Sciencexpress Research Article

High-Resolution Crystal Structure of an Engineered Human β_2-Adrenergic G Protein–Coupled Receptor

Vadim Cherezov,[1]* Daniel M. Rosenbaum,[2]* Michael A. Hanson,[1] Søren G. F. Rasmussen,[2] Foon Sun Thian,[2] Tong Sun Kobilka,[2] Hee-Jung Choi,[2,3] Peter Kuhn,[4] William I. Weis,[2,3] Brian K. Kobilka,[2]† Raymond C. Stevens[1]†

[1]Department of Molecular Biology, Scripps Research Institute, La Jolla, CA 92037, USA. [2]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, USA. [3]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [4]Department of Cell Biology, Scripps Research Institute, La Jolla, CA 92037, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: stevens@scripps.edu (R.C.S.); kobilka@stanford.edu (B.K.K.)

G protein–coupled receptors comprise the largest family of eukaryotic signal transduction proteins that communicate across the membrane. We report the crystal structure of a human β_2-adrenergic receptor–T4 lysozyme fusion protein bound to the partial inverse agonist carazolol at 2.4 Å resolution. The structure provides a high-resolution view of a human G protein–coupled receptor bound to a diffusible ligand. Ligand-binding site accessibility is enabled by the second extracellular loop which is held out of the binding cavity by a pair of closely spaced disulfide bridges and a short helical segment within the loop. Cholesterol, a necessary component for crystallization, mediates an intriguing parallel association of receptor molecules in the crystal lattice. Although the location of carazolol in the β_2-adrenergic receptor is very similar to that of retinal in rhodopsin, structural differences in the ligand binding site and other regions highlight the challenges in using rhodopsin as a template model for this large receptor family.

G protein–coupled receptors (GPCRs) comprise the largest integral membrane protein family in the human genome, with over one thousand members (*1*, *2*). These receptors actively participate in the transduction of signals across cellular membranes in response to an astonishing variety of extracellular stimuli, including light, proteins, peptides, small molecules, hormones, protons and ions. Once activated, GPCRs trigger a cascade of intracellular responses, primarily through interactions with their cognate heterotrimeric G proteins, although G protein independent signaling pathways have also been described (*3*–*5*). GPCRs are major contributors to the information flow into cells and, as such, are associated with a multitude of diseases that make members of this family important pharmacological targets (*6*).

GPCRs have been grouped into five classes (*2*) based on sequence conservation, with class A being the largest and most studied. Class A receptors are further divided into groups associated with particular ligand specificity, such as the opsin, amine, peptide, cannabinoid, and olfactory receptors. Historically, the adrenergic receptors in the amine group are some of the most thoroughly investigated of the class A GPCRs (*7*–*12*), and are composed of two main subfamilies, α and β, which differ in tissue localization and ligand specificity, as well as in G protein coupling and downstream effector mechanisms (*13*). Genetic modifications of adrenergic receptors are associated with diseases as diverse as asthma, hypertension, and heart failure (*14*). β_2-adrenergic receptors (β_2ARs) reside predominantly in smooth muscle throughout the body, and β_2AR agonists are used in the treatment of asthma and preterm labor (*15*–*17*).

Despite extensive efforts, structural information for only one member of the eukaryotic GPCR family, bovine rhodopsin, is available to date (*18*–*21*). Rhodopsin is unusual in that it is highly abundant from natural sources and structurally stabilized by the covalently bound ligand 11-*cis*-retinal, which maintains the receptor in a dark-adapted, non-signaling conformation. In contrast, all other GPCRs are activated by diffusible ligands and are expressed at relatively low levels in native tissues. These receptors are structurally more flexible and equilibrate among multiple conformational states, some of which are prone to instability (*22*). While the structure determination of rhodopsin was important, many questions remain on the conformational changes between different activation states for each receptor, as well as the structural differences amongst receptors that accommodate the very large diversity of ligands. Specifically: (i) What structural features enable GPCRs to recognize and bind diffusible ligands? (ii) How structurally conserved are the class A GPCRs, and what is the importance of both similarities and differences?

To address these questions, we modified the human β_2AR to facilitate the growth of diffraction quality crystals by inserting T4-lysozyme (T4L) in place of the third intracellular loop (β_2AR-T4L) and solved the three-dimensional crystal structure in the presence of a partial inverse agonist carazolol (2-propanol, 1-(9*H*-carbazol-4-yloxy)-3-[(1-methylethyl)amino]) at 2.4 Å resolution (*23*, *24*). We provide a comprehensive analysis of the crystal packing and intramolecular contacts between the β_2AR and T4L to identify potential receptor perturbing interactions. The overall receptor topology and the ligand binding pocket are described, as are the main similarities and differences between β_2AR-T4L and rhodopsin, and the implications for modeling of other GPCR-ligand complexes.

Structure determination. The engineering, functional properties, expression and purification of crystallization grade β_2AR-T4L protein are described fully in the companion paper (*25, 26*). Briefly, β_2AR-T4L was expressed in Sf9 insect cells, solubilized in 1% dodecylmaltoside, and purified by sequential antibody and ligand affinity chromatography. Following the reported success with microbial rhodopsins in lipidic cubic phase (LCP) (*27*), we were able to produce crystals of β_2AR-T4L that diffract to a resolution of 2.2 Å with a modified LCP procedure, and to solve and refine the structure at 2.4 Å resolution (*28*). Compared to crystallization in detergents, LCP provides a more native, lipid environment for crystallization, as well as a confinement of protein molecules to two-dimensional membrane sheets that may facilitate the crystallization process through the formation of Type I packing interactions (*29–31*). In agreement with prior biological evidence that cholesterol improves β_2AR stability (*32*) and may mediate receptor-receptor interactions, crystals were grown from a cholesterol-doped monoolein cubic phase. An automated, nanovolume LCP crystallization protocol (*33*) significantly reduced the time and amount of protein required for the exhaustive, multi-dimensional optimization trials needed to arrive at these conditions. Crystals of b2AR-T4L were also obtained in lipid bicelles, but they did not diffract as well as those obtained in LCP (*28*).

Diffraction data for β_2AR-T4L were measured to a resolution of 2.4 Å from a total of 27 microcrystals (average size 30 x 15 x 5 µm) using a high intensity, highly parallel minibeam with a diameter of 10 microns at the GM/CA-CAT beamline of the Advanced Photon Source, Argonne National Laboratory (*34*). Phase information was obtained by molecular replacement using both T4 lysozyme (PDB ID Code 2LZM) and a polyalanine model of the transmembrane regions of rhodopsin (PDB ID Code 1U19) as search models. Additional crystallization, data collection, processing, and refinement statistics are reported in Table 1 and discussed in detail in (*28*).

Overall receptor topology. The final model of β_2AR-T4L includes 442 amino acids. The model also includes a palmitic acid covalently bound to Cys341 and an acetamide molecule bound to Cys265$^{6.27}$ (throughout the text residues are designated by their position within the β_2AR sequence and their Ballesteros-Weinstein designation as a superscript where applicable) (*35, 36*), as well as one carazolol molecule, three cholesterol molecules, two sulfate ions and two butanediol molecules that interact with β_2AR. There are also four sulfate ions, a putative disaccharide (modeled as maltose) and a molecule of PEG 400 bound to T4L. For β_2AR, excellent electron density is observed for residues 29-342, including the ligand carazolol and the two disulfide bonds Cys106$^{3.25}$-Cys191$^{5.30}$ and Cys184$^{4.76}$-Cys190$^{5.29}$. The palmitic acid at Cys341 is clearly visible in F_o-F_c omit maps; however, the quality of the electron density is lower than for the rest of the receptor. The N-terminus (residues 1 to 28) and the majority of the C-terminus (residues 343 to 365) are disordered and not visible in the structure.

The β_2AR has a fold composed of seven transmembrane helices forming a helical bundle (Fig. 1A). The residues that make up the helices (I to VII) in β_2AR are as follows: helix I 29$^{1.28}$ to 60$^{1.59}$, helix II 67$^{2.38}$ to 96$^{2.67}$, helix III 103$^{3.22}$ to 136$^{3.55}$, helix IV 147$^{4.39}$ to 171$^{4.63}$, helix V 197$^{5.36}$ to 229$^{5.68}$, helix VI 267$^{6.29}$ to 298$^{6.60}$, and helix VII 305$^{7.32}$ to 328$^{7.55}$. The residues forming the intracellular loops (ICL) and extracellular loops (ECL) of β_2AR are: ICL1 61$^{1.60}$ to 66$^{2.37}$, ECL1 97$^{2.68}$ to 102$^{3.21}$, ICL2 137$^{3.56}$ to 146$^{4.38}$, ECL2 172$^{4.64}$ to 196$^{5.35}$, ICL3 230$^{5.69}$ to 266$^{6.28}$ (residues 231 to 262 are replaced by T4-lysozyme residues 2 to 161), and ECL3 299$^{6.61}$ to 304$^{7.31}$. Helices II, V, VI and VII each have a proline-induced kink at conserved positions along the span of the transmembrane segments. These kinks are thought to enable the structural rearrangements required for activation of G protein effectors (*37*). In addition to the seven membrane spanning helices, β_2AR possesses two other helical segments: helix VIII, which is believed to be common to all rhodopsin-like GPCRs (*38*), and an unexpected, short helical segment in the middle of ECL2, which is not present in rhodopsin, and was not predicted by computational secondary structure analysis (Fig. 1A).

In the β_2AR-T4L construct, T4L is fused to the truncated cytoplasmic ends of helices V and VI. In the crystal structure, the T4L moiety is tilted slightly away from the center axis of β_2AR drawn normal to the membrane (Fig. 1B). As a result, interactions between T4L and β_2AR are minimal, with only 400Å^2 of surface area buried between them. The intramolecular contacts between T4L and β_2AR include salt bridges between the side chains of T4L-Asp159 and the side-chain amine of β_2AR-Lys227$^{5.66}$ (distance 3.4 Å) and between the guanidinium group of T4L-Arg8 with the side-chain carboxyl of β_2AR-Glu268$^{6.30}$ on helix VI (distance 3.2 Å) (Fig. 1C and table S2). The latter interaction is noteworthy, as in rhodopsin, Glu$^{6.30}$ forms an ionic bond with Arg$^{3.50}$ of the conserved D(E)RY motif (*18*). This interaction is postulated to be important for maintaining rhodopsin in the inactive state, but the charged groups of the two residues [Arg131$^{3.50}$ (NH1) and Glu268$^{6.30}$ (OE1)] are 10 Å apart in the β_2AR-T4L structure. Possible functional implications of this disruption are discussed in the companion manuscript (*26*). The remainder of the lysozyme molecule provides important crystal packing interactions, but does not appear to influence significantly the receptor structure.

Crystal packing interactions. The β_2AR-T4L protein is packed in a C-centered monoclinic lattice with one molecule per asymmetric unit (Fig. 2A). As observed in all previous lipidic mesophase grown crystals (*39*), the β_2AR-T4L crystals adopt Type I packing (*40*), featuring a multilayered arrangement in accordance with proposed crystallization mechanism (*29, 41*). Within each layer, protein molecules form arrays of parallel, symmetry-related dimers. There are four distinct crystal-packing interactions within each layer, three of which are mediated by T4L. The fourth interaction in the array is between two receptor molecules related by a crystallographic two-fold rotation axis. This is the sole interaction between symmetry-related receptors, and is mediated primarily by ordered lipids consisting of six cholesterol and two palmitic acid molecules, the latter being covalently attached to Cys341 in the C-terminal portion of the receptor (*42*) (Fig. 2B). These eight lipid molecules form a two-fold symmetric sheet between receptors. The only direct receptor-receptor contact involves a 2.7 Å pair of ionic interactions between the charged amine group of Lys60$^{1.59}$ in helix I and the carboxylate of Glu338 in helix VIII from the symmetry-related receptor. Remarkably, of the 515 Å^2 buried at the receptor symmetry interface, 73% of the crystal contact surface area is mediated by ordered lipid, while only 27% is

contributed by protein-protein contacts. The stacking interactions between layers are formed between T4L and extracellular loops ECL2 and ECL3 of the receptor (Fig. 2A). It is unlikely that these contacts affect the orientation of these loops due to the small size of ECL3 and the rigid architecture of ECL2.

Lipid mediated receptor association. Many GPCRs including β_2AR are thought to exist as dimers in the plasma membrane, although the location of the dimer interface and the functional significance of dimerization is not clear (*43*). The observation of ordered lipids in the helix I and VIII interface between two symmetry related molecules makes it tempting to speculate on the physiological relevance of this association (*44–46*). Associations between the equivalent regions of rhodopsin have been found in crystal structures (*21*, *47*) (Fig. 2C). On the other hand, studies in native membranes suggest that helix VI may form the dimer interface for the β_2AR (*48*), and helix IV may form the dimer interface for the closely related D_2 dopamine receptor (*49*).

While the role of cholesterol in promoting β_2AR association is speculative, its role in the physiologic function of β_2AR is well documented. Depletion of cholesterol from the membranes of neonatal cardiac myocytes alters the signaling behavior of endogenous β_2AR (*50*). In untreated cells, activation of β_2AR results in sequential coupling to the G proteins G_s and G_i, producing a biphasic effect on myocyte contraction rate. Upon depletion of cholesterol, the β_2AR couples more strongly to G_s. This effect may be due to a role of cholesterol in regulating interactions between the β_2AR and G proteins, or possibly to the effect of cholesterol on β_2AR dimerization. The β_2AR couples efficiently to G_s as a monomer (*51*), so it is possible that cholesterol mediated association (dimerization) reduces the efficiency of β_2AR coupling to G_s. The effects of cholesterol depletion on β_2AR signaling may also be a secondary effect of altering subcellular signaling compartments. There is evidence that cells may concentrate signaling molecules, such as GPCRs and their cognate G proteins, by way of membrane microdomains or compartments, such as caveolae (*52*). This compartmentalization may be a major regulator of receptor-effector coupling. Thus, the importance of cholesterol in forming the observed crystallographic association is consistent with its role in β_2AR signaling. Additional experiments will be required to determine how relevant the association of monomers observed in the crystal is to β_2AR packing within membrane microdomains.

Electrostatic charge distribution. Electrostatic charge distribution was calculated using APBS (*53*) and mapped onto a molecular surface representation of β_2AR. The analysis reveals three polarized areas within the molecule (Fig. 3A). First, the cytoplasmic face of the receptor is involved in G protein interaction and carries a net positive charge even in the absence of ICL3, which also has a predicted overall positive charge (Fig. 3B). The second site is an electrostatically negative region located within the membrane between helices III, IV and V potentially exposed to the lipid alkyl chains, which is unexpected as the burial of charge within the plasma membrane is thermodynamically unfavorable. A glutamate residue at position $122^{3.41}$ may partially account for the observed charge distribution. Finally, the binding site cleft is negatively charged and exposed to solvent by an unusual ECL2 architecture and lack of N-terminal interactions. This negative charge may facilitate ligand binding through electrostatic funneling of positively charged catecholamines (Fig. 3B).

Extracellular region. The ECLs and amino termini of GPCRs, together with the extracellular halves of the transmembrane helices, are believed to define the ligand-binding site of each receptor (*44*). Therefore, the ECLs play an important role in the overall pharmacology of any particular receptor. In general, small molecule ligands are thought to bind deeper within the space created by the transmembrane domain helices, whereas larger ligands such as peptides bind closer to the membrane surface near the ECLs (*54*, *55*). Mutagenesis studies suggest that the β_2AR binds its ligand deep within the transmembrane helix bundle, which may be related to the observation that the extracellular regions have a rather simple structure with short loops connecting transmembrane helices II and III, and VI and VII (Fig. 4A). ECL2, which links helices IV and V, has a somewhat more extensive architecture that is unanticipated. In contrast to the buried, β-sheet structure of this loop in rhodopsin (Fig. 4B), ECL2 in β_2AR is more exposed to the solvent and contains an extra helical segment. Additionally, there is an intra-loop disulfide bond between $Cys184^{4.76}$ and $Cys190^{5.29}$ that may help stabilize the more exposed ECL2. A second disulfide bond between $Cys191^{5.30}$ and $Cys106^{3.25}$ in helix III effectively ties ECL2 to the transmembrane core (*56*). The distal portion of ECL2 makes close contacts with ECL1 and contains a glycosylation site at $Asn187^{5.26}$ (*57*), which may serve to mask a grouping of aromatic residues on ECL1; in this construct, $Asn187^{5.26}$ has been mutated to glutamate to aid in crystallization.

Electron density corresponding to the N-terminus was not apparent in the maps and, therefore, residues 1-28 are not included in the model. This disorder contrasts with rhodopsin, in which the N-terminus interacts extensively with the ECLs, forming a small four-strand β-sheet in conjunction with ECL2. This sheet structure forms a cap that effectively isolates the retinal binding site in a hydrophobic pocket (Fig. 5B). The lack of interactions between the N-terminus of β_2AR and ECL2 further enables diffusible ligand access to the binding site. However a completely disordered N-terminus may be an artifact induced by the presence of the N-terminal Flag tag which carries an overall positive charge and may disrupt N-terminal interactions.

The short helical region on ECL2 adds a rigid structural element that, along with the two disulfide bonds, constrains the loop to a small range of conformations and helps stabilize the receptor by linking three transmembrane helices (Fig. 5A). This rigid conformation may help to stabilize the core of the receptor and lock ECL2 in a conformation that does not hinder access to the binding pocket.

Ligand binding site and comparison to rhodopsin. Carazolol is a partial inverse agonist that binds with picomolar affinity to β_2AR-T4L producing a reduction of the basal activity of the receptor (*24*). The crystal structure reveals extensive interactions between the receptor and carazolol that position the carbazole moiety adjacent to $Phe289^{6.51}$, $Phe290^{6.52}$, and $Trp286^{6.48}$ (Fig. 5A, fig. S1, and table S3). In contrast, *cis*-retinal is a full inverse agonist covalently bound to rhodopsin, which suppresses all activity towards transducin (*58*). Carazolol and retinal occupy similar spaces in their respective receptors, with significant overlap

of the non-aromatic regions of carazolol. However, the β-ionone ring of retinal extends deep into the binding pocket of rhodopsin and contacts residues on helix V and VI, where it is sandwiched between Phe212$^{5.47}$ and Tyr268$^{6.51}$, and interacts with the highly conserved Trp265$^{6.48}$ (Fig. 5B). It has been proposed that changes in the rotamer of Trp265$^{6.48}$ occur upon activation of rhodopsin and related family members, and constitutes the "toggle switch" for receptor activation (*59*). Accordingly, the interactions between *cis*-retinal and Trp265$^{6.48}$ are likely to contribute to the absence of basal activity in rhodopsin. Carazolol does not interact directly with the toggle switch on helix VI, however it lowers the basal activity of the receptor, and may do so by interacting with Phe289$^{6.51}$ and Phe290$^{6.52}$, which form an extended aromatic network surrounding the highly conserved Trp286$^{6.48}$. As a result, Trp286$^{6.48}$ adopts the rotamer associated with the inactive state. Thus, the steric constraints imposed by Phe290$^{6.52}$ appear to structurally mimic the interaction of the β-ionone ring of retinal with the conserved Trp265$^{6.48}$ and Phe212$^{5.47}$ on rhodopsin (*60*) (Fig. 5C).

Structural alignment and helix bundle reorganization. It has long been thought that class A GPCRs share a similar architecture due to their predicted seven transmembrane helical bundles and sequence conservation within the membrane spanning regions (*61*). Nonetheless, given the common ability to activate G proteins, yet the astonishing variety of ligand specificities among the class A receptor family, the similarities and differences in ligand binding modes remains an open question. To this end, we aligned the structure of β_2AR-T4L to highest resolution structure of rhodopsin (PDB ID Code 1U19). We used difference distance matrices to select non-divergent areas between the two structures that align to reveal the differences in helix orientation between β_2AR-T4L and rhodopsin (*62*).

Relative to rhodopsin, the following helical shifts are seen in β_2AR-T4L: the extracellular portions of helices I and III angle away from the center of the receptor, helix IV is translated away from the center of the receptor, helix V is translated closer to the center of the receptor and helix VI angles away from the receptor on the cytoplasmic end (Fig. 6). The largest difference is in helix I, which lacks a proline-induced kink found in rhodopsin and is comparatively straight. The angle between the rhodopsin and β_2AR positions of helix I is approximately 18° with a shift of 7 Å at the apex on the extracellular face. This structural difference may arise from the need for an accessible binding site in β_2AR, which is provided in part by a lack of interactions between the N-terminus and extracellular loop segments. In contrast the N-terminal region in rhodopsin occludes the retinal-binding site through extensive interactions with the extracellular loops (Fig. 4B). Helix V of β_2AR is closer to the binding pocket by approximately 3.5 Å on average and its lumenal end is angled more towards helix VI. Helix IV of β_2AR is further from the binding site, possibly to remove steric clashes resulting from the modified position of helix V (Fig. 6, B and C). Helix III pivots further from the binding site about a fulcrum located close to the cytoplasmic end (Fig. 6C). The angle formed between rhodopsin helix III and the β_2AR helix III is approximately 7°, yielding a 4 Å displacement out of the binding pocket at the cytoplasmic end of the helix. Helix VI is positioned further from the center of the receptor at the cytoplasmic end as compared to rhodopsin, which is caused by a slight difference in the angle about the proline-induced kink in the helix (Fig. 6C).

The ligand-binding pocket is formed by both structurally conserved and divergent helices as compared to rhodopsin (Fig. 6D). Helices III and V are two of the most conformationally shifted helices and contain the canonical catecholamine binding residues associated with activation of adrenergic family of receptors (*63–65*). The comparison with rhodopsin suggests that the structurally conserved helices provide a common core present throughout the class A GPCRs, whereas the variable helices confer binding site plasticity with a resulting architecture capable of binding a large spectrum of ligands.

Comparison to rhodopsin-based GPCR models. Since the determination of the inactive dark-state rhodopsin structure (*18*), a number of homology models of other class A GPCRs have been reported (*66–70*). Typically, homology models start by alignment of so-called fingerprint motifs that are common among the family. These fingerprint motifs are extrapolated to assign coordinates for the entire helical bundle. Loop regions are either ignored or modeled based on databases of loop conformations depending on the application (*66*). A number of models exist for β_2AR, some of which have been improved upon with supporting biochemical data (*66*, *70–73*). When compared to the β_2AR structure reported here, however, all of these models were more similar to rhodopsin, as were models for other receptors (e.g. dopamine, muscarinic, and chemokine) (*28*). This is not entirely surprising but highlights a general shortcoming in homology models generated from a single structural template. The structural divergence between β_2AR and rhodopsin would be quite difficult to predict accurately using only rhodopsin as a template. The addition of a second class A GPCR structure should make it possible to correlate the sequence differences between rhodopsin and β_2AR with the observed structural differences and extrapolate to other class A GPCRs. Highlighting interactions that constrain class A receptors into each of the two observed states will allow a more comprehensive analysis of structural divergence and should result in more accurate models. Furthermore, evidence provided in the companion publication (*26*) indicates that β_2AR-T4L may not be in a completely inactive conformation like rhodopsin, providing an alternative signaling state on which to base homology models that will be more relevant for virtual ligand screening and structure-based drug design (*66*, *73*). The addition of further structural templates and conformational states to the pool of information on GPCRs should pave the way to a new generation of more potent therapeutics targeting this expansive receptor family and enhance our understanding of the signaling properties within their associated pathways.

References and Notes

1. S. Takeda, S. Kadowaki, T. Haga, H. Takaesu, S. Mitaku, *FEBS Lett.* **520**, 97 (2002).
2. R. Fredriksson, M. C. Lagerstrom, L. G. Lundin, H. B. Schioth, *Mol. Pharmacol.* **63**, 1256 (2003).
3. K. L. Pierce, R. T. Premont, R. J. Lefkowitz, *Nat. Rev. Mol. Cell Biol.* **3**, 639 (2002).
4. R. J. Lefkowitz, S. K. Shenoy, *Science* **308**, 512 (2005).
5. Y. Sun *et al.*, *EMBO J.* **26**, 53 (2007).
6. J. Drews, *Science* **287**, 1960 (2000).

7. B. Kobilka, *Annu. Rev. Neurosci.* **15**, 87 (1992).
8. M. G. Caron, R. J. Lefkowitz, *Recent Prog. Horm. Res.* **48**, 277 (1993).
9. A. D. Strosberg, *Protein Sci.* **2**, 1198 (1993).
10. L. Hein, B. K. Kobilka, *Trends Cardiovasc. Med.* **7**, 137 (1997).
11. D. K. Rohrer, *J. Mol. Med.* **76**, 764 (1998).
12. Y. Xiang, B. Kobilka, in *The Adrenergic Receptors in the 21st Century*, D. M. Perez, Ed. (Humana, Totowa, NJ, 2006), pp. 267–292.
13. G. Milligan, P. Svoboda, C. M. Brown, *Biochem. Pharmacol.* **48**, 1059 (1994).
14. M. R. Taylor, *Pharmacogenomics J.* **7**, 29 (2007).
15. T. R. Bai, *Lung* **170**, 125 (1992).
16. P. J. Barnes, *Life Sci.* **52**, 2101 (1993).
17. R. M. Smiley, M. Finster, *J. Matern. Fetal Med.* **5**, 106 (1996).
18. K. Palczewski *et al*., *Science* **289**, 739 (2000).
19. T. Okada *et al*., *J. Mol. Biol.* **342**, 571 (2004).
20. J. Li, P. C. Edwards, M. Burghammer, C. Villa, G. F. Schertler, *J. Mol. Biol.* **343**, 1409 (2004).
21. D. Salom *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16123 (2006).
22. B. K. Kobilka, X. Deupi, *Trends Pharmacol. Sci.* **28**, 397 (2007).
23. Inverse agonists act to reduce the basal activity of a receptor through interactions that shift the equilibrium to more of an inactive state. In contrast, antagonists bind to and block the active site, but do not affect the equilibrium between inactive and active states, and agonists shift the equilibrium to an active receptor state.
24. S. G. F. Rasmussen *et al*., *Nature*, published online 21 October 2007; 10.1038/nature06325.
25. β_2AR-T4L was generated by three distinct modifications to β_2AR: (1) a fusion protein was created by replacement of the third intracellular loop with T4L, (2) the carboxyl terminal 48 amino acids were deleted, and (3) a glycosylation site at Asn187 was eliminated through a glutamate substitution. This modified version was created to assist in improved crystal formation.
26. D. M. Rosenbaum *et al*., *Science*, in press; published online 25 October 2007 (10.1126/science.1150609).
27. E. M. Landau, E. Pebay-Peyroula, R. Neutze, *FEBS Lett.* **555**, 51 (2003).
28. See supporting material on *Science* Online.
29. M. Caffrey, *Curr. Opin. Struct. Biol.* **10**, 486 (2000).
30. J. Deisenhofer, H. Michel, *EMBO J.* **8**, 2149 (1989).
31. E. M. Landau, J. P. Rosenbusch, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14532 (1996).
32. Z. Yao, B. Kobilka, *Anal. Biochem.* **343**, 344 (2005).
33. V. Cherezov, A. Peddi, L. Muthusubramaniam, Y. F. Zheng, M. Caffrey, *Acta Crystallogr. D* **60**, 1795 (2004).
34. The successful diffraction screening and data collection that led to the structure determination of β_2AR-T4L required overcoming a number of technological barriers that encompassed the growth and harvest of microcrystals, crystal imaging, and the collection of diffraction data. Due to their transparency, crystals were often visually obstructed by the frozen lipidic mesophase material and, therefore, could not be confidently imaged by traditional beamline cameras, while their extremely small size made them susceptible to rapid radiation damage. Further details are provided in (*28*).
35. GPCRs are frequently post-translationally modified with palmitoylate on cysteine residues at the C-terminal tail. β_2AR-T4L was treated with iodoacetamide during purification to eliminate free thiols.
36. Ballesteros-Weinstein numbering is used throughout the text as superscripts to the protein numbering. Within each helix is a single most conserved residue among the class A GPCRs. This residue is designated x.50 where x is the number of the transmembrane helix. All other residues on that helix are numbered relative to this conserved position.
37. S. Yohannan, S. Faham, D. Yang, J. P. Whitelegge, J. U. Bowie, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 959 (2004).
38. M. Katragadda, M. W. Maciejewski, P. L. Yeagle, *Biochim. Biophys. Acta* **1663**, 74 (2004).
39. V. Cherezov, J. Clogston, M. Z. Papiz, M. Caffrey, *J. Mol. Biol.* **357**, 1605 (2006).
40. Membrane protein generally can form two types of crystal packing: Type I represents stacks of two dimensional crystals ordered in the third dimension via interactions of hydrophilic parts of membrane proteins. Type II crystals are composed of membrane proteins whose hydrophobic part is shielded by a detergent micelle and all crystal contacts are formed through hydrophilic, solvent exposed parts of protein molecules.
41. P. Nollert, H. Qiu, M. Caffrey, J. P. Rosenbusch, E. M. Landau, *FEBS Lett.* **504**, 179 (2001).
42. B. F. O'Dowd, M. Hnatowich, M. G. Caron, R. J. Lefkowitz, M. Bouvier, *J. Biol. Chem.* **264**, 7564 (1989).
43. G. Milligan, *Mol. Pharmacol.* **66**, 1 (2004).
44. S. Angers *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3684 (2000).
45. J. A. Javitch, *Mol. Pharmacol.* **66**, 1077 (2004).
46. J. F. Mercier, A. Salahpour, S. Angers, A. Breit, M. Bouvier, *J. Biol. Chem.* **277**, 44925 (2002).
47. G. F. Schertler, *Curr. Opin. Struct. Biol.* **15**, 408 (2005).
48. T. E. Hebert *et al*., *J. Biol. Chem.* **271**, 16384 (1996).
49. W. Guo, L. Shi, J. A. Javitch, *J. Biol. Chem.* **278**, 4385 (2003).
50. Y. Xiang, V. O. Rybin, S. F. Steinberg, B. Kobilka, *J. Biol. Chem.* **277**, 34280 (2002).
51. M. R. Whorton *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7682 (2007).
52. R. S. Ostrom, P. A. Insel, *Br. J. Pharmacol.* **143**, 235 (2004).
53. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037 (2001).
54. T. H. Ji, M. Grossmann, I. Ji, *J. Biol. Chem.* **273**, 17299 (1998).
55. U. Gether, *Endocr. Rev.* **21**, 90 (2000).
56. K. Noda, Y. Saad, R. M. Graham, S. S. Karnik, *J. Biol. Chem.* **269**, 6743 (1994).
57. J. Mialet-Perez, S. A. Green, W. E. Miller, S. B. Liggett, *J. Biol. Chem.* **279**, 38603 (2004).
58. K. Palczewski, *Annu. Rev. Biochem.* **75**, 743 (2006).
59. T. W. Schwartz, T. M. Frimurer, B. Holst, M. M. Rosenkilde, C. E. Elling, *Annu. Rev. Pharmacol. Toxicol.* **46**, 481 (2006).
60. L. Shi *et al*., *J. Biol. Chem.* **277**, 40989 (2002).
61. R. J. Lefkowitz, *Nat. Cell Biol.* **2**, E133 (2000).
62. For the alignment, residues on β_2AR were aligned to equivalent residues on Rhodopsin, respectively: 43-59 to 47-63; 67-95 to 71-99; 122-135 to 126-139; 285-296 to 264-275.
63. C. D. Strader *et al*., *J. Biol. Chem.* **263**, 10267 (1988).
64. C. D. Strader, M. R. Candelore, W. S. Hill, I. S. Sigal, R. A. Dixon, *J. Biol. Chem.* **264**, 13572 (1989).

65. G. Liapakis *et al*., *J. Biol. Chem.* **275**, 37779 (2000).
66. C. Bissantz, P. Bernard, M. Hibert, D. Rognan, *Proteins* **50**, 5 (2003).
67. A. Fano, D. W. Ritchie, A. Carrieri, *J. Chem. Inf. Model.* **46**, 1223 (2006).
68. J. V. Hobrath, S. Wang, *J. Med. Chem.* **49**, 4470 (2006).
69. M. Nowak, M. Kolaczkowski, M. Pawlowski, A. J. Bojarski, *J. Med. Chem.* **49**, 205 (2006).
70. Y. Zhang, M. E. Devries, J. Skolnick, *PLoS Comput. Biol.* **2**, e13 (2006).
71. P. L. Freddolino *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2736 (2004).
72. K. E. Furse, T. P. Lybrand, *J. Med. Chem.* **46**, 4450 (2003).
73. P. R. Gouldson *et al*., *Proteins* **56**, 67 (2004).
74. M. A. Lomize, A. L. Lomize, I. D. Pogozheva, H. I. Mosberg, *Bioinformatics* **22**, 623 (2006).
75. W. L. DeLano, The PyMOL Molecular Graphics System (2002) (www.pymol.org).
76. Author Contributions: RCS and BKK independently pushed the GPCR structural biology projects for more than 15 years. BKK managed the protein design, production and purification. RCS managed novel crystallization and data collection methods development and experiments. VC developed novel methods for, and performed LCP crystallization, LCP crystal mounting, LCP data collection, model refinement, analyzed the results, and was involved in manuscript preparation. DMR supplied protein materials for all crystallization trials, grew and collected data from the bicelle crystals, collected, processed and refined the 3.5 Å LCP structure, refined the 2.4 Å structure, analyzed the results, and was involved in manuscript preparation. MAH designed the blind crystal screening protocol and collected the 2.4 Å data set, processed the 2.4 Å data, solved the structure by MR at 3.5 Å and 2.4 Å resolution, wrote the initial draft of the manuscript and created all figures. SGFR assisted with the final stages of β_2AR-T4L purification. FST expressed β_2AR-T4L in insect cells and, together with TSK, performed the initial stage of β_2AR purification. HJC assisted with the refinement. PK assisted in developing novel methods to screen the transparent crystals, data collection, refinement, and was involved in manuscript preparation. WIW assisted with low resolution data collection and processing, solved the β_2AR-T4L molecular replacement problem at 3.5 Å, participated in the 2.4 Å refinement process, and participated in structure analysis and manuscript preparation. BKK additionally assisted with β_2AR-T4L purification, β_2AR-T4L 3.5 Å synchrotron data collection, structure analysis and manuscript preparation. BKK and DMR designed the β_2AR-T4L fusion protein strategy. RCS additionally assisted with β_2AR-T4L crystallization, 2.4 Å data collection, structure solution, refinement, structure analysis and manuscript preparation.

77. This work was supported in part by the NIH Roadmap Initiative grant P50 GM073197 and Protein Structure Initiative P50 GM62411 (to R.C.S.) and NIH Roadmap Initiative grant R21 GM075811 and NINDS Grant NS028471 (to B.K.K.). The authors acknowledge the support of Janet Smith, Robert Fischetti and Nukri Sanishvili at the GM/CA-CAT beamline at the Advanced Photon Source, for assistance in development and use of the minibeam and beamtime. The GM/CA-CAT beamline (23-ID) is supported by the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-1104). D.M.R. was supported in part by NIH grant F32 GM082028; S.G.F.R. was supported in part by the Lundbeck Foundation; H.-J.C. and W.I.W. were supported in part by NIH grant R01 GM056169. We thank Gebhard Schertler for help with the initial diffraction experiments on LCP crystals performed at ID-13 at the ESRF. The authors acknowledge Chris Roth, Veli-Pekka Jaakola, Alexander Alexandrov, Ellen Chien, Michael Bracey, Vsevolod Katrich, Ian Wilson and Mark Yeager for careful review of the manuscript, Yuan Zheng and Martin Caffrey from The Ohio State University for the generous loan of the *in meso* robot, and Angela Walker for assistance with manuscript preparation. Coordinates and structure factors have been deposited in the Protein Data Bank with identification code 2RH1.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1150577/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S3
References

17 September 2007; accepted 11 October 2007
Published online 25 October 2007; 10.1126/science.1150577
Include this information when citing this paper.

Fig. 1. Overall fold of the β_2AR-T4L fusion with its predicted orientation in the plasma membrane and key intramolecular interactions. (**A**) Stereoview of the overall fold of β_2AR-T4L. The receptor and T4L are colored gray and green, respectively. Carazolol is colored blue and the lipid molecules bound to the receptor are colored yellow. (**B**) The receptor is aligned to a rhodopsin model that was positioned in a lipid membrane (boundaries indicated by horizontal black lines) as found in the orientations of proteins in membranes (OPM) database (*74*). T4L is fused internally into the third intracellular loop of β_2AR and maintains minimal intramolecular packing interactions by tilting away from the receptor. (**C**) Specific intramolecular interactions between β_2AR and T4L are represented.

Fig. 2. Crystal packing interactions in the lipidic mesophase crystallized β_2AR-T4L. (**A**) There are four main contact areas, two of which are mediated by T4L in the plane of the membrane with itself through a two-fold symmetry axis and translation. The third interaction is normal to the membrane plane between T4L and lumen exposed loops of β_2AR. The fourth interaction is generated by the two-fold symmetry axis, packing one receptor to receptor in the plane of the membrane. (**B**) The receptor crystal packing interface is composed mainly of lipids with two cholesterol molecules and two palmitic acid molecules forming the majority of the interactions. A network of ionic charge interactions exists on the cytoplasmic end of the interface forming the only inter-receptor protein contacts. (**C**) Comparison between β_2AR-T4L and rhodopsin (PDB ID Code 2I35) parallel receptor association interface. Helices I (blue) and VIII (magenta) are highlighted in both structures. Only one monomer is shown for each receptor representation along with helices I' and VIII' only from the opposing symmetry related molecule. The rhodopsin interface is twisted significantly relative to β_2AR-T4L resulting in a significant offset from the parallel orientation required for a physiological dimer interface.

β_2AR-T4L associated monomers are in a highly parallel orientation.

Fig. 3. Surface representation of β_2AR colored by calculated charge from red ($-10\ k_bT/e_c$) to blue ($+10\ k_bT/e_c$) using a dielectric constant of 70. (**A**) Three main areas of interest are indicated. The binding site cleft is negatively charged as is a groove between helices III, IV and V. The third region is an overall positive charge in the region of the ionic lock and DRY motif on the cytoplasmic face. The overall result is a highly polarized molecule that may utilize its negative charge to facilitate binding of catecholamine ligands. The presence of a negative charge in the groove between helices III, IV and V is unexpected as it is in the middle of the lipid membrane. This charge may be partially derived from the presence of an unpaired glutamate at position $122^{3.41}$. The effective charge in this region is likely greater than shown here due to its location in the low dielectric environment of the lipid membrane. (**B**) View rotated 90° from A. Showing both the negatively charged binding site cleft (top) and positively charged cytoplasmic face (bottom). Poisson-Boltzmann electrostatics were calculated using the program APBS (*53*) as implemented in Pymol (*75*). Pymol was used exclusively in the preparation of all figures.

Fig. 4. Comparison of the extracellular sides of β_2AR-T4L and rhodopsin. (**A**) The N-terminus is missing from the experimental density in the β_2AR-T4L structure and is not shown. ECL2 is shown in green and contains a short α-helix and two disulfide bonds (yellow). The intraloop disulfide bond constrains the tip of ECL2 which interacts with ECL1. The second disulfide bond links ECL2 with helix III. There is one interaction between ECL2 and carazolol (blue) through $Phe193^{5.32}$. The entire loop is held out of the ligand binding site by a combination of the rigid helical segment and the two disulfide bonds. (**B**) In contrast, ECL2 (green) in rhodopsin assumes a lower position in the structure that occludes direct access to the retinal-binding site and forms a small β-sheet in combination with the N-terminal region (magenta) directly above the bound retinal (pink).

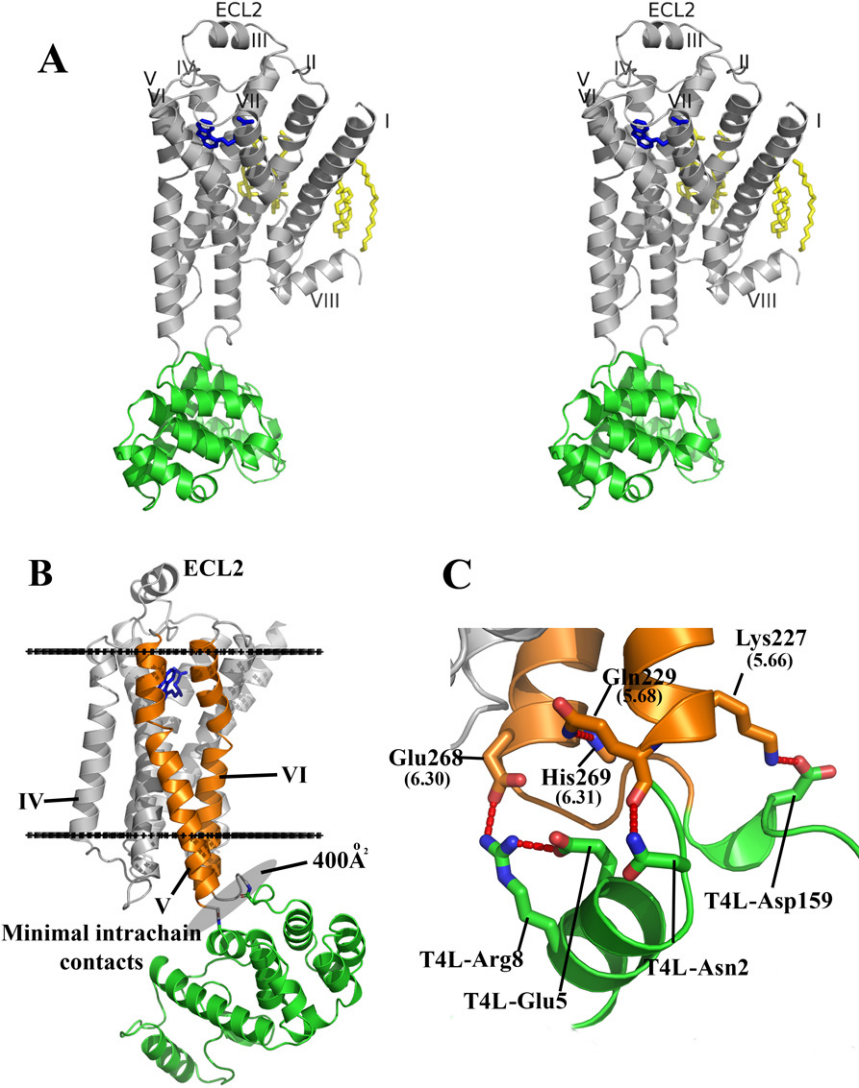
Fig. 5. Ligand binding characterization and comparison to rhodopsin. (**A**) A view looking down on the plane of the membrane from the extracellular surface showing a detailed representation of the carazolol binding site in β_2AR-T4L. Carazolol is shown as sticks with carbon atoms colored yellow. β_2AR-T4L residues contributing to carazolol binding are shown in green and labeled. Electron density is contoured at 5σ from an F_o-F_c omit map calculated without the contribution of carazolol. (**B**) Binding orientation comparison between 11-*cis*-retinal in rhodopsin and carazolol in β_2AR-T4L. Van der Waals' surfaces for carazolol and retinal are represented as dots to accentuate the close packing interactions. Retinal in the all-*cis* conformation (pink), binds deep in the active site of rhodopsin as compared to carazolol (blue), packing its β-ionone ring between $Tyr268^{6.51}$ and $Phe212^{5.47}$ (cyan), blocking movement of $Trp265^{6.48}$ (magenta) into the space. The β-ionone ring of *trans*-retinal in activated rhodopsin would not block $Trp265^{6.48}$ from rotating into the space allowing a rotameric shift into its proposed active form. (**C**) There are four residues involved in the toggle switch mechanism of β_2AR-T4L as shown. $Phe290^{6.52}$ (magenta) is sandwiched between $Phe208^{5.47}$ (tan) and $Phe289^{6.51}$ (tan) forming a ring-face aromatic interaction. Like rhodopsin, an activation step is thought to occur by a rotameric change of $Trp286^{6.48}$ (magenta) which would displace $Phe290^{6.52}$. Carazolol is shown to interact extensively with the sandwich motif as shown: however, few interactions are seen with $Trp286^{6.48}$. The 6.52 position in β_2AR-T4L is occupied by $Phe290^{6.52}$ as opposed to $Ala269^{6.52}$ in rhodopsin where the β-ionone ring replaces an aromatic protein side chain in forming the sandwich interactions. The aromatic character of the sandwich is otherwise maintained by $Phe289^{6.51}$ and $Phe208^{5.47}$ in β_2AR-T4L.

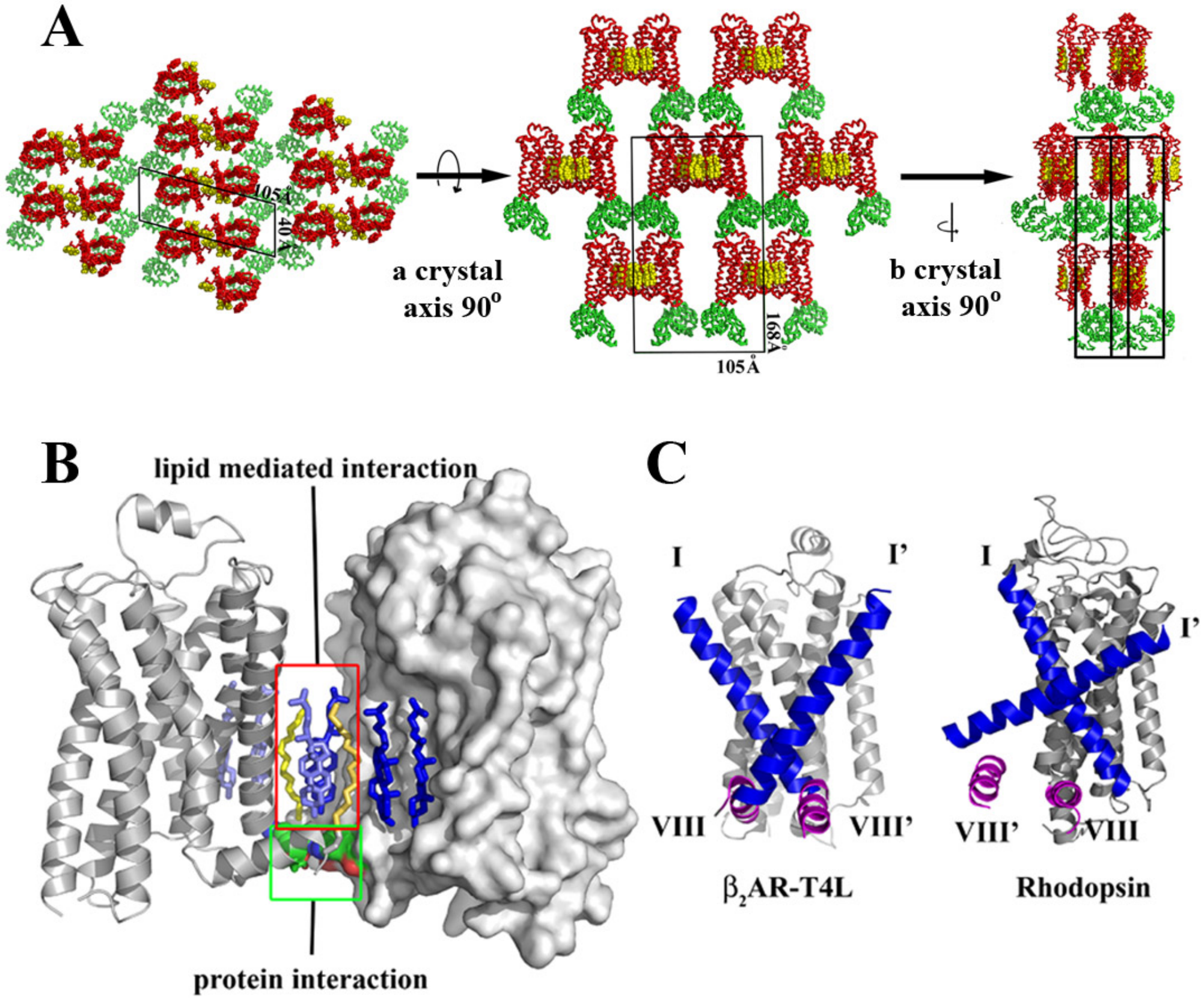
Fig. 6. Comparison of β_2AR-T4L helical orientations with rhodopsin (PDB ID Code 1U19). (**A**) β_2AR-T4L is rendered as a ribbon trace colored with a blue to red spectrum corresponding to observed distances between Cα positions in the two structures (RMSD 2.7 Å between all residues in the transmembrane region). Helix II shows very little movement, whereas the entire lengths of helices III, IV, V shift significantly. Helix VIII and loops were not included in the comparison and are colored in tan. (**B**) Movements of helices I and V of rhodopsin (grey) are shown relative to β_2AR-T4L. (**C**) Movements of helices III, IV and VI. (**D**) Ligand binding site representation. Carazolol is shown with yellow carbons. Entire helices are assigned a single designation based on their divergence from the rhodopsin position in the area of the ligand binding site as shown. Helix I is highly divergent, Helices II and VI are similar to rhodopsin. Helices IV and VII are moderately constant. Helices III and V are moderately divergent.

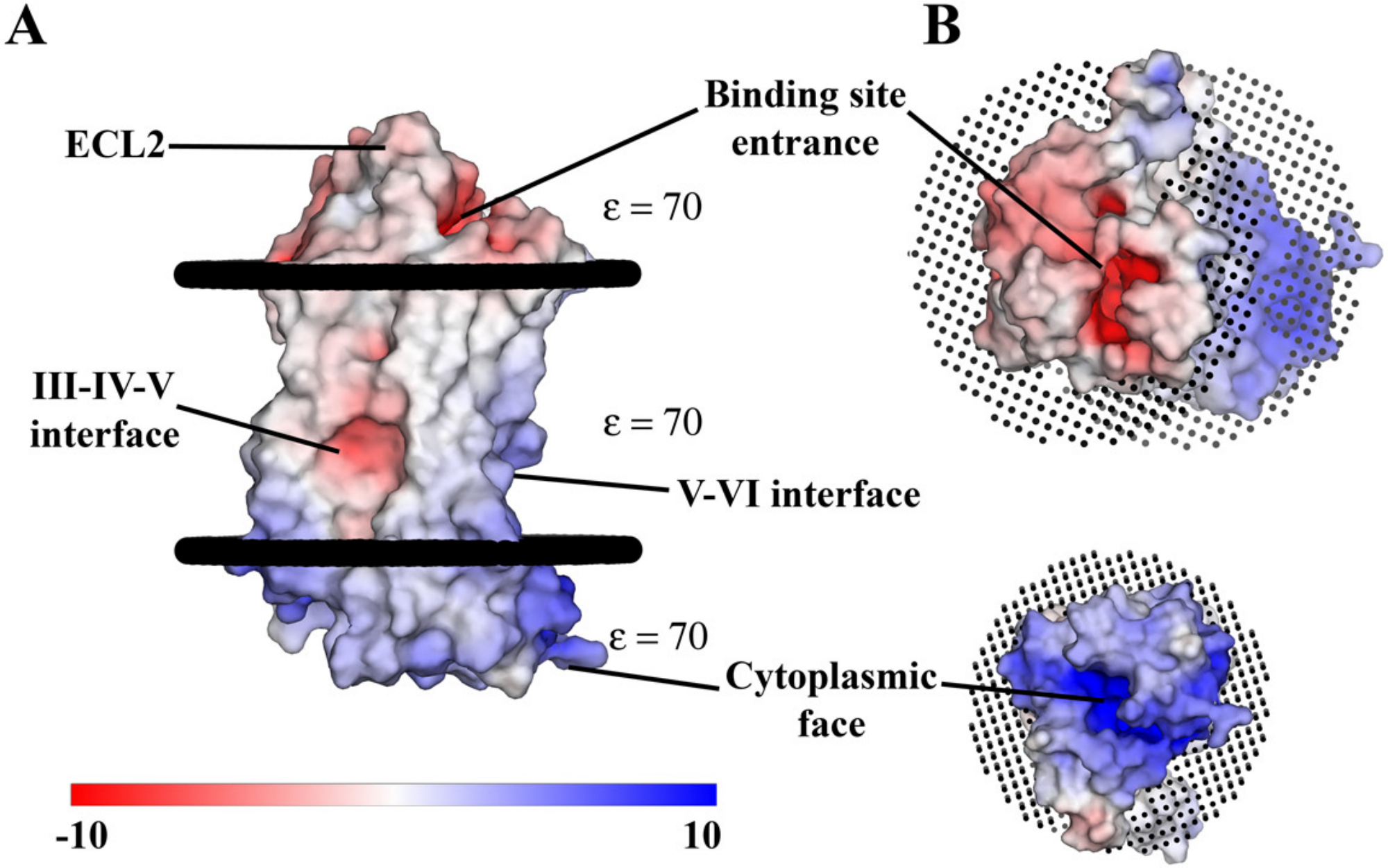
Table 1. Data collection and refinement statistics.

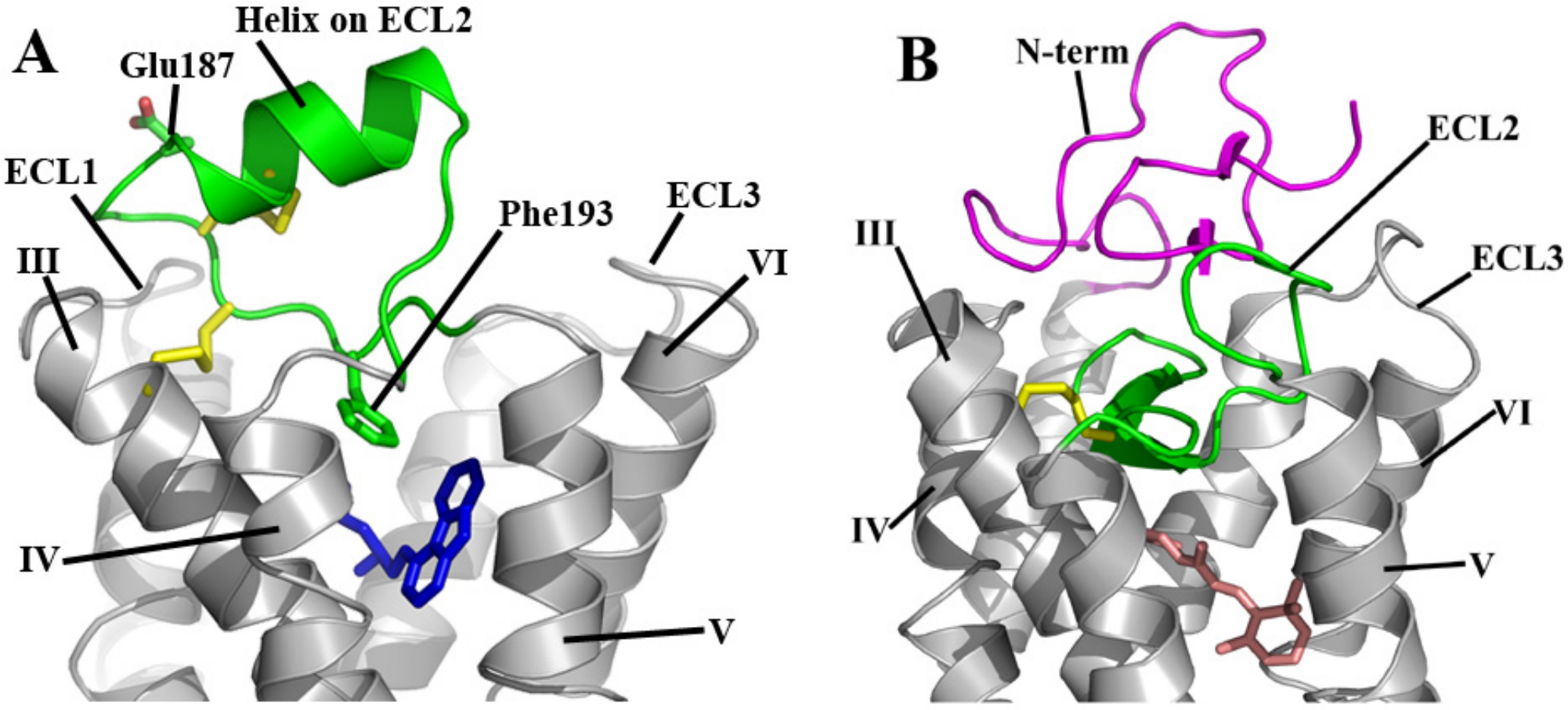
	β_2AR-T4L
Data collection (APS GM/CA CAT 23ID-B, 10 µm beam)*	
Space group	C2
Cell dimensions	
a, *b*, *c* (Å)	106.32, 169.24, 40.15
β (°)	105.62
No. of reflections processed	245,571
No. unique reflections	26,574
Resolution (Å)	50 – 2.4 (2.5 – 2.4)
R_{sym}	12.7 (67.8)
Mean I/σ(I)	9.6 (2.2)
Completeness (%)	99.5 (99.1)
Redundancy	9.4 (4.8)
Refinement*	
Resolution (Å)	20 – 2.4 (2.46 – 2.4)
No. reflections (test set)	25,247 (1,310)
R_{work} / R_{free}	19.8(27.0) / 23.2(30.1)
No. atoms	3,805
Protein	3,544
Ions, lipids, ligand and other	213
Water	48
Overall *B*-values (Å^2)	82
β_2AR	77
T4-Lysozyme	75
Carazolol	55
Lipid	100
R.m.s deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.5
Ramachandran plot statistics (%) (excl. Gly, Pro):	
Most favored regions	94.8
Additionally allowed regions	5.0
Generously allowed regions	0.2
Disallowed regions	0

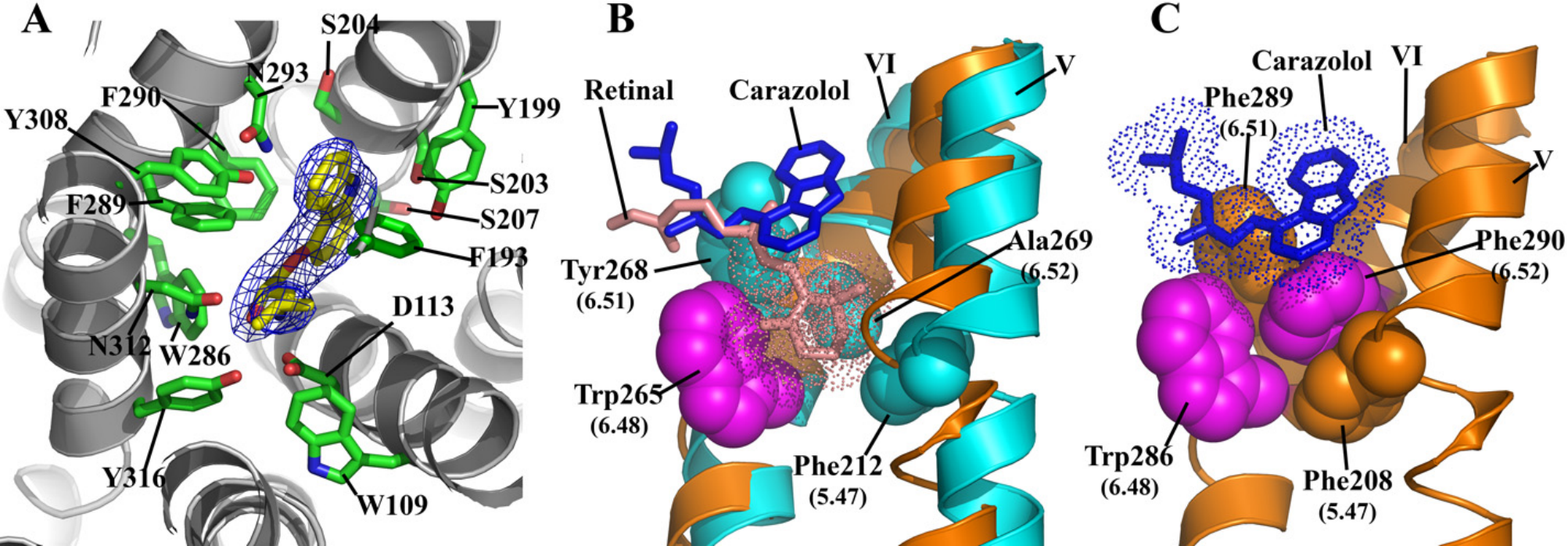
*Highest resolution shell is shown in parenthesis.
$R_{sym} = \Sigma_{hkl} |I(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}(hkl)$, where $\langle I(hkl)\rangle$ is the mean of the symmetry equivalent reflections of I(*hkl*).

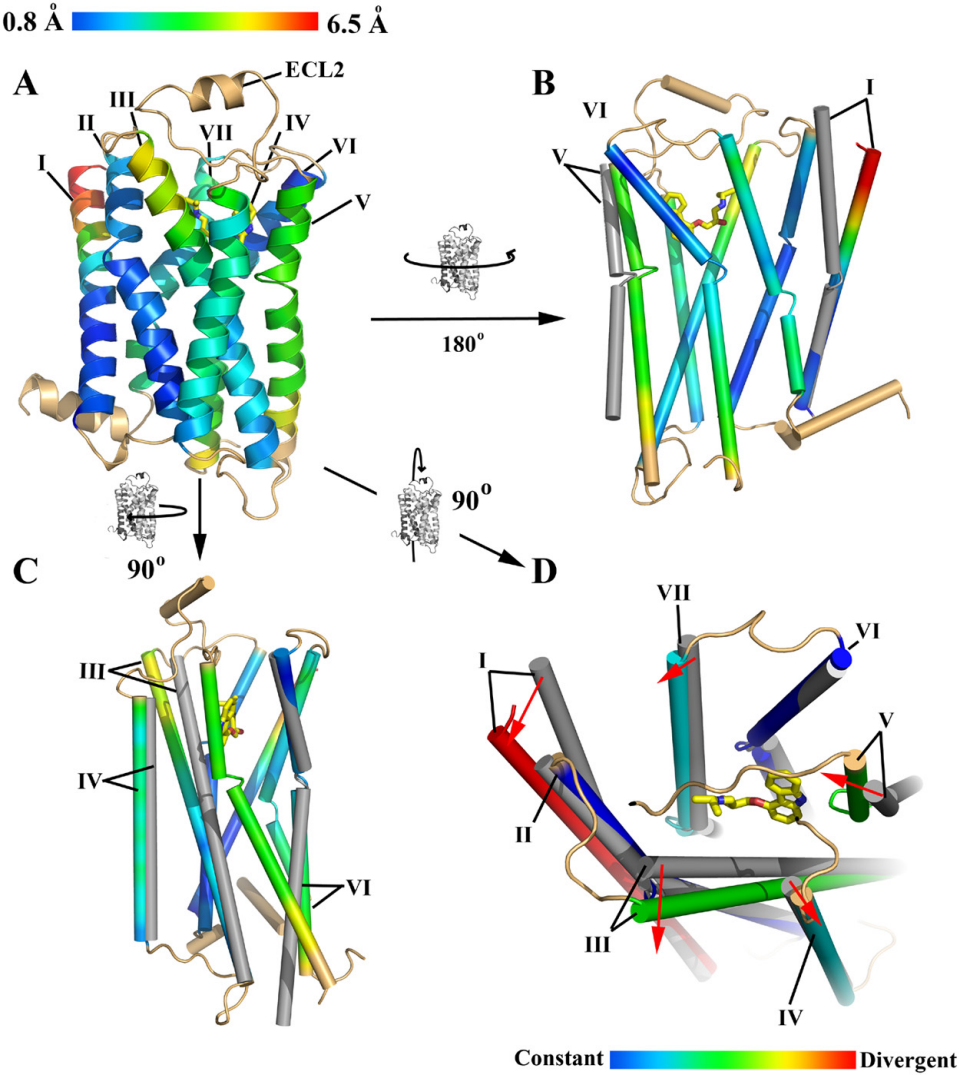

A
ECL2
III
II
IV
V
VI
VII
I
VIII
ECL2
III
II
IV
V
VI
VII
I
VIII
B
ECL2
IV
VI
V
400Å²
Minimal intrachain contacts
C
Lys227 (5.66)
Gln229 (5.68)
Glu268 (6.30)
His269 (6.31)
T4L-Asp159
T4L-Arg8
T4L-Asn2
T4L-Glu5


A
105Å
40Å
a crystal
axis 90°
168Å
105Å
b crystal
axis 90°
B
lipid mediated interaction
protein interaction
C
I
I'
VIII
VIII'
β2AR-T4L
I
I'
VIII'
VIII
Rhodopsin


A
B
ECL2
Binding site entrance
ε = 70
III-IV-V interface
ε = 70
V-VI interface
ε = 70
Cytoplasmic face
-10
10


A
Helix on ECL2
Glu187
ECL1
III
Phe193
ECL3
VI
IV
V
B
N-term
ECL2
III
ECL3
VI
IV
V


A
S204
N293
F290
Y308
Y199
S203
F289
S207
F193
D113
N312
W286
Y316
W109
B
Retinal
Carazolol
VI
V
Tyr268
(6.51)
Ala269
(6.52)
Trp265
(6.48)
Phe212
(5.47)
C
Carazolol
VI
Phe289
(6.51)
V
Phe290
(6.52)
Trp286
(6.48)
Phe208
(5.47)


0.8 Å
6.5 Å
A
ECL2
III
II
VII
IV
VI
I
V
180°
B
VI
I
V
90°
90°
C
III
IV
VI
D
VII
VI
I
V
II
III
IV
Constant
Divergent