM6 to be redirected outward, similar to its orientation in activated opsin[6,7]. In the binding pocket, the most significant difference between the two agonist-bound structures and the carazolol-bound β_2AR are in the hydrogen bonding contacts with

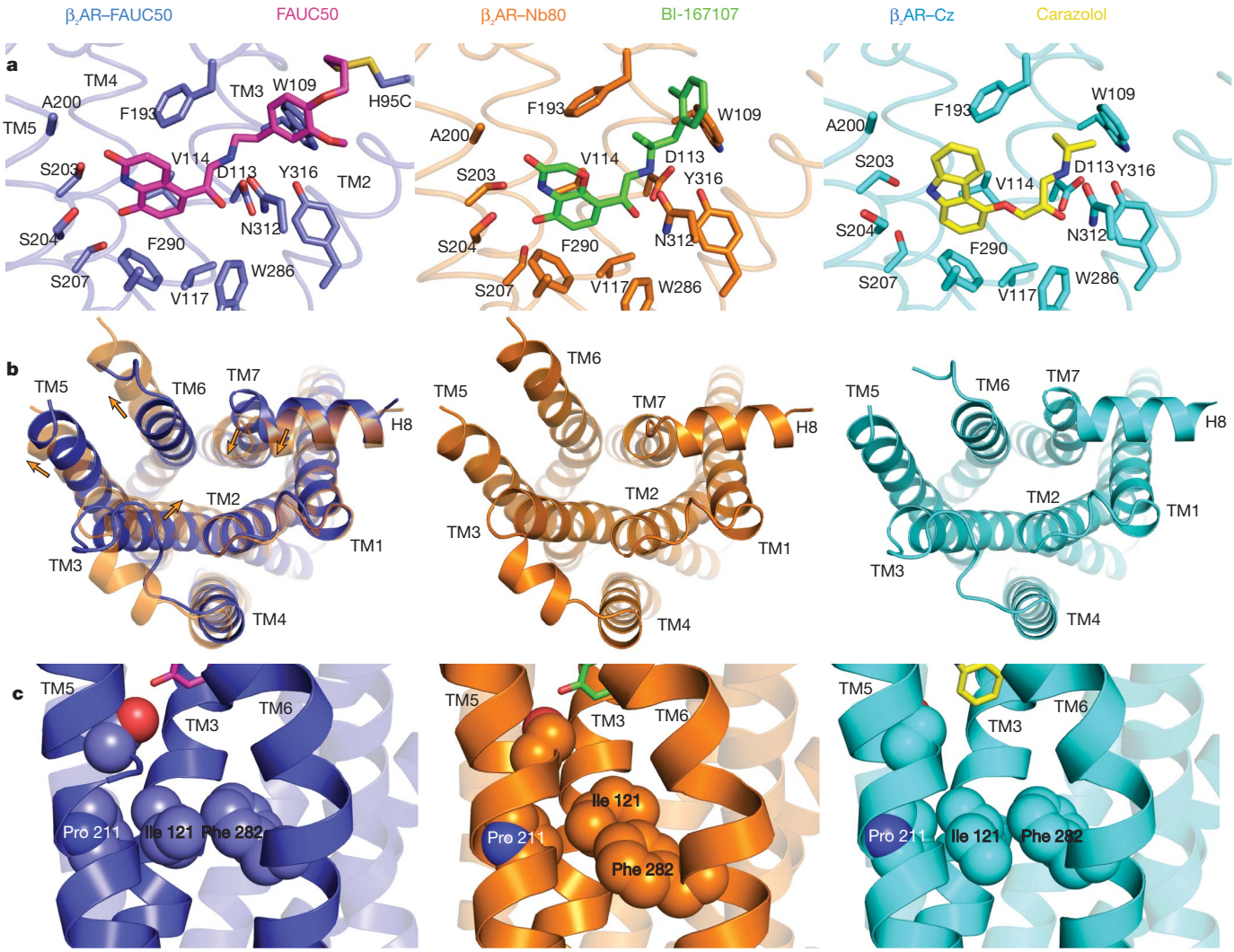

Figure 2 | Comparison of agonist and inverse agonist bound β_2AR structures. Comparison of the covalent FAUC50-bound β_2AR^{H93C}T4L (left panels, blue cartoon, ligand carbons in purple), BI-167107-bound β_2AR–T4L/nanobody (β_2AR–Nb80) complex (middle panels, orange cartoon, ligand carbons in green), and carazolol-bound β_2AR–T4L (right panels, cyan cartoon, ligand carbons in yellow). **a**, Hormone binding site with interactions between ligands and receptors. TMs 6 and 7 and residues Phe 289, Asn 293 and Tyr 308 are omitted for clarity. **b**, Comparison of the cytoplasmic faces showing differences in TMs 5 and 6. Superimposed β_2AR–Nb80 complex is also shown in left panel as a transparent cartoon, and arrows indicate rigid body movements. **c**, Conformational switch region with residues Ile $121^{3.40}$, Pro $211^{5.50}$ and Phe $282^{6.44}$ from TMs 3, 5 and 6 (other TMs transparent). Side chains are shown in van der Waals sphere representation.

Ser $203^{5.42}$ and Ser $207^{5.46}$ on TM5 (Fig. 2a). However, only in the nanobody-stabilized structure are these differences coordinated with further changes towards the cytoplasmic surface of the molecule (Fig. 2b and c, middle).

To determine whether an agonist-bound receptor would sustain a stable active conformation in the absence of a cytoplasmic binding partner, we initiated an unbiased molecular dynamics (MD) simulation of the receptor from the nanobody complex structure, but with the nanobody removed (Fig. 3a). In the first several microseconds of simulated time, the intracellular ends of TM5 and TM6 showed high mobility (Supplementary Fig. 5), drifting by as much as 5 Å from their crystallographic positions. After approximately 11 μs, the agonist-bound receptor spontaneously transitioned to a more rigid conformation resembling the inactive, carazolol-bound structure and the covalent agonist-bound structure (superposition in Fig. 3b), which remained stable for the remainder of the 30-μs simulation (Fig. 3a, top, and Supplementary Fig. 6). In particular, TM5 and TM6 reverted to the positions they adopt in the inactive structure, as did a number of side chains, including Ile $121^{3.40}$ and Phe $282^{6.44}$ (Fig. 3a, bottom). Following the transition, the TM3–TM6 ionic lock was usually intact, as observed previously in simulations initiated from the inactive structure[19]. Although this 30-μs simulation is an order of magnitude longer than any previously published atomistic simulation of a membrane protein, the transition to an inactive-like conformation took place more quickly than the millisecond timescales observed experimentally for adrenergic receptor activation[20]. An additional simulation with different protonation states for Asp $79^{2.50}$ and Asp $130^{3.49}$—two conserved residues whose protonation states have been suggested to change upon receptor activation[21,22]—showed similar behaviour (Supplementary Fig. 7).

In the dynamic conformational equilibrium of a GPCR that links diffusible agonist binding and G protein association, the energies of different states reflect both the ligand binding energies and the conformations of the receptor and its binding partners. This can be depicted in a hypothetical energy landscape of the β_2AR as shown in Fig. 3c, where R, R′, R′′ and R* represent members of the ensemble of receptor conformations along an activation pathway. The constitutive activity displayed by the β_2AR implies that the energy differences and barriers between inactive (R) and active (R*) conformations are low enough to allow a significant population of active state receptors even in the absence of agonist. Agonist binding decreases the energy difference and thus increases the population of receptors in an active conformation; however, the inactive state is still the most stable

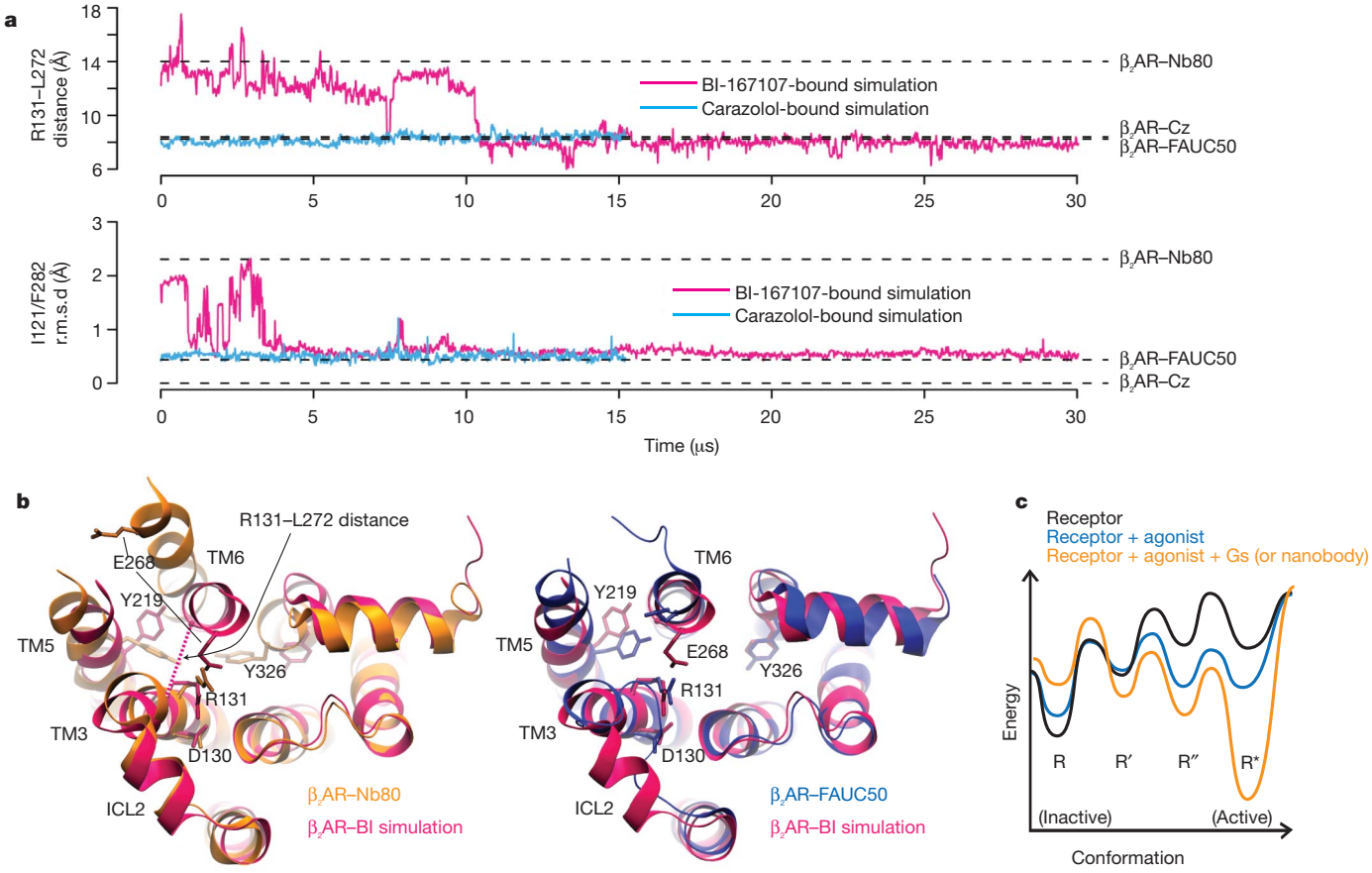

Figure 3 | Molecular dynamics simulations. **a**, An unbiased simulation initiated from the nanobody complex (β_2AR–Nb80) structure, with the agonist BI-167107 bound but the nanobody removed (magenta), and a carazolol-bound simulation initiated from the inactive structure (β_2AR–Cz) (blue). Top, distance between C_α atoms of Arg 131$^{3.50}$ and Leu 272$^{6.34}$; and bottom, r.m.s.d. (root mean squared deviation) from the inactive structure of non-symmetric non-hydrogen atoms in residues Ile 121$^{3.40}$ and Phe 282$^{6.44}$. Dashed lines indicate corresponding quantities from crystal structures. **b**, Cytoplasmic view of the simulated agonist-bound receptor after 30 μs, compared to the β_2AR–Nb80 (left) and β_2AR–FAUC50 (right) structures. The conformations of intracellular loop 2, Tyr 219$^{5.58}$ and Glu 268$^{6.30}$ shown for the agonist-bound simulation differ from β_2AR–FAUC50, but have been observed in inactive β_2AR simulations[19] and in other inactive-state GPCR structures[1]. **c**, Proposed energy landscape model, in which both an agonist and a cytoplasmic binding partner are required to stabilize the fully active receptor conformation (R*) over intermediate (R′ and R′′) and inactive (R) states.

conformation. Our crystallographic studies and MD simulations are compatible with fluorescence lifetime experiments on the purified receptor, demonstrating that saturating concentrations of a full agonist are incapable of pushing the β_2AR conformational equilibrium towards a homogenous active state[23]. These results indicate that binding energy from a G protein or nanobody interaction is required to stabilize conformational changes such as those observed in the vicinity of Ile 121$^{3.40}$ and Phe 282$^{6.44}$ in the active state (Fig. 2c). Alternative biophysical approaches such as NMR[24,25] and further molecular dynamics simulations will be crucial to understand these transitions and identify potential pathways and intermediates that are compatible with these structures.

METHODS SUMMARY

Synthetic chemistry methodology to generate FAUC50 and FAUC72 is described in Supplementary Information. Protocols for radioligand binding and Gs activation assays to characterize ligands are described in Methods.

Crystallization. FAUC50–β_2AR^{H93C}T4L was expressed, purified and crystallized as described in Methods. The receptor was crystallized in a cholesterol-doped (10%) monoolein cubic phase overlaid with precipitant in glass sandwich plates. Optimized precipitant consisted of 24–27%(v/v) PEG 400, 200 mM Li_2SO_4, 4%(v/v) DMSO, 3.5%(v/v) 1,4-butanediol, 100 mM MES pH 6.7. After 3–5 days of growth, crystals were harvested after adding an excess of precipitant solution for cryoprotection, and flash-frozen in liquid nitrogen.

Data collection, structure solution and refinement. Diffraction data were collected at beamline 23-ID (GM/CA-CAT) of the Advanced Photon Source, using a 10 μm diameter collimated microbeam. Oscillation data were measured in 1.0° frames with 10 s or 25 s exposures using 2× or 5× attenuated beam, respectively. The complete data set consisted of 106° of data from 19 crystals. The structure of FAUC50–β_2AR^{H93C}T4L was solved by molecular replacement, and refined by group B factor and TLS refinement. Full details are provided in Methods.

Molecular dynamics. All-atom classical molecular dynamics simulations with explicitly represented lipids and water were performed using the CHARMM force field[26] on Anton[27], a special-purpose computer that accelerates such simulations by orders of magnitude; details are provided in the Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 August; accepted 11 November 2010.

1. Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459,** 356–363 (2009).
2. Caffrey, M. Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu. Rev. Biophys.* **38,** 29–51 (2009).
3. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* doi:10.1038/nature09648 (this issue).
4. Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343,** 1409–1438 (2004).
5. Palczewski, K. *et al.* Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289,** 739–745 (2000).
6. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454,** 183–187 (2008).

7. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455,** 497–502 (2008).
8. Deupi, X. & Kobilka, B. K. Energy landscapes as a tool to integrate GPCR structure, dynamics, and function. *Physiology (Bethesda)* **25,** 293–303 (2010).
9. Dohlman, H. G., Caron, M. G., Strader, C. D., Amlaiky, N. & Lefkowitz, R. J. Identification and sequence of a binding site peptide of the β_2 adrenergic receptor. *Biochemistry* **27,** 1813–1817 (1988).
10. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein coupled receptors. *Meth. Neurosci.* **25,** 366–428 (1995).
11. Buck, E. & Wells, J. A. Disulfide trapping to localize small-molecule agonists and antagonists for a G protein-coupled receptor. *Proc. Natl Acad. Sci. USA* **102,** 2719–2724 (2005).
12. Kolb, H. C., VanNieuwenhze, M. S. & Sharpless, K. B. Catalytic asymmetric dihydroxylation. *Chem. Rev.* **94,** 2483–2547 (1994).
13. Martinelli, M. J. *et al.* Catalytic regioselective sulfonylation of α-chelatable alcohols: scope and mechanistic insight. *J. Am. Chem. Soc.* **124,** 3578–3585 (2002).
14. Whorton, M. R. *et al.* A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proc. Natl Acad. Sci. USA* **104,** 7682–7687 (2007).
15. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature Protocols* **4,** 706–731 (2009).
16. Cherezov, V., Peddi, A., Muthusubramaniam, L., Zheng, Y. F. & Caffrey, M. A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases. *Acta Crystallogr. D* **60,** 1795–1807 (2004).
17. Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. *Science* **318,** 1266–1273 (2007).
18. Tota, M. R., Candelore, M. R., Dixon, R. A. & Strader, C. D. Biophysical and genetic analysis of the ligand-binding site of the β-adrenoceptor. *Trends Pharmacol. Sci.* **12,** 4–6 (1991).
19. Dror, R. O. *et al.* Identification of two distinct inactive conformations of the β_2-adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl Acad. Sci. USA* **106,** 4689–4694 (2009).
20. Vilardaga, J. P., Bunemann, M., Krasel, C., Castro, M. & Lohse, M. J. Measurement of the millisecond activation switch of G protein-coupled receptors in living cells. *Nature Biotechnol.* **21,** 807–812 (2003).
21. Vanni, S., Neri, M., Tavernelli, I. & Rothlisberger, U. A conserved protonation-induced switch can trigger ''ionic-lock'' formation in adrenergic receptors. *J. Mol. Biol.* **397,** 1339–1349 (2010).
22. Ghanouni, P. *et al.* The effect of pH on β_2 adrenoceptor function. Evidence for protonation-dependent activation. *J. Biol. Chem.* **275,** 3121–3127 (2000).
23. Ghanouni, P. *et al.* Functionally different agonists induce distinct conformations in the G protein coupling domain of the β_2 adrenergic receptor. *J. Biol. Chem.* **276,** 24433–24436 (2001).
24. Ahuja, S. *et al.* Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. *Nature Struct. Mol. Biol.* **16,** 168–175 (2009).
25. Bokoch, M. P. *et al.* Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* **463,** 108–112 (2010).
26. MacKerell, A. D. Jr *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616 (1998).
27. Shaw, D. E. *et al.* in *Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis* (ACM Press, 2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge support from National Institutes of Health Grants NS028471 and GM083118 (B.K.K.), GM56169 (W.I.W.), P01 GM75913 (S.H.G), GM75915 and P50GM073210 (M.C), and P60DK-20572 (R.K.S.), the Mathers Foundation (B.K.K and W.I.W), the Lundbeck Foundation (Junior Group Leader Fellowship, S.G.F.R.),Science Foundation Ireland (07/IN.1/B1836) and FP7 COST Action CM0902 (M.C.), the Bavaria California Technology Center (P.G.), and the University of Michigan Biomedical Sciences Scholars Program (R.K.S). We thank Stefan Löber, Harald Hübner, Albert Pan and Paul Maragakis for discussion and suggestions. We thank Foon Sun Thian for help with insect cell expression.

Author Contributions D.M.R. designed project and agonists, did binding assays to characterize agonists, developed purification, optimized crystallization conditions, optimized construct, purified protein, grew and harvested crystals, collected data, solved structure, refined structure, wrote manuscript. C.Z. performed G protein activation assays for covalent agonist, prepared recombinant baculovirus, performed large-scale expression of recombinant β_2AR in insect cells, purified protein, grew crystals. J.L. helped to optimize crystallization conditions, grew and harvested crystals, collected data. D.A. was involved in LCP optimization, harvested crystals, collected data. R.H. synthesized agonists. S.G.F.R. identified the use of MNG3 detergent for β_2AR stabilization and assisted with manuscript preparation. H.J.C. assisted with data processing and refinement. R.K.S. and B.D. provided ApoA1 and Gs protein for functional characterization of the covalent ligand. P.S.C. and S.G. provided MNG3 detergent for stabilization of purified β_2AR. D.H.A. and R.O.D. performed and analyzed MD simulations; R.O.D. and D.E.S. oversaw MD simulations and analysis. W.I.W. assisted with data processing and refinement and with manuscript preparation. M.C. helped to guide the LCP crystallization efforts at Stanford and in Ireland, and oversaw automated lipidic cubic phase crystallography screens. P.G. designed the strategy for the synthesis of the covalent agonist. B.K. was responsible for the overall project strategy and management, oversaw manuscript preparation, and assisted with synchrotron data collection.

Author Information Coordinates and structure factors for β_2AR–FAUC50 are deposited in the Protein Data Bank (accession code 3PDS). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.K.K. (kobilka@stanford.edu), M.C (martin.caffrey@tcd.ie) or P.G. (peter.gmeiner@medchem.uni-erlangen.de).

METHODS

G protein activation assay with ICI-118,551 reversal. Wild-type β2AR or $\beta2AR^{H93C}$ were purified in unliganded form[28]. Samples of these receptors were reconstituted into rHDL (recombinant high density lipoprotein) particles as described[14]. For GTPγS binding assay, receptor-rHDL particles were pre-incubated with 5 μM FAUC50 or 5 μM FAUC72 for 4 h at 4 °C. Samples were diluted 20-fold into binding buffer and split. Half of the samples were used in GTPγS binding assays without ICI-118,551 competition, while the other half was incubated with 6 μM ICI-118,551 at room temperature for 1 h with shaking. Control samples of receptor-rHDL particles with no ligand, 10 μM isoproterenol, or 10 μM ICI-118,551 were also prepared. Purified Gs heterotrimer[29] was added to each sample and incubated for 10 min at 23 °C. The final concentrations of reconstituted receptor and Gs were 100 nM and 600 nM, respectively. GTPγS binding reactions were initiated by the addition of 0.4 nM [^{35}S]GTPγS. Free [^{35}S]GTPγS was removed by rapid filtration of the particles using glass fibre filters. Filter-bound radioactivity was determined by liquid scintillation counting using a Beckman LS6000 scintillation counter. Data shown in Fig. 1c are from three independent experiments each performed in triplicate.

Purification of β_2AR^{H93C}T4L. The construct used for crystallography consisted of the β2AR–T4L fusion protein[17] with the following modifications: a TEV (tobacco etch virus) protease site was inserted after residue 23 of the human β2AR sequence; Histidine 93 was mutated to cysteine; the construct was terminated by a 6×His tag after residue 348 of human β2AR. This construct was cloned into the baculovirus transfer vector pVL1392, and the resulting vector was used to make a high-titre baculovirus stock using the Bestbac system (Expression Systems). The β_2AR^{H93C}T4L receptor was expressed in Sf9 insect cell cultures infected with this baculovirus, and Sf9 membranes were solubilized as described[17]. Purification was achieved using M1 Flag affinity chromatography (Sigma), functional alprenolol-Sepharose chromatography[30], and a second M1 chromatography concentrating step. While the receptor was still bound to the second M1 Flag column, bound alprenolol was washed out and replaced with the covalent ligand using 3 h of washing with 15 column volumes of buffer containing 30 μM FAUC50 (tenfold total molar excess over receptor). Likewise, the dodecylmaltoside detergent used in all previous steps was exchanged for 0.1% (w/v) MNG-3 amphiphile[31]. Covalent agonist-bound and detergent-exchanged $\beta2AR^{H93C}$T4L was eluted in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.1% (w/v) MNG-3, 30 μM FAUC50. In contrast to previously published protocols, receptor was not alkylated before alprenolol-Sepharose chromatography. Instead, receptor was alkylated after agonist exchange by treatment with 2 mM iodoacetamide for 30 min at 4 °C. Incubation of the eluate with AcTEV protease (Invitrogen) succeeded in removing the N terminus of the receptor (23 amino acids plus TEV site, leaving a glycine scar preceding residue 24), as verified by SDS–PAGE analysis. The sample at this stage was further purified and concentrated using Ni chelating chromatography, taking advantage of the C-terminal 6×His tag. Receptor was bound to a 0.7-ml column Ni-Sepharose column, and eluted in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.1% (w/v) MNG-3 and 30 μM FAUC50, 200 mM imidazole. Finally, the receptor was concentrated to 50 mg ml^{-1} with a 100 kDa cutoff Vivaspin concentrator (Vivascience).

Crystallization. Purified FAUC50–β_2AR^{H93C}T4L was crystallized using the *in meso* method[2]. Cubic phase reconstitution consisted of two parts 50 mg ml^{-1} receptor and three parts molten lipid mixture (10:1 monoolein:cholesterol by mass, lipids purchased from Sigma). Note that cholesterol is required for successful crystallization, and one cholesterol molecule is included in the refined model. Aqueous and lipid components were combined at room temperature using a syringe mixing apparatus[32]. Crystallization experiments in glass sandwich plates, set up by either by hand or using an *in meso* robot[16], consisted of 30–50 nl cubic phase overlaid with 800 nl precipitant. Precipitant conditions producing diffraction quality crystals were identified by screening around previous conditions[17], and testing additives and alternative buffers. Final optimized conditions consisted of 24–27%(v/v) PEG 400, 200 mM Li_2SO_4, 4%(v/v) DMSO, 3.5%(v/v) 1,4-butanediol, 100 mM MES pH 6.7. Crystals grew at 20 °C to a maximum size of 50 × 15 × 5 μm^3 within 3 to 5 days (see Supplementary Fig. 3, Supplementary Materials). For harvesting and cryocooling, exposed crystallization drops were overlaid with an excess of precipitant solution, and cryoloops (MiTeGen) containing single crystals were flash-frozen in liquid nitrogen.

Data collection and processing. Diffraction data were collected at beamline 23-ID of the General Medicine and Cancer Institutes Collaborative Access Team (GM/CA-CAT) of the Advanced Photon Source, Argonne, Illinois, USA. All data was acquired using a 10-μm diameter collimated microbeam. Attenuated 1.0° rotation images were used to locate and centre crystals within the opaque mesophase in each cryoloop. Oscillation data were measured in 1.0° frames with 10 s or 25 s exposures using 2× or 5× attenuated beam, respectively. Significant radiation damage caused decay in the signal that prevented merging more than the first 5–10° of oscillation data from each crystal. A total of 106° of data from 19 crystals were integrated with Mosflm[33] and merged with Scala[34].

Structure solution and refinement. Molecular replacement to obtain initial phases was performed with the program Phaser[35]. The separated structures of the receptor and T4L components of the high-resolution carazolol-bound β_2AR structure (PDB ID 2RH1 with all non-protein atoms removed), were used as search models. The model was refined in Phenix[36] and Buster[37], using group B factor refinement (one B factor per residue) followed by TLS refinement (using two TLS groups, one for the receptor and one for T4L). Refinement statistics are given in Supplementary Table 1 of Supplementary Information. The crystallographic data was strongly anisotropic, as seen in the anisotropic B factor corrections; however, the electron density was clear enough for the placement of side chains. Although there was clear electron density present for the agonist (see omit map in Supplementary Fig. 4, Supplementary Information), we did observe a small discontinuity at the end of the linker connecting to Cys 93. This discontinuity could arise from flexibility of the polymethylene component of the linker, as well as potential radiation damage[38] localized to the disulphide bond. As a control, we carried out refinement in which the ligand was modelled as a non-covalent species without a disulphide connection to the receptor. No significant differences in final model R/R_{free} or $2Fo - Fc$ and $Fo - Fc$ electron density maps were observed.

Molecular dynamics simulations. In all simulations, β_2AR was embedded in a hydrated lipid bilayer, with all atoms, including those in the lipids and water, represented explicitly. The BI-167107-bound β_2AR simulations were initiated with the receptor and ligand in the conformation of the β_2AR–Nb80 crystal structure (companion paper[3]), with the nanobody (Nb80) removed. The carazolol-bound β_2AR simulations were initiated with the receptor and ligand in the conformation of the high-resolution carazolol-bound crystal structure[17]. Both crystal structures were determined using a β_2AR–T4L fusion protein, in which intracellular loop 3 (ICL3) of the receptor was replaced by T4 lysozyme (although the T4L was not resolved in the BI-167107-bound nanobody complex structure). We omitted the T4L in all simulations. Experimentally, removal of the bulk of ICL3 by partial tryptic digest does not seem to affect receptor function[39].

Production simulations were performed on Anton[27], a special-purpose computer designed to accelerate standard molecular dynamics simulations by orders of magnitude relative to the previous state of the art. Prior to production simulation, systems were equilibrated using Desmond[40] on a commodity cluster, according to the protocol described below.

Hydrogens were added to the crystal structures of carazolol-bound β_2AR (β_2AR–Cz; PDB ID 2RH1) and β_2AR–Nb80/BI-167107 (companion paper[3]) using Maestro (Schrödinger) as described previously[19]. T4L and Nb80 were deleted, and chain termini were capped with neutral groups (acetyl and methylamino).

All titratable residues other than Glu $122^{3.41}$, Asp $79^{2.50}$, and Asp $130^{3.49}$ were left in the dominant protonation state at pH 7.0. Glu $122^{3.41}$ was protonated in all simulations. It faces the lipid bilayer and is thus likely protonated[41]; in addition, a similarly positioned residue in rhodopsin (Glu $122^{3.37}$) has been found to be protonated during the entire photocycle[22].

We performed both BI-167107-bound and carazolol-bound simulations using two different sets of protonation states for Asp $79^{2.50}$ and Asp $130^{3.49}$. Previous studies have suggested that Asp $130^{3.49}$ is protonated upon activation[22], and FTIR (Fourier transform infrared) spectroscopy data has shown that this is the case for the corresponding residue of rhodopsin, Glu $134^{3.49}$ (ref. 42). Asp $79^{2.50}$ is homologous to Asp $83^{2.50}$ of rhodopsin, which has been found by FTIR spectroscopy to remain protonated during the entire photocycle[41]. On the other hand, neutralization of Asp $79^{2.50}$ in β_2AR by mutation to asparagine uncouples agonist binding from G protein activation[43], and a recent study suggested that Asp 79 may be deprotonated upon activation[21]. Thus, we repeated each simulation with two sets of protonation states: first with Asp 79 protonated and Asp 130 deprotonated ('Ash 79/Asp 130', representing potential protonation states in the inactive receptor), and then with Asp 79 deprotonated and Asp 130 protonated ('Asp 79/Ash 130', representing possible protonation states in the active receptor). Results shown in Supplementary Figs 3, 5 and 6 are for the Ash 79/Asp 130 simulations, whereas results in Supplementary Fig. 7 are for the Asp 79/Ash 130 simulations.

β_2AR residues that were truncated or not resolved in the crystal structures were omitted from the simulations. In particular, N-terminal residues 1–28, C-terminal residues 343–413, and ICL3 residues 231–262 were omitted from the carazolol-bound system, whereas N-terminal residues 1–22, C-terminal residues 345–413 and ICL3 residues 228–265 were omitted from the BI-167107-bound system. Both simulations included a glutamate at position 187, reflecting an Asn 187 Glu mutation made in both crystallization constructs to eliminate a glycosylation site.

We prepared the carazolol-bound Ash 79/Asp 130 β_2AR system following our previously described protocol[19], and the other systems according to an updated protocol (details below). We do not expect the differences in the simulation setup

protocols to affect the quantities of interest in this study. As a control, however, we performed an additional 5 μs simulation of a carazolol-bound Ash 79/Asp 130 β_2AR system prepared according to the new protocol and obtained results (not shown) similar to those of the 15 μs carazolol-bound Ash 79/Asp 130 β_2AR simulation prepared according to the old protocol.

Prepared protein structures were inserted into an equilibrated phospholipid bilayer solvated with 0.15 M NaCl as described previously[19]. Carazolol-bound Ash 79/Asp 130 β_2AR was simulated in a POPE (palmitoyl oleoyl phosphatidyl ethanolamine) bilayer system initially measuring $85 \times 73 \times 88$ Å, containing 170 lipid molecules, 8,798 water molecules, 19 sodium ions, and 24 chloride ions, for a total of 52,396 atoms. Carazolol-bound Asp 79/Ash 130 β_2AR was simulated in a POPC (palmitoyl oleoyl phosphatidyl choline) bilayer system initially measuring $83 \times 83 \times 87$ Å, containing 160 lipid molecules, 11,313 water molecules, 30 sodium ions, and 35 chloride ions, for a total of 60,153 atoms. (The 5 μs carazolol-bound Ash 79/Asp 130 control simulation prepared according to the new protocol also used a POPC bilayer.) BI-167107-bound β_2AR (both Ash 79/Asp 130 and Asp 79/Ash 130) was simulated in a POPC bilayer system initially measuring $79 \times 79 \times 87$ Å, containing 138 lipid molecules, 10,116 water molecules, 26 sodium ions, and 31 chloride ions, for a total of 53,603 atoms.

All systems were equilibrated in the NPT ensemble at 310 K and 1 bar using the Berendsen coupling scheme with 5 kcal mol^{-1} Å^{-2} harmonic position restraints applied to all non-hydrogen atoms of the protein; these restraints were tapered off linearly over 5 ns. Unrestrained systems were then simulated for an additional 5 ns to equilibrate the aspect ratio of the simulation box further. During the equilibration process, van der Waals and short-range electrostatic interactions were cut off at 9 Å for the carazolol-bound Ash 79/Asp 130 β_2AR system and at 12 Å for the other systems. Long-range electrostatic interactions were computed using the particle mesh Ewald method[44], with a $64 \times 64 \times 64$ grid and $\sigma = 2.26$ Å for the carazolol-bound Ash 79/Asp 130 β_2AR system, and with a $32 \times 32 \times 32$ grid and $\sigma = 3.23$ Å for the other systems; fifth-order B-splines were used for interpolation in both cases. All bond lengths to hydrogen atoms were constrained using M-SHAKE[45]. A RESPA integrator[46] was used with a time step of 2 fs, and long-range electrostatics were computed every 6 fs.

Production simulations on Anton were initiated from the final snapshot of each corresponding equilibration run on Desmond, using the same integration schemes. Van der Waals and short-range electrostatic interactions were cut off at 9 Å for the carazolol-bound Ash 79/Asp 130 β_2AR system and at 13.5 Å for the other systems. Long-range electrostatics were computed using the *k*-space Gaussian split Ewald method[47], with a $64 \times 64 \times 64$ grid, $\sigma = 2.01$ Å, and $\sigma_s = 1.41$ Å for the carazolol-bound Ash 79/Asp 130 β_2AR system, and with a $32 \times 32 \times 32$ grid, $\sigma = 3.33$ Å, and $\sigma_s = 2.35$ Å for the other systems.

Force field parameters: the CHARMM27 (ref. 26) parameter set (with CMAP terms[48]) and the CHARMM TIP3P[49] water model were used for all protein molecules, POPE lipid molecules, water molecules, and salt ions. A modified CHARMM lipid force field[50], which became available after we performed the 15-μs simulation of carazolol-bound β_2AR in POPE, was used for POPC lipids. Force-field parameters for carazolol and palmitoyl-cysteine were designed previously[19]. Force-field parameters for BI-167107 were transferred from previously parameterized model compounds: parameters for the hydroxyethylamine 'tail' group were transferred from the alkylamine parameters we previously designed for carazolol[19], and parameters for the benzoxazine 'head' group were transferred from the model compounds anisole and *p*-phenol acetamide from the CHARMM General Force Field[51]. Full parameter sets are available upon request.

Analysis protocols: trajectory snapshots, each containing a record of all atom positions at a particular instant in time, were saved every 180 ps during production simulation. Distance and r.m.s.d. measurements were computed using the HiMach parallel analysis framework[52].

The distance and r.m.s.d. measurements shown in Fig. 3a, Supplementary Fig. 5 and Supplementary Fig. 7 are smoothed; the smoothed time series were computed from the original time series by a weighted running average, using a filter kernel of half-width 12.51 ns whose shape corresponded to that of the cosine function from $-\pi/2$ to $\pi/2$.

VMD[53] was used to visualize trajectories and to produce the molecular renderings of Fig. 3b.

28. Yao, X. *et al.* Coupling ligand structure to specific conformational switches in the β_2-adrenoceptor. *Nature Chem. Biol.* **2,** 417–422 (2006).
29. Kozasa, T. & Gilman, A. G. Purification of recombinant G proteins from Sf9 cells by hexahistidine tagging of associated subunits. Characterization of α_{12} and inhibition of adenylyl cyclase by α_z. *J. Biol. Chem.* **270,** 1734–1741 (1995).
30. Kobilka, B. K. Amino and carboxyl terminal modifications to facilitate the production and purification of a G protein-coupled receptor. *Anal. Biochem.* **231,** 269–271 (1995).
31. Chae, P. S. *et al.* Maltose–neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nature Methods* **7,** 1003–1008 (2010).
32. Cheng, A., Hummel, B., Qiu, H. & Caffrey, M. A simple mechanical mixer for small viscous lipid-containing samples. *Chem. Phys. Lipids* **95,** 11–21 (1998).
33. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62,** 48–57 (2006).
34. Collaborative Computational Project, number 4. The *CCP*4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
35. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Cryst.* **40,** 658–674 (2007).
36. Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. A robust bulk-solvent correction and anisotropic scaling procedure. *Acta Crystallogr. D* **61,** 850–855 (2005).
37. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in *BUSTER-TNT*. *Acta Crystallogr. D* **60,** 2210–2221 (2004).
38. Weik, M. *et al.* Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc. Natl Acad. Sci. USA* **97,** 623–628 (2000).
39. Rubenstein, R. C. *et al.* The hydrophobic tryptic core of the β-adrenergic receptor retains G_s regulatory activity in response to agonists and thiols. *J. Biol. Chem.* **262,** 16655–16662 (1987).
40. Bowers, K. J. *et al.* in *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing* (ACM Press, 2006).
41. Fahmy, K. *et al.* Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. *Proc. Natl Acad. Sci. USA* **90,** 10206–10210 (1993).
42. Vogel, R. *et al.* Functional role of the ''ionic lock''–an interhelical hydrogen-bond network in family A heptahelical receptors. *J. Mol. Biol.* **380,** 648–655 (2008).
43. Liapakis, G., Chan, W. C., Papadokostaki, M. & Javitch, J. A. Synergistic contributions of the functional groups of epinephrine to its affinity and efficacy at the β_2 adrenergic receptor. *Mol. Pharmacol.* **65,** 1181–1190 (2004).
44. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an *N*·log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089 (1993).
45. Kräutler, V., van Gunsteren, W. F. & Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **22,** 501–508 (2001).
46. Tuckerman, M., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97,** 1990 (1992).
47. Shan, Y., Klepeis, J. L., Eastwood, M. P., Dror, R. O. & Shaw, D. E. Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. *J. Chem. Phys.* **122,** 054101 (2005).
48. Mackerell, A. D. Jr, Feig, M. & Brooks, C. L. III. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25,** 1400–1415 (2004).
49. Beglov, D. & Roux, B. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J. Chem. Phys.* **100,** 9050–9063 (1994).
50. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114,** 7830–7843 (2010).
51. Vanommeslaeghe, K. *et al.* CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31,** 671–690 (2010).
52. Tu, T. *et al.* in *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing* (ACM Press, 2008).
53. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38 (1996).