Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser

Yanyong Kang*, X. Edward Zhou*, Xiang Gao*, Yuanzheng He*, Wei Liu, Andrii Ishchenko, Anton Barty, Thomas A. White, Oleksandr Yefanov, Gye Won Han, Qingping Xu, Parker W. de Waal, Jiyuan Ke, M. H. Eileen Tan, Chenghai Zhang, Arne Moeller, Graham M. West, Bruce D. Pascal, Ned Van Eps, Lydia N. Caro, Sergey A. Vishnevetsky, Regina J. Lee, Kelly M. Suino-Powell, Xin Gu, Kuntaal Pal, Jinming Ma, Xiaoyong Zhi, Sebastien Boutet, Garth J. Williams, Marc Messerschmidt, Cornelius Gati, Nadia A. Zatsepina, Dingjie Wang, Daniel James, Shibom Basu, Shatabdi Roy-Chowdhury, Uwe Weierstall, Oliver P. Ernst, Vsevolod Katritch, Vsevolod V. Gurevich, Patrick R. Griffin, Wayne L. Hubbell, Raymond C. Stevens, Vadim Cherezov, Karsten Melcher & H. Eric Xu

G-protein-coupled receptors (GPCRs) signal primarily through G proteins or arrestins. Arrestin binding to GPCRs blocks G-protein interaction and redirects signalling to numerous G-protein-independent pathways. Here we report the crystal structure of a constitutively active form of human rhodopsin bound to a pre-activated form of the mouse visual arrestin, determined by serial femtosecond X-ray laser crystallography. Together with extensive biochemical and mutagenesis data, the structure reveals an overall architecture of the rhodopsin–arrestin assembly in which rhodopsin uses distinct structural elements, including transmembrane helix 7 and helix 8, to recruit arrestin. Correspondingly, arrestin adopts the pre-activated conformation, with a ~20° rotation between the amino and carboxyl domains, which opens up a cleft in arrestin to accommodate a short helix formed by the second intracellular loop of rhodopsin. This structure provides a basis for understanding GPCR-mediated arrestin-biased signalling and demonstrates the power of X-ray lasers for advancing the frontiers of structural biology.

G-protein-coupled receptors (GPCRs) comprise the largest family of cell surface receptors, which signal primarily via G proteins or arrestins. Upon activation, GPCRs recruit heterotrimeric G proteins and subsequently G-protein-coupled receptor kinases (GRKs), which phosphorylate GPCRs to allow the high-affinity binding to arrestin. Arrestin binding to the receptors blocks their interactions with G proteins and leads to the receptor’s desensitization. The binding of arrestins to GPCRs also initiates numerous cellular signalling pathways that are independent of G proteins. Arrestin-mediated signalling is therefore a central component of the GPCR functional network.

GPCRs are targets of one-third of the currently used drugs. Recent studies have demonstrated that G-protein and arrestin pathways are distinct and can be pharmacologically modulated independently using biased GPCR ligands. Biased GPCR ligands are often preferred over unbiased agonists and antagonists, as they selectively direct the receptor to a subset of partners and can deliver therapeutic benefits with fewer undesirable side effects. Research towards biased ligands has become a new trend for GPCR-targeting therapeutics.

The molecular mechanisms of GPCR signalling have been unravelled by recent breakthroughs in GPCR structural biology. In the antagonist-bound state, GPCRs assume a closed conformation with the cytoplasmic ends of the transmembrane (TM) helices packed closely with each other, thus blocking the interactions with G proteins or arrestins. In contrast, agonist binding promotes conformational changes in GPCRs, including a dramatic movement within the cytoplasmic side...
of the TM domain, allowing activated receptors to recruit G proteins or arrestins to mediate downstream signalling. However, arrestin coupling to GPCRs may require a conformation of the receptor different from that required for coupling with G proteins.

Rhodopsin is a prototypical GPCR responsible for light perception. Along with the β2-adrenergic receptor (β2AR), rhodopsin has served as a model system for studying GPCR signalling. Figure 1a shows rhodopsin binding to G protein and arrestin. Light induces isomerization of 11-cis retinal to all-trans-retinal (ATR), which activates rhodopsin and promotes its interactions with G protein (GRK1), leading to high affinity recruitment of arrestin that terminates the G protein signalling. Activation of rhodopsin can also be achieved through mutations, including the E113Q and M257Y mutation, which yields a constitutively active rhodopsin (superscripts in residues refer to the Ballesteros–Weinstein numbering). The crystal structure of bovine rhodopsin has been determined in the inactive, resting state, the ligand-free state, and the ligand-activated state in complex with a G-protein peptide. Arrestin structures have been determined in the inactive and pre-activated form. Recent electron microscopy analysis has revealed the assembly and conformational dynamics of the β2AR–β-arrestin complex. Here we report the crystal structure of an active form of rhodopsin bound to a pre-activated mouse visual arrestin, determined by serial femtosecond crystallography (SFX). The structure has been confirmed by electron microscopy, double electron-electron resonance (DEER) spectroscopy, hydrogen–deuterium exchange mass spectrometry (HDX), cell-based rhodopsin–arrestin interaction assays, and site-specific disulfide cross-linking experiments. Our study provides a molecular basis for understanding GPCR-mediated arrestin-biased signalling.

Characterization and crystallization

To characterize the rhodopsin–arrestin interaction, we expressed and purified E113Q and E113Q/M257Y mutant receptors (Extended Data Fig. 1a). These mutations were introduced in the context of the N2Ser/C282E/C367Y mutant that is known to create a disulfide bond that increases rhodopsin stability without affecting its activity. Wild-type arrestin has weak background binding to wild-type rhodopsin (Fig. 1b). The E113Q mutation increased wild-type arrestin binding by twofold to threefold, and the E113Q/M257Y mutation further increased the binding of wild-type arrestin in the presence of all-trans-retinal (fourfold to eightfold). In contrast to the relatively weak binding of wild-type arrestin, the binding of 3A arrestin, a pre-activated form of arrestin that obviates the need for receptor phosphorylation for high affinity binding through three alanine mutations in L374, V375, and F376 in the C-terminal tail of arrestin, is much stronger. In the absence of all-trans-retinal, we observed a nearly 50-fold increase of 3A arrestin binding to the E113Q/M257Y receptor. All-trans-retinal further increased 3A arrestin binding to the E113Q/M257Y receptor by ∼60-fold above the binding of wild-type arrestin to wild-type rhodopsin (Fig. 1b and Extended Data Fig. 1b).

We also measured rhodopsin–arrestin interactions using AlphaScreen assays (Extended Data Fig. 1c) with His-tagged rhodopsin and biotin-tagged arrestin. Wild-type arrestin interacted weakly with the E113Q/M257Y rhodopsin, regardless of the presence of all-trans-retinal (Fig. 1c). As a positive control, the GzCT-HA peptide, a high affinity peptide variant of the C terminus of G-transducin (Gt), readily interacted with the E113Q/M257Y receptor in the absence of all-trans-retinal, and addition of all-trans-retinal slightly increased this interaction (Extended Data Fig. 1d). Quantitative competition using unlabelled 3A arrestin or GzCT-HA with the E113Q/M257Y receptors revealed an IC50 value of 15 nM and 700 nM for the binding of 3A arrestin and the GzCT-HA peptide, respectively (Fig. 1d and Extended Data Fig. 1d). The strength of the interaction between the E113Q/M257Y rhodopsin and the 3A arrestin is in a similar range as the estimated Kd value of 30–80 nM for the binding of arrestin to the fully activated phosphorylated rhodopsin.

Mixing individually purified proteins did not yield a stable 1:1 complex, nor did it lead to crystallization. Extensive biochemical data support a 1:1 stoichiometry in the rhodopsin–arrestin complex. Therefore, we engineered a fusion protein in which 3A arrestin is linked by a 15-residue linker to the C terminus of E113Q/M257Y rhodopsin. We expressed and purified the rhodopsin–arrestin fusion protein, as well as a T4 lysozyme (T4L)–rhodopsin–arrestin fusion, in which a T4L is fused to the N terminus of rhodopsin to increase the soluble surface for crystallization (Extended Data Fig. 1e).
Diffraction patterns from 18,874 crystals could be indexed and integrated using CrystFEL. The data were processed according to the apparent tetragonal lattice with a large unit cell (a = b = 109.2 Å and c = 452.6 Å). The diffraction was anisotropic with resolution limits of 3.8 Å and 3.3 Å along the a* /b* and c* axes, respectively (Supplementary Table 1).

The crystals appeared to be pseudo-merohedrally twinned in P212121 (Supplementary Table 2) and the structure was solved by molecular replacement using known structures of active rhodopsin and pre-activated arrestin (details in Methods). The structure contains four rhodopsins (residues 1–326), four arrestins (residues 12–361 with a small missing loop of residues 340–342), and three T4Ls (residues 2–161 in complexes A and D; residues 2–12 and 58–162 in complex C; no T4L was modelled in complex B owing to poor density) (Fig. 2). The final structure was refined to Rwork and Rfree of 25.2% and 29.3%, respectively, with excellent geometry (Supplementary Table 1b). The overall arrangement of the T4L–rhodopsin–arrestin complex is well supported by the electron density maps (Extended Data Fig. 3), including a 3,000 K simulated annealing omit map. Because of the twinned nature of the data sets, we performed extensive structure-validation experiments, including DEER, HDX, cell-based rhodopsin–arrestin interaction assays and site-specific disulfide cross-linking. Below we describe the rhodopsin–arrestin structure and the results of validation experiments.

Overall structure of the rhodopsin–arrestin complex

The most striking feature of the rhodopsin–arrestin complex is the asymmetric binding of arrestin to rhodopsin (Fig. 2) and this asymmetric arrangement is similar in all four complexes in the asymmetric unit, providing an independent confirmation of the rhodopsin–arrestin complex assembly (Extended Data Fig. 4). Figure 2a shows one rhodopsin–arrestin complex in four 90° orientations. From the intracellular (IC) view, rhodopsin and arrestin have similar heights, but the width of arrestin is nearly three times that of rhodopsin. Figure 2b shows the rhodopsin–arrestin complex in a transparent surface, whose overall arrangement of the domains can be fit into the electron microscopy images (Fig. 1e). Figure 2c shows the layered or type I packing of the complex in the crystal lattice with alternating hydrophilic and hydrophobic layers comprising arrestin, T4L and rhodopsin, respectively (Fig. 2c). This arrangement allows the complex to form extensive packing interactions that involve all soluble portions of the proteins, with the arrestin being the central mediator for packing with T4L, rhodopsin and arrestin from neighbouring symmetry-related molecules.

To validate the assembly of the rhodopsin–arrestin complex, we used DEER to determine intermolecular distances within the complex. The DEER distances from residue Y74 in rhodopsin to three arrestin residues (T61, V140, and S241) measured in a non-fused rhodopsin–arrestin complex were 28 Å, 23 Å and 33 Å, closely matching the distances of 28 Å, 22 Å and 34 Å, respectively, as observed in the crystal structure (Fig. 3). The intramolecular distances in the active arrestin bound to light-activated phosphorylated rhodopsin have also been studied extensively by DEER and all of them match exceedingly well with the crystal structure (Supplementary Table 3). Together, these data support the conclusion that the complex formed by fusion proteins closely resembles the physiologically relevant complex formed by individual proteins.

The rhodopsin–arrestin interface

The four rhodopsin–arrestin complexes in the asymmetric unit adopt nearly identical interfaces (Extended Data Fig. 4a), which are stabilized by intermolecular interactions as summarized in Supplementary Table 4. The total surface area buried in the interface is 1,350 Å², which is substantially smaller than the area (2,576 Å²) buried in the β2AR–Gs complex. Unlike the continuous interface observed in the β2AR–Gs complex, the rhodopsin–arrestin complex has four distinct
The first arrestin interface patch consists of the finger loop (residues Q70 to L78), which adopts a short α-helix and forms extensive interactions with the C terminus of TM7 and the N terminus of helix 8, as well as the loop residues (ICL1) of rhodopsin (Fig. 5a). Interactions of arrestin with TM7 of rhodopsin are of particular interest because conformational changes in TM7 have been implicated in arrestin activation and receptor binding. The second arrestin interface patch is formed by the middle loop (residue V140 region) and the C-loop (residue Y251 region at the central loop in the arrestin C domain) that interact with the ICL2 of rhodopsin, and the arrestin back loop (R319 and T320) that interacts with the C terminus of TM5. The middle and C-loops are close to each other in the inactive arrestin, but move apart upon its activation to form a cleft that accommodates the ICL2 of rhodopsin, which adopts a short helix (Fig. 4a, b). The positions of the finger loop and the C-loop are supported by a composite omit 2Fo − Fo electron density map (Extended Data Fig. 3a, d). The third arrestin interface patch is the β-strand (residues 79–86), which follows the finger loop and interacts with residues from TM5, TM6 and ICL3 of rhodopsin. The fourth arrestin putative interface patch is mostly between its N-terminal β-strand (residues 11–19) and the C-terminal tail of rhodopsin, which was not visible in the electron density map owing to the apparent flexibility of this region, but was computationally modelled based on HDX and disulfide cross-linking data described below (Extended Data Fig. 5). Consistent with the crystal structure, these arrestin elements have been implicated in various aspects of arrestin activation and receptor binding.

To further characterize the rhodopsin–arrestin interfaces, we performed three additional sets of validation experiments. The first was HDX, which probes the dynamics and stability of protein complexes. Compared with free arrestin, the rhodopsin-bound arrestin has several regions that are protected from exchange, including the finger loop and the N-terminal β-sheets, consistent with their location in the rhodopsin-binding interface (Fig. 4c, d and Extended Data Fig. 6a). The hydrogen to deuterium exchange rate of arrestin in the complex is lower than that for free arrestin across the whole protein, indicating that arrestin is stabilized by complex formation, consistent with the results of previous HDX experiments and thermal stability assays, which revealed that the melting temperature of free arrestin is six degrees lower than that of the complex (53 °C versus 59 °C, Extended Data Fig. 6b).

The second set consisted of Tango assays, which have been used for probing GPCR–arrestin interactions (Extended Data Fig. 7a). Wild-type rhodopsin and wild-type arrestin had a very low basal interaction and all-trans-retinal increased the binding by approximately threefold. In contrast, E113<sup>S</sup>Q/M257<sup>Y</sup> rhodopsin showed a high level of interaction with the pre-activated 3A arrestin, and addition of all-trans-retinal further increased the binding signal by approximately fivefold. Mutations in finger loop (D74, M76, G77, and L78), middle loop (Q134 and D139), and C-loop (L250 and Y251) decreased rhodopsin–arrestin binding (Extended Data Fig. 7b). Correspondingly, mutations in rhodopsin residues involved in arrestin binding also weakened the interaction (Extended Data Fig. 7c), consistent with the complex crystal structure.

The third set consisted of site-specific disulfide cross-linking experiments, which have been used to validate structures based on the geometry requirements for disulfide bond formation (Cα–Cα

---

**Figure 3** | DEER validation of rhodopsin–arrestin complex assembly. a, An overall view of rhodopsin–arrestin assembly showing the three intermolecular distances based on the models of the R1 nitroxide pairs at rhodopsin residue Y74<sup>41</sup> and three arrestin residues T61, V140, and S241 based on the crystal structure. b–d, The experimental distance distributions between the nitroxide spin labelled R1 pairs of rhodopsin Y74<sup>41</sup> and bovine arrestin S60, V139, and L240, which are in equivalent positions to mouse arrestin T61(b), V140(c), and S241(d) as labelled in the figure.

**Figure 4** | The rhodopsin–arrestin interface and its validation by HDX. a, b, Two overall views showing the four interface patches of the rhodopsin–arrestin complex. c–e, Mapping of HDX on the rhodopsin-bound arrestin structure. Rhodopsin is coloured in red and arrestin is coloured based on the exchange rate differences between free 3A arrestin and rhodopsin-bound arrestin as shown in Extended Data Fig. 6a. This figure was made using a computational model of the full rhodopsin–arrestin complex.
Figure 5 | Validation of the rhodopsin–arrestin interface by disulfide bond cross-linking. a–e, Structure and cross-linking of arrestin with rhodopsin. Panels are arrestin finger loop with rhodopsin TM7 and helix 8 (a); arrestin middle loop with rhodopsin ICL2 (b); arrestin C-loop residue Y251 with rhodopsin TM5 (c); arrestin β-strand interface residues with residues of rhodopsin TM5, ICL3, and TM6 (d); and arrestin’s N terminus with rhodopsin’s C-tail (e). Rhodopsin K311 is marked with a red asterisk and the side chain of arrestin M76 is shown in full from computation modelling of the full rhodopsin–arrestin complex, which was also used in panel e. Black asterisks, arrestin; arrowheads, rhodopsin/arrestin crosslinking adduct.

distances of 5–9Å and appropriate side-chain orientations). We engineered cysteine pairs at the binding interface of arrestin and rhodopsin, which were tagged with Flag and HA, respectively. Over 314 co-expression combinations were tested and monitored by SDS–PAGE followed by western blotting (Extended Data Fig. 8). The results are summarized in Supplementary Table 5. Every interface residue in arrestin was included in the study and the results closely agree with the crystal structure. For example, the distances from the Cβ atom of the finger loop residue G77 of arrestin to the Cβ atoms of N310, K311, and K312 in rhodopsin fit the requirement for disulfide bond formation (Fig. 5a). G77C cross-linked efficiently with N310, K311, and K312, but not with K311 and K312, because the Cβ of K311 points away from G77 (Fig. 5a). Neither did G77C show cross-linking with a large set of other rhodopsin residues, indicating the high specificity of the cross-linking experiments (Extended Data Fig. 8c and Supplementary Table 5). In contrast, several other mutants in the finger loop region (D74C, M76C, and L78C) readily cross-linked with Q312C from helix 8 (Fig. 5a). The cross-linking results of these four finger loop residues not only matched the crystal structure, but also agreed well with the results from the Tango assays (Extended Data Fig. 7b). In addition, mutants of three N-terminal finger loop residues (Q70C, E71C, and D72C) were cross-linked to mutants in rhodopsin ICL1 T79 and K67C, respectively (Extended Data Fig. 7d).

We also observed cross-linking of the arrestin middle loop (D139) with rhodopsin ICL2 (G149C) (Fig. 5b), of the arrestin C-loop (Y251) with rhodopsin TM5 (T229 and A233) (Fig. 5c), and of the arrestin β-strand (residues 79–86) that follows the finger loop with rhodopsin TM5, TM6, and ICL3 (Fig. 5d). Additional cross-linking was observed in two back-loop residues R319C and T320C of arrestin (Extended Data Fig. 8e). Furthermore, extensive cross-linking of the arrestin N terminus with the C-terminal tail of rhodopsin was detected, including R19 of arrestin with S334C, K311 of rhodopsin (Fig. 5e), of the arrestin C-loop with S333C and K339C of rhodopsin, and V11 and S10 of arrestin with the final eight residues of rhodopsin (Supplementary Table 5). Together, these cross-linking experiments further validated the interface assembly of the rhodopsin–arrestin complex.

Possible structural mechanisms for biased signalling

The rhodopsin–arrestin complex represents the first crystal structure of a GPCR bound to arrestin and provides an opportunity to examine the mechanism of arrestin-biased signalling. Although a crystal structure of G-protein-bound rhodopsin is not available, several structures of rhodopsin bound to GsCT and analogue peptides have been determined11,22,23,39 and reveal that the arrangement of TM helices in light-activated rhodopsin is similar to that in the G-protein-bound β,γ-AR complex, with the exception of TM6, whose outward movement in β,γ-AR is much more pronounced upon binding to G protein4. The arrestin-bound rhodopsin has its intracellular end of TM6 moved outward by approximately 10Å relative to its inactive structure (Fig. 6a, b and Extended Data Fig. 9). This is in contrast to the 14Å outward movement of TM6 reported in the G-protein-bound β,γ-AR complex5. Compared to the active conformation of rhodopsin bound to GsCT peptides11,22,23,39, arrestin-bound rhodopsin has additional conformational differences in TM1, TM4, TM5, and TM7 (Fig. 6c, d and Extended Data Fig. 9), and these unique structural features may constitute essential elements for arrestin-biased signalling.

The molecular assembly observed in the rhodopsin–arrestin complex also provides a general model for arrestin recruitment by phosphorylated rhodopsin-like class A GPCRs. In the computational model of the full complex, the highly cationic N-terminal domain of arrestin is paired with the C-terminal tail of rhodopsin (Extended Data Fig. 10). Based on the extensive disulfide cross-linking data and computation modelling, phosphorylated S334, S338 and S343 can form tight ionic interactions with three positively charged pockets at the N terminus of arrestin (Extended Data Fig. 11a–d). These results support a model of arrestin activation by phosphorylated rhodopsin through the C-tail exchange mechanism (Fig. 6e)2. The displacement of the arrestin C terminus by the phosphorylated rhodopsin C-tail destabilizes the polar core of arrestin44, thus allowing for a 20° rotation of the N- and C-domains of arrestin that opens a cleft between the middle and C-loops into which the ICL2 helix of rhodopsin can fit. The ionic interaction between rhodopsin and arrestin is consistent with the fact that it is highly salt sensitive in our AlphaScreen assay (Extended Data Fig. 11e), in agreement with the salt-sensitive binding of phosphorylated rhodopsin to arrestin45,48. Importantly, the cytoplasmic face of the rhodopsin TM bundle is highly positively charged, whereas the finger loop (residues 70–78) contains three conserved negatively charged residues (E71, D72, and D74) (Extended Data Fig. 10). Thus, the interaction of arrestin with the rhodopsin TM bundle is mediated not only by shape but also by charge complementarity. Arrestins are highly conserved with only four subtypes in vertebrates. In contrast, there are hundreds of GPCRs, with cytoplasmic interfaces that are mainly non-conserved. However, the positive charge property is a common feature on the cytoplasmic side of a number of GPCR structures (Extended Data Fig. 12). Electrostatic interactions between arrestins and GPCRs...
may represent an adaptive mechanism for arrestins to pair promiscuously with the large number of GPCRs.

The asymmetric orientation of the bound arrestin with regard to the relative positions of its N–C domains in respect to the membrane has important implication in its binding to rhodopsin. In the dark state, the receptor is inactive (R-state) and arrestin is in the closed state (basal state). Receptor activation and phosphorylation (P-R* state) allow the phosphorylated C-terminal tail of rhodopsin to bind to the N-domain of arrestin (pre-activated state), thus displacing the arrestin C-terminal tail. This displacement destabilizes the polar core of arrestin, which allows a 20° rotation between the arrestin N- and C-domains, leading to the opening of the middle loop (ML) and C-loop (CL) to accommodate the ICL2 helix of rhodopsin (fully activated state). The activated receptor also opens the cytoplasmic side of the TM bundle to adopt the finger loop (FL) of arrestin. In this model, the tip of arrestin’s C-domain contacts the membrane (red asterisk).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 December 2013; accepted 12 June 2015.
Published online 22 July 2015.

1. Kim, Y. J. et al. Crystal structure of pre-activated arrestin p44. Nature 497, 142–146 (2013).
2. Shukla, A. K. et al. Structure of active β-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. Nature 497, 137–141 (2013).
3. Pitcher, J. A., Freedman, N. J. & Lefkowitz, R. J. G-protein-coupled receptor kinases. Annu. Rev. Biochem. 67, 653–692 (1998).
4. Wilden, U., Hall, S. W. & Kuhn, H. Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-40a protein of rod outer segments. Proc. Natl Acad. Sci. USA 83, 1174–1178 (1986).
5. Reiter, E., Ahn, S., Shukla, A. K. & Lefkowitz, R. J. Molecular mechanism of beta-arrestin-biased agonism at seven-transmembrane receptors. Annu. Rev. Pharmacol. Toxicol. 52, 179–197 (2012).
6. Kenakin, T. P. Biased signaling and allosteric machines: new vistas and challenges for drug discovery. Br. J. Pharmacol. 165, 1659–1669 (2012).
7. Palczewski, K. et al. Crystal structure of rhodopsin: a G protein-coupled receptor. Science 289, 739–745 (2000).
8. Rasmussen, S. G. et al. Crystal structure of the βδ adrenergic receptor-Gs protein complex. Nature 477, 549–555 (2011).
9. Cherezov, V. et al. High-resolution crystal structure of an engineered human βδ-adrenergic G protein-coupled receptor. Science 318, 1258–1265 (2007).
10. Katritch, V., Cherezov, V. & Stevens, R. C. Structure-function of the G protein-coupled receptor superfamily. Annu. Rev. Pharmacol. Toxicol. 53, 531–556 (2013).
11. Standfuss, J. et al. The structural basis of agonist-induced activation in constitutively active rhodopsin. Nature 471, 656–660 (2011).
12. Xu, F. et al. Structure of an agonist-bound human A2A adenosine receptor. Science 332, 322–327 (2011).
13. Wang, C. et al. Structural basis for molecular recognition at serotonin receptors. Science 340, 610–614 (2013).
14. Wacker, D. et al. Structural features for functional selectivity at serotonin receptors. Science 340, 615–619 (2013).
15. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wuthrich, K. Biased signaling pathways in βα-adrenergic receptor characterized by 19F-NMR. Science 335, 1106–1110 (2012).
16. Zhou, X. E., Melcher, K. & Xu, H. E. Structure and activation of rhodopsin. Acta Pharmacol. Sin. 33, 291–299 (2012).
17. Gurevich, V. V., Hanson, S. M., Song, F. X., Vishnivetskiy, S. A. & Gurevich, E. V. The functional cycle of visual arrestins in photoreceptor cells. Prog. Retin. Eye Res. 30, 405–430 (2011).
18. Smith, S. O. Insights into the activation mechanism of the visual receptor rhodopsin. Biochim. Soc. Trans. 40, 389–393 (2012).
19. Han, M., Smith, S. O. & Sakmar, T. P. Constitutive activation of opsins by mutation of methionine 257 on transmembrane helix 6. Biochemistry 37, 8253–8261 (1998).
20. 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. Methods in Neurosciences 25, 366–428 (1995).
21. 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).
Portions of this research were carried out at the Linac Coherent Light Source (LCLS) and the National Institute of General Medical Sciences of sector 23 (ID-23), funded in part with Federal funds from the National Cancer Institute (NUDT) for computational resources.

**Author Contributions** Y.K. initiated the project, developed the expression and purification methods for rhodopsin–arrestin complex, and bulk-purified expression constructs and proteins used in LCP crystallization for the SFX method; X.E.Z. collected the XFEL data, helped with design of the 9.7 MAG phase diagram and performed computational modelling on-site crystal characterization as well as data analysis; B. Weis for advice on twin refinement and validation of the structure. M.C. generated the Y74C/I140S/C316S stable cell line, characterized and provided the rhodopsin mutant constructs and proteins used in LCP crystallization for the SFX method; X.E.Z. collected the XFEL data and for advising on refinement; B. Weis for advice on twin refinement and structure validation; J.S. provided a set of the 9.7 MAG phase diagram and commented on the paper, A.B., T.A.W., C.G., O.Y., and H.N.C. helped with XFEL data collection and commented on the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper. R.C.S. supervised crystal growth, data collection, and commented on the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper. H.E.X. conceived the project, designed the research, performed synchrotron data collection at APS and LCLS, structure solution and validation, and helped with writing the paper.
METHODS

No statistical methods were used to predetermine sample size.

**Protein preparation.** We used human rhodopsin and mouse visual arrestin-1 in this study. The T4L–rhodopsin–arrestin fusion protein was expressed using a tetracycline vector (Novagen), and the biotin ligase cassette encoding a fusion protein with His$_6$–MBP–MBP followed by a 3C protease cleavage site at the N terminus of the T4L–rhodopsin–arrestin. In this engineered construct, we have fused a cysteine-free T4L (residues 2–161 with C54T and C97A) to the N terminus of a rhodopsin that contains four mutations: N$_{2}$N$_{2}$C$_{6}$M and N$_{28}$B$_{3}$C$_{1}$C to form a disulfide bond, and E113$\gamma$28Q and M257$\gamma$40Y for constitutive receptor activity. The C terminus of rhodopsin was fused to 3A arrestin (L374A, V375A, F376A, and E113$\gamma$3.28Q and M257$\gamma$6.40Y for constitutive receptor activity.

The 3A arrestin was expressed as a His$_6$–SUMO fusion protein from the expression vector pSUMO (LifeSensors). The expression and purification of 3A arrestin were identical as for the rhodopsin–3A arrestin fusion protein used for crystallization, with the difference that the proteins were not eluted, but remained bound to beads.

To generate biotinylated proteins for the AlphaScreen assays, wild-type and 3A arrestin (residues 10–392) open reading frames were cloned with N-terminal avitag–MBP–MBP tag into the first expression cassette of a modified pET-Duet expression vector. The biotin ligase cassette encoding a biotinylated construct was cloned into the second cassette. The 14 amino acid avitag contains a single lysine that is efficiently biotinylated in vivo by the BirA protein$^{35}$. BL21 (DE3) cells transformed with the expression plasmid were grown in LB broth at 16 °C for one hour on a rotating platform to allow for transcription from a T7 promoter. The TNT Quick Coupled Transcription and Translation kit was used according to the manufacturer’s protocol (Promega), to express [35S]methionine-labelled wild-type arrestin and 3A arrestin (L374A, V375A, and F376A). Radiolabelled wild-type arrestin and mutant arrestin proteins were incubated with His$_6$–MBP–rhodopsin fusion protein immobilized to 50 μl of maltose agarose bead suspension. Proteins and beads were incubated at 4 °C for one hour on a rotating platform in binding buffer containing 20 mM Tris–HCl, pH 7.4, 100 mM NaCl and 0.02% DDM/0.004% CHS. The beads were then washed three times with binding buffer and resuspended in 400 μl elution buffer (20 mM Tris–HCl, pH 7.4, 100 mM NaCl, 0.02% DDM/0.004% CHS and 100 mM maltose). The eluates were concentrated and incubated at room temperature for 15 min with 2 × loading dye. Samples were separated on a 12% sodium dodecyl sulfate (SDS)-denaturing polyacrylamide gels. Gels were stained with Coomassie R-250, dried at 70 °C for 90 min, and exposed overnight to a phosphor storage screen. Results were visualized on a PhosphorImager (Fuji).

**Assays for the interactions between rhodopsin and arrestin or G protein.** Interactions between rhodopsin and arrestin were assessed by luminence-based AlphaScreen assay (Perkin Elmer), which our group has used extensively to determine ligand-dependent protein–protein interactions of nucleareceptors. The AlphaScreen principle is illustrated in Extended Data Fig. 1c. Briefly, biotinylated arrestin was bound to streptavidin-coated donor beads and His$_8$-tagged rhodopsin was bound to nickel-chelated acceptor beads. The donor and acceptor beads were brought into close proximity by the interactions between rhodopsin and arrestin, which were measured in the presence or absence of all-trans-retinal (Sigma). When excited by a laser beam of 680 nm, the donor bead emits singlet oxygen that activates thioxene derivatives in the acceptor beads, which releases photons of 520–620 nm as the binding signal. The experiments were conducted with 100 nM of rhodopsin and arrestin proteins in the presence of 5 μg ml$^{-1}$ donor and acceptor beads in a buffer of 50 mM MOPS-Na, pH 7.4, 50 mM NaF, 50 mM CHAPS, and 0.1 mg ml$^{-1}$ bovine serum albumin. The results were based on an average of three experiments with standard errors typically less than 10%.

**GrCT-HA.** GrCT-HA was synthesized by Peptide 2.0. For the competition assay, different amounts of untagged arrestin or GrCT-HA were added to the reaction to compete with tagged arrestin or GrCT-HA for rhodopsin binding.

**Thermal stability assay.** The thermal stability assay was performed with the thiol-specific fluorochrome N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM) as described previously$^{35}$. Briefly, 10 μg of protein was diluted with dilution buffer (20 mM Tris–HCl, pH 7.5, 200 mM NaCl, 0.005% MNG-3/CHS) to 195 μl, and CPM dye stock (4 mg ml$^{-1}$ in DMSO) was freshly diluted to 0.2 mg ml$^{-1}$ in the dark. After 5 min of incubation at room temperature for both protein and CPM dye separately, 5 μl of diluted CPM dye was added to the protein sample and the protein/dye mix transferred into a sub-micro quartz fluorometer cuvette (Starna Cells) and measured in a Cary Eclipse spectrofluorometer (Agilent). Assays were performed from 20 °C to 80 °C with a slope of 2 °C increase per minute at an excitation wavelength of 387 nm and an emission wavelength of 463 nm. All data were processed using GraphPad Prism and fit using the Boltzmann sigmoidal equation to determine the melting temperature (T_m) as inflection point of the melting curves.

**Cryocrystallization.** T4L–rhodopsin–arrestin crystals were grown in lipid cubic phase (LCP)$^{56}$. Protein solution (~30 mg ml$^{-1}$) was mixed with monopalmityl-lecithin (MP) from Nu-Cholesterol (MAG) containing 1:1 ratio (w/w) using a coupled syringe mixer$^{51}$ and 50 nl boluses of protein-laden LCP were dispensed on 96-well glass sandwich plates (Molecular Dimensions or Marienfeld-Suprior) and overlaid with 0.8 μl precipitant solutions using a Gryphon LCP robot (Art Robbins Instruments) or an NTB-LCP robot (Formulatrix). Multiple initial hits were identified by using screens of 30% PEG 400 in combination of 100 mM or 400 mM salts from the StockOptions Salt kit (Hampton Research)$^{58}$. Crystals that reached full size (around 10–20 μm) within 40 days at 20 °C were harvested from the mesophase and were flash frozen in liquid nitrogen without additional cryoprotectant. Crystals used for synchrotron and XFEL diffraction data collection were grown in 0.05 M magnesium acetate, 0.05 M sodium acetate, pH 5.0 and 28% PEG 400.

Crystals for LCP-SFX were prepared in 100 μl gas-tight Hamilton syringes as described$^{65,67}$. About 5 μl of protein-laden LCP in the presence of fivefold molar excess of all-trans-retinal was slowly injected into 60 μl mother liquid buffer (0.15 M ammonium phosphate, pH 6.4 and 32% PEG 400) using a coupled syringe mixing device$^{66}$. Crystals were grown in several syringes at 20 °C, consoliated and transferred into the LCP injector$^{50}$ for XPEF diffraction data collection. The average crystal size was 5–10 μm as determined under a polarized light microscope. The phased diagram for 9.7 °C MAG suggests that this MAG is a suitable host lipid for extruding LCP in vacuum at 20 °C, where evaporative cooling created problems when 9.9 MAG was used as a host LCP lipid$^{35}$.
Argonne National Laboratory. A full 7.7 Å data set was collected from a single crystal (~20 µm in size) using 10 µm beam size and 0.1 s exposures per 0.1° oscillation with a Pilatus 6M pixel detector at the X10SA beam line at the Swiss Light Source. The observed reflections were reduced, merged, and scaled with XDS* with statistics shown in Supplementary Table 1a. The L-test plot of the 7.7 Å data set is consistent with a perfectly twinned crystal.

Data collection (XFEL). LCP-SFX experiments were carried out at the Coherent X-ray Imaging (CXI) instrument* at the Linac Coherent Light Source (LCLS) in the SLAC National Accelerator Laboratory (Menlo Park, California, USA). X-ray pulses of 20 fs duration at a wavelength of 1.3 Å (9.5 keV) were attenuated to ~3% (3 x 10^16 photons per pulse) and focused to ~1.5 µm diameter at the interaction point using Kirkpatrick–Baez mirrors. Rhodopsin–arrestin complex crystals in LCP were injected across the XFEL beam using an LCP injector with a 50 µm diameter nozzle at a flow rate of ~0.2 µl min^-1. Diffraction patterns were collected at 120 Hz using the Cornell-SLAC Pixel Array Detector (CSPAD). Over 5 million data frames were collected corresponding to ~12 h of data acquisition time. Of these frames, ~0.45% images contained potential crystal hits as identified using Cheetah* (more than 40 Bragg peaks of 1–20 pixels in size and a signal to noise ratio better than 6 after local background subtraction). Of the potential crystal hits, 18,874 diffraction patterns could be auto-indexed by CrystFEL* using a combination of MOSFLM 62, XDS 58 and DirAx* 63. An integration radius of only two pixels was used to avoid overlapping with neighbouring peaks due to the high spot density resulting from the large unit cell dimensions.

Partial reflections from different crystals in random orientations were gathered using a Monte Carlo integration across the crystal rocking curve of each reflection. The resolution was anisotropic with ~3.3 Å resolution along the c*-axis and ~3.8 Å resolution along the a*b* axes. The data used for the structure refinement were trimmed at 3.8 Å/3.8 Å/3.8 Å using the get.hkl program of CrystFEL* based on the criteria of data correlation coefficient (CC*), which is 0.87 at the highest resolution shell (Supplementary Table 1a). The use of CC* of 0.5 as resolution cutoff has been recently recommended for X-ray data correlation coefficients (CC*).

Structure determination. The Xfel data were initially merged according to the apparent Laue group of 4/mmm, and molecular replacement searches were performed in all possible space groups of 4/m and 4/mmm. The best structure solution was found in P21. Based on analysis of the Zanuda program* we determined that the most likely space group was P212121, and that the crystals were physically twinned. The data were reprocessed with the Laue group of mm2 and molecular replacement searches were performed in P212121. This space group assignment resulted in the best statistics and map quality out of several possible space groups (Supplementary Table 1b).

The crystals appeared to be pseudo-merohedrally twinned based on L-test analysis*. Despite the challenge of twinned data, the rhodopsin–arrestin complex structure was solved by the molecular replacement method implemented in Phaser* using the models of constitutively active rhodopsin, pre-activated arrestin, and the complex of rhodopsin with arrestin respectively. Four molecules of rhodopsin and four molecules of arrestin were found sequentially by molecular replacement search, resulting in four very similar rhodopsin–arrestin assemblies. Four T4Ls were also found in the aqueous layer with its C-terminal Flag tag, and full-length rhodopsin with C-terminal haemagglutinin (HA) tag were cloned into pcDNA6. Cysteine mutations (41 for arrestin and 51 for rhodopsin) were systematically introduced into arrestin and rhodopsin in these two DNA vectors. AD293 cells were split one day before transfection at 50,000 cells per well in a 24-well plate. Cells were grown for one day, then transfected with 100 ng rhodopsin constructs (pcDNA6-rho-3HA) plus 100 ng arrestin plasmid (pcDNA6-Arr-3Flag) by Lipofectamine 2000 (DNA/Lipofectamine 2000 ratio of 1:2) in each well. Cells were grown for 2 days after transfection, and were then treated at room temperature with H2O2, which was freshly diluted in the cell culture medium to a final concentration of 1 mM. After 5 min treatment with H2O2, the medium was aspirated and 100 µl of Cellytic M (Sigma C2978) was added to each well and the plate was shaken for 10 min at room temperature. Cell lysates were transferred to 1.5 ml tubes and mixed with an equal volume of 2 × SDS loading buffer (without reducing agents) for 5 min at room temperature, and loaded onto a protein gel for western blot analysis. Horseradish peroxidase-conjugated anti-Flag (Sigma M2) and anti-HA (Sigma) antibodies were used to probe for free and cross-linked arrestin and rhodopsin proteins.

Hydrogen-deuterium exchange mass spectrometry (HDX). HDX was carried out as described previously*, with the following modifications: (1) the solution handling and mixing was performed with a LEAP Technologies Twin HTS PAL liquid handling robot housed inside a temperature-controlled cabinet held at 4 °C (and (2) deuterium oxide (DDM) was used in place of DDD in the exchange buffer. Briefly, all stock solutions and dilutions were made using the 7T HDX buffer (50 mM HEPES (pH 7.5), 150 mM NaCl, 2% (v/v) glycerol, 0.01% (m/v) CHS and 0.05% (m/v) DMNG in either H2O or in D2O for on-exchange). All HDX protein stock solutions were prepared at 15 µM in the 7T
HDX H2O buffer. On-exchange was carried out in triplicate for predetermined times (10, 30, 60, 900, and 3,600 s) at 4 °C by mixing 5 μl of stock protein solution with 20 μl of D2O on-exchange buffer. Exchange was quenched by adding 25 μl of quench solution (100 mM Na2HPO4, 0.02% DMNG, and 15 mM TCEP at pH 2.4) to the reaction. Digestion was performed in line with chromatography using an in-house packed peptides column. Peptides were captured and desalted on a C8 trap. Peptides were then separated across a 5 μl 10 × 1 mm Betasil C8 column (Thermo Fisher Scientific) with a linear gradient of 12–40% acetonitrile in 0.3% formic acid over a short 5 min gradient to limit back exchange with the solvent.

Mass spectra were acquired in the range of m/z 300–2,000 at a resolution of 60,000 for 8 min in positive ion mode on a Q Exactive mass spectrometer (Thermo Fisher Scientific) equipped with an ESI source operated at a capillary temperature of 225 °C and spray voltage of 3.5 kV. The intensity weighted average m/z value (centroid) of each peptide's isotopic envelope was calculated with mascot search engine as described previously45. For DEER measurements, the spin-labelled proteins were mixed in 1:1 ratio in the dark and loaded into quartz capillaries (1.5 mm internal diameter and 1.8 mm outer diameter). The samples were irradiated for 30 s within the capillaries using a tungsten light source with a 500 nm cut-off filter. Immediately after irradiation, the samples were flash frozen in liquid nitrogen, and loaded into an EN 510/D2 resonator for Q band DEER measurements. Measurements were performed at 80 K on a Bruker Elexys 580 spectrometer with a Super Q-FTu Bridge. A 36-ns π-pump pulse was applied to the low field peak of the nitroxide field sweep spectrum, and the observer π/2 (16 ns) and π (32 ns) pulses were positioned 50 MHz (17.6 G) upfield, which corresponds to the nitroxide centre line. Model-free distance distributions were obtained from the raw dipolar evolution data using the LabVIEW (National Instruments) program “LongDistances” that can be downloaded from http://www.biochemistry.ucla.edu/biochem/Faculty/Hubbell/.

To estimate the median distances, the distance distributions were integrated and normalized to the maximum amplitude. The median distance was estimated as that corresponding to 0.5 of the integrated intensity. The modelled distances between nitroxide spin labels are based on the crystal structure of the rhodopsin–arrestin complex. R1 nitroxide side chains were modelled into the structure using common R1 rotamers56. 9.7 MAG/water temperature-composition phase diagram. The phase diagram was constructed based on small- and wide-angle X-ray scattering measurements made in the heating direction. Sample preparation and X-ray scattering measurements and analysis were as previously described57. The phases identified include the lamellar crystalline (Lc) or solid phase, the fluid isotropic (FI) or liquid phase, and the following liquid crystalline phases: lamellar liquid crystal (LcL), cubic-Ia3d and cubic-Pn3m. A separate aqueous phase observed in equilibrium with the solid or liquid crystalline phases is indicated by Aq. The phase diagram shows that the solid Lc phase stabilizes under equilibrium conditions below ~8 °C. The latter is some 10 °C below that observed with 9.9 MAG (monolein)58 and is similar to what was found with 7.9 MAG59. This low solidification temperature enabled use in the current project of 9.7 MAG as the host lipid for LCP-SFX data collection in an evacuated sample chamber at 20 °C, where evaporative cooling created problems for measurements with 9.9 MAG but not with 7.9 MAG59. The maximum water carrying capacity of 9.7 MAG resides at ~50% (w/w) water, which is considerably greater and smaller than that of 9.9 MAG56 and 7.7 MAG59, respectively. These observations indicate that the cubic mesophase of 9.7 MAG has larger aqueous channels compared to 9.9 MAG that are more like those of 7.7 MAG. This is consistent with 9.7 MAG supporting the growth of rhodopsin–arrestin–T4L crystals where the complex has sizable extra-membrane bulk best accommodated in a large aqueous channel. This parallels the rational use of 7.7 MAG as a host lipid for the β2AR–Gs complex crystallization and structure determination5,6.

Molecular modelling of the full-length rhodopsin–arrestin complex. Energy-based conformational modelling of the rhodopsin–arrestin complex was performed with the ICM-Pro molecular suite60, using a global energy optimization procedure similar to the one described recently for modelling of the full-length complex of CRFR161. Protein sequences of human rhodopsin and mouse arrestin were obtained from the Uniprot database (http://www.uniprot.org/). Starting from the crystallographically determined structure of the complex, the modelling procedure was used to add unresolved residues of the C terminus (residues 327–342 of mouse arrestin) as well as missing residues of the arrestin C-terminal loop (residues 340–342) of the mouse arrestin structure. The final model did not include the last 3 residues of rhodopsin (346–348), which lack well-defined cross-linking contacts and appear flexible. Initial conformations of the short loop in arrestin were predicted with the fast “build model” ICM algorithm, followed by extensive energy optimization in internal coordinates. Conformational optimization of the rhodopsin C-terminal peptide was guided by soft pairwise harmonic distance restraints derived from disulfide crosslinking data. The restraints introduced between Cβ atoms of the crosslinked residues were graded according to the crosslinking strength listed in Supplementary Table 5, from very strong, with the penalty function starting at 2 A˚ distance, to medium at 7 A˚ distance and very weak at 12 A˚. The C-terminal peptide conformation and confirmation of the contact side chains of the arrestin were optimized to convergence (3 independent simulations of 106 steps) using global optimization procedure in internal coordinates with improved conformational energy terms for protein and peptides57. A special
backbone closure sampling procedure was applied to the loop regions to allow efficient optimization. The global optimization runs were executed in parallel on a Linux multicore server resulting in similar best energy conformations (<3 Å r.m.s.d.; root mean squared deviation) for the C-terminal peptide residues.

The best energy-optimized conformation of this region suggests that the extended C terminus peptide runs antiparallel along the N-terminal β-strand of arrestin. This conformation of the C terminus satisfied all 17 medium to strong disulfide crosslinking restraints for this region (Supplementary Table 5), while making a number of specific polar interactions and salt bridges of D131, E332 and E341 side chains with arrestin basic residues (Extended Data Fig. 5). These modelling results suggest that although the rhodopsin C terminus is rather flexible, some low energy conformations may be preferable even in non-phosphorylated rhodopsin. Moreover, independent modelling of the complex with phosphorylated serine residues Ser334, Ser338 and Ser343 in the C termus of rhodopsin resulted in a similar conformation of this domain. The interactions within phosphorylated complex, however, are greatly enhanced by as many as seven additional salt bridges between negatively charged phosphates and the positively charged lysine and arginine residues within the N-terminal domain of arrestin (Extended Data Fig. 11).

51. Beckert, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in boiot holoenzyme synthetase-catalyzed biotinylation. Protein Sci. 8, 921–929 (1999).
52. Mossewsova, E. & Lima, C. D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Mol. Cell 5, 865–876 (2000).
53. Alexandrov, A. I., Mileni, M., Chien, E. Y. T., Hanson, M. A. & Stevens, R. C. Microscale fluorescent thermal stability assay for membrane proteins. Structure 16, 351–359 (2008).
54. Caffrey, M. & Cherezov, V. Crystalizing membrane proteins using lipidic mesophases. Nature Protocols 4, 706–710 (2009).
55. Chen, A. H., Hummel, B., Qiu, H. & Caffrey, M. A simple mechanical mixer for small viscous lipid-containing samples. Chem. Phys. Lipids 95, 11–21 (1998).
56. Xu, F., Liu, W., Hanson, M. A., Stevens, R. C. & Cherezov, V. Development of an automated high throughput LCP-FRAP assay to guide membrane protein crystallization in lipid mesophases. Cryst. Growth Des. 11, 1193–1201 (2011).
57. Liu, W., Ishchenko, A. & Cherezov, V. Preparation of microcysts in lipidic cubic phase for serial femtosecond crystallography. Nature Protocols 9, 2123–2134 (2014).
58. Kabsch, W. Integration, scaling, space-group assignment and post-refinement, Acta Crystallogr. D 66, 133–144 (2010).
59. Boutet, S. & Williams, G. J. The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). New J. Phys. 12, 035024 (2010).
60. Siewert, F. et al. Ultra-precise characterization of LCLS hard X-ray focusing mirrors by high resolution slope measuring deflectometry. Opt. Express 20, 4525–4536 (2012).
61. White, T. A. et al. Crystallographic data processing for free-electron laser sources. Acta Crystallogr. D 69, 1221–1250 (2013).
62. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. IMOSFIL: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D 67, 271–281 (2011).
63. Duisenberg, A. J. M. Indexing in single-crystal diffractometry with an octahedral list of reflections. J. Appl. Crystallogr. 25, 92–96 (1992).
64. Kirian, R. A. et al. Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals. Acta Crystallogr. A 67, 131–140 (2011).
65. Karpowicz, P. A. & Diederichs, K. Linking crystallographic model and data quality. Science 336, 1030–1033 (2012).
66. Lebedev, A. A. & Isupov, M. N. Space-group and orientation ambiguity in macromolecular structures with pseudo-symmetry and its treatment with the program Zanuda. Acta Crystallogr. D 70, 2430–2443 (2014).
67. Padilla, J. E. & Yeates, T. O. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. Acta Crystallogr. D 59, 1124–1130 (2003).
68. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
69. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D 66, 486–505 (2010).
Extended Data Figure 1 | Constitutively active rhodopsin interacts with arrestin and GαCT-HA. **a**, SDS–PAGE of N-terminally MBP-tagged wild-type and mutant rhodopsin. **b**, Non-cropped versions of the pull-down assay gels shown in Fig. 1b. The interactions between mouse wild-type arrestin and human wild-type or E1133.28Q rhodopsin are very weak. In contrast, the interaction between constitutively active rhodopsin (E1133.28Q/M2576.40Y) and pre-activated L374A/V375A/F376A arrestin (3A arrestin) is strong and is further increased in the presence of 10 μM all-trans-retinal. Input: 5% of the binding reaction. Bottom panels show the rhodopsin loading controls. **c**, Schematic representation of the AlphaScreen assay. **d**, AlphaScreen binding assay between E1133.28Q/M2576.40Y rhodopsin and GαCT-HA (TGGRVLEDLKCGLF) in the presence and absence of 5 μM all-trans-retinal. The two left columns show the controls with ‘peptide only’ and ‘rhodopsin only’. (n = 3, error bars, s.d.). **e**, Determination of the affinity of the interaction between rhodopsin E1133.28Q/M2576.40Y and GαCT-HA by homologous competition. His6–MBP–rhodopsin mutant protein was immobilized on Ni-acceptor beads and biotinylated GαCT-HA on streptavidin donor beads. Binding between rhodopsin and arrestin brings donor and acceptor beads into close proximity, resulting in the indicated binding signal. Non-biotinylated GαCT-HA competed for the interaction with an IC50 of ~700 nM (n = 3, error bars, s.d.).
Extended Data Figure 2 | Purification and crystallization of T4L–rhodopsin–arrestin. a, Purification of the T4L–rhodopsin–arrestin (T4L–Rho–Arr) complex. His₈–MBP–MBP–T4L–Rho–Arr complex was first purified by amylose column chromatography (lane 1). The His₈–MBP–MBP tandem tag was then released by cleavage with 3C protease (lane 2) and removed by binding to Ni-NTA beads to recover pure T4L–rhodopsin–arrestin (T4L–Rho–Arr) protein (lane 3).
b, Analytical gel filtration profile of the T4L–rhodopsin–arrestin complex. T4L–rhodopsin–arrestin eluted mostly as monomers with a small proportion of oligomers. The molecular weights of protein standards are indicated at the top. c, Thermal stability shift analysis of T4L–rhodopsin–arrestin. T4L–rhodopsin–arrestin is relatively stable with a T_m of 59 °C. d, e, Crystals of T4L–rhodopsin–arrestin in lipid cubic phase under bright-field illumination (d) and polarized light (e). f, X-ray diffraction pattern of a T4L–rhodopsin–arrestin crystal recorded at LS-CAT of APS. The green ring indicates the position of reflections at 8.0 Å resolution.
Extended Data Figure 3 | Electron density map for the overall complex and the key interfaces based on the XFEL data. a, A $2F_o - F_c$ electron density map contoured at 1$\sigma$ of the arrestin finger loop, which forms the key interface with TM7 and helix 8. b, A $2F_o - F_c$ electron density map contoured at 1$\sigma$ of the loop between TM5 and TM6, which forms the key interface with the $\beta$-strand following the finger loop. c, A 3,000 K simulated annealing omit map ($2F_o - F_c$ electron density map contoured at 1$\sigma$) calculated from the 3.8 Å/3.8 Å/3.3 Å XFEL data supports the overall arrangement of the rhodopsin–arrestin complex. In all panels, the complex structure is shown with rhodopsin coloured in green and arrestin in brown. d, The C-loop with a $2F_o - F_c$ composite omit map at 1$\sigma$ calculated from the 3.8 Å/3.8 Å/3.3 Å truncated XFEL data. Key residues are labelled.
Extended Data Figure 4 | Structure similarity of the four rhodopsin–arrestin complexes in the asymmetric units and the interface between rhodopsin and arrestin. a, Two 90° views of the superposition of the four rhodopsin–arrestin complexes are shown as cartoon representation. The four complexes have an r.m.s.d. of less than 0.5 Å in the Cα atoms of rhodopsin and arrestin. b, Close-up view of arrestin-binding sites in rhodopsin. The four arrestin-binding sites (P1–P4) are highlighted in brown on the rhodopsin surface. The rhodopsin C-terminal tail/arrestin interface (P4) is based on computational modelling and disulfide cross-linking data. c, Rhodopsin-binding sites in arrestin. The four rhodopsin-binding sites (P1–P4) are highlighted in green on the arrestin surface.
Extended Data Figure 5 | Conformational modelling of the rhodopsin–arrestin full length complex. 

a, An overview of the computational model.  
b, Predicted interactions of the rhodopsin C terminus with arrestin, showing strong to medium pairwise restraints between Cβ atoms of rhodopsin and arrestin residues identified by disulfide crosslinking.  
c, Same as in b, but showing predicted hydrogen bonding and ionic interactions for the C-terminal residues of rhodopsin.
a

Average % Deuterium Exchange

b

Arrestin, $T_m=53 \, ^\circ C$
T4L-Rho-Arr, $T_m=59 \, ^\circ C$
Extended Data Figure 6 | Dynamics of free 3A arrestin and rhodopsin-bound arrestin determined by HDX. a, HDX perturbation map between rhodopsin-bound arrestin and free arrestin, which is derived from the difference in the HDX rate between rhodopsin-bound arrestin and free arrestin. The bars below the arrestin sequence represent the peptide fragments resolved by mass spectrometry and the colours of the bars indicate the relative decrease in deuterium exchange (colour code at bottom). b, The thermal stability of free 3A arrestin and the rhodopsin–arrestin complex shows that the rhodopsin–arrestin complex is more stable than free 3A arrestin.
Extended Data Figure 7 | Cell-based Tango assays to validate the rhodopsin–arrestin interface. a, Cartoon illustration of the Tango assay for rhodopsin–arrestin interactions in cells. b, c, Mutations of key arrestin (b) and rhodopsin (c) residues that mediate the rhodopsin–arrestin interactions. Tango assay were performed in the absence or presence of 10 μM all-trans-retinal (ATR). (n = 3, error bars, s.d.).
**Extended Data Figure 8** | Control experiments for disulfide bond cross-linking specificity.  
**a**, The product of the cross-linking reaction of finger loop residue G77C with N310^7.57C of TM7 was confirmed by western blots using anti-Flag antibody (which detects arrestin–Flag fusion) and anti-HA antibody (which detects rhodopsin–HA fusion). The cross-linked products are marked with arrow heads, and free-arrestin and free-rhodopsin are indicated by asterisks. Arrestin (3A) and rhodopsin (4M) without cysteine mutations do not form cross-linked products. 
**b**, The cross-linked product of finger loop residue G77C with N310^7.57C of TM7 was sensitive to treatment with reducing agents, indicating the cross-linking is mediated through disulfide bond formation. 
**c**, A close-up view of arrestin finger loop residues M76C and G77C and their cross-linking with rhodopsin, which shows that G77C was specifically cross-linked to N310^7.57C of TM7 and Q312^8.49C of helix 8, and M76C was cross-linked to N310^7.57C of TM7 and Q312^8.49C of helix 8, but not to other residues. 
**d**, Structure and cross-linking of finger loop N-terminal residues Q70C, E71C, and D72C of arrestin to T70C and K67C from ICL1 of rhodopsin. 
**e**, Structure and cross-linking of arrestin back loop residues R319C and T320C to Q237^ICL3C from TM5 of rhodopsin.
Extended Data Figure 9 | Structure comparison of the arrestin-bound rhodopsin with the $\beta_2$-adrenergic receptor in complex with Gs protein (PDB code 3SN6) and the inactive rhodopsin (PDB code 1F88). a, Superposition of arrestin-bound rhodopsin (green) with Gs protein-bound $\beta_2$-adrenergic receptor (light yellow). The major conformational changes are indicated by arrows. b, An intracellular view of a superposition of arrestin-bound rhodopsin (green) and Gs protein-bound $\beta_2$-adrenergic receptor (light yellow). c, Overlays of arrestin-bound rhodopsin (green) with inactive rhodopsin (pink) reveals specific conformational changes in each TM helix. The arrows indicate outward movements of TM helices. d, r.m.s.d. of Cα atom differences between arrestin-bound rhodopsin and inactive rhodopsin shows the large conformational changes in TM5 and TM6.
Extended Data Figure 10 | Structure of rhodopsin-bound arrestin and its comparison with inactive and ‘pre-activated’ arrestin.  

**a, b.** The charge potential surface map of rhodopsin from the rhodopsin–arrestin bound complex shows that the cytoplasmic rhodopsin TM bundle surface is positively charged (blue) whereas its C-terminal tail is negatively charged (red).

**c, d.** Charged surface of arrestin from the rhodopsin–arrestin bound complex shows that the arrestin finger loop is negatively charged (red) and its N-terminal β-strand interface is positively charged (blue). The charge distribution in rhodopsin and arrestin is complementary to each other for their interactions.  

**e.** Comparison of rhodopsin-bound arrestin (light blue) with inactive arrestin (brown, PDB code: 1CF1), showing an ~20° rotation between the N- and C- domains of arrestin.  

**f.** Comparison of rhodopsin-bound arrestin (dark brown) with pre-activated arrestin (light brown, PDB code 4J2Q), showing conformational changes in the finger loop, which adopts an α-helical conformation (cyan) in the complex. The extended finger loop conformation would protrude into the rhodopsin TM bundle and is not compatible with receptor binding. Computational model for the full rhodopsin–arrestin complex is shown in panels **b** and **d**.
Extended Data Figure 11 | A computational model of phosphorylated rhodopsin in complex with arrestin and salt sensitivity of the rhodopsin–arrestin interaction. a–d. An overall view (a) and close-up views (b–d) of the computational model of the rhodopsin C-tail with phospho-serine at positions 334, 338 and 343 in complex with arrestin. e. The AlphaScreen control (biotin–His₆) shows much less salt sensitivity than the interaction between His-tag–rhodopsin and biotin arrestin, which is very sensitive to salt, with an IC₅₀ of around 200 mM NaCl (100 mM NaCl added to 100 mM salt of the original assay buffer) (n = 3, error bars, s.d.).
Extended Data Figure 12 | A positive charge property is commonly found at the cytoplasmic side of GPCRs. a–e, Surface charge potential of the cytoplasmic side of selected agonist bound GPCR structures: β2AR, PDB code 2Y02 (a); β2AR, PDB code 3PDS (b); A2A adenosine receptor, PDB code 3QAK (c); serotonin receptor 5HT1B, PDB code 4IAR (d); serotonin receptor 5HT2B, PDB code 4IB4 (e). Positive and negative charge potentials are shown in blue and red, respectively. f, Sequence alignment of the finger loop region highlighting negatively charged residues (shown in red), which are conserved in all subtypes of arrestins.
A possible role of the arrestin C-edge in lipid binding. 

**a, b.** The asymmetric assembly of the rhodopsin–arrestin complex in the presence of a lipid membrane bilayer, showing the C-edge of arrestin dipping into the lipid layer.

**c, d.** A close-up view of the C-edge of arrestin in the membrane layer, where the conserved hydrophobic side chains are shown. The figure was made using the computational model for the full rhodopsin–arrestin complex.