ARTICLE

doi:10.1038/nature14656

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

Yanyong Kang[1]*, X. Edward Zhou[1]*, Xiang Gao[1]*, Yuanzheng He[1]*, Wei Liu[2], Andrii Ishchenko[3], Anton Barty[4], Thomas A. White[4], Oleksandr Yefanov[4], Gye Won Han[3], Qingping Xu[5], Parker W. de Waal[1], Jiyuan Ke[1], M. H. Eileen Tan[1,6], Chenghai Zhang[1], Arne Moeller[7], Graham M. West[8], Bruce D. Pascal[8], Ned Van Eps[9]†, Lydia N. Caro[10], Sergey A. Vishnivetskiy[11], Regina J. Lee[11], Kelly M. Suino-Powell[1], Xin Gu[1], Kuntal Pal[1], Jinming Ma[1], Xiaoyong Zhi[1], Sébastien Boutet[12], Garth J. Williams[12], Marc Messerschmidt[12,13], Cornelius Gati[4], Nadia A. Zatsepin[2,14], Dingjie Wang[2,14], Daniel James[2,14], Shibom Basu[2,14], Shatabdi Roy-Chowdhury[2,14], Chelsie E. Conrad[2], Jesse Coe[2], Haiguang Liu[2,15], Stella Lisova[2], Christopher Kupitz[2,16], Ingo Grotjohann[2], Raimund Fromme[2], Yi Jiang[17], Minjia Tan[17], Huaiyu Yang[17], Jun Li[6], Meitian Wang[18], Zhong Zheng[19], Dianfan Li[20], Nicole Howe[20], Yingming Zhao[13,21], Jörg Standfuss[22], Kay Diederichs[23], Yuhui Dong[24], Clinton S. Potter[7], Bridget Carragher[7], Martin Caffrey[20], Hualiang Jiang[17], Henry N. Chapman[4,25], John C. H. Spence[2,14], Petra Fromme[2], Uwe Weierstall[2,14], Oliver P. Ernst[10,26], Vsevolod Katritch[19], Vsevolod V. Gurevich[11], Patrick R. Griffin[8], Wayne L. Hubbell[9], Raymond C. Stevens[3,19,27], Vadim Cherezov[3], Karsten Melcher[1] & H. Eric Xu[1,28]

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 carboxy 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[1,2]. 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[3]. Arrestin binding to the receptors blocks their interactions with G proteins and leads to the receptor's desensitization[4]. 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 current clinically used drugs. Recent studies have demonstrated that G-protein and arrestin pathways are distinct and can be pharmacologically modulated independently using biased GPCR ligands[5]. 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[6].

The molecular mechanisms of GPCR signalling have been unravelled by recent breakthroughs in GPCR structural biology[7–10]. In the antagonist-bound state, GPCRs assume a closed conformation with the cytoplasmic ends of the transmembrane (TM) helices packed closely with each other[7,9], 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

---

[1] Laboratory of Structural Sciences, Center for Structural Biology and Drug Discovery, Van Andel Research Institute, Grand Rapids, Michigan 49503, USA. [2] Department of Chemistry and Biochemistry, and Center for Applied Structural Discovery, Biodesign Institute, Arizona State University, Tempe, Arizona 85287-1604, USA. [3] Department of Chemistry, Bridge Institute, University of Southern California, Los Angeles, California 90089, USA. [4] Center for Free Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany. [5] Joint Center for Structural Genomics, Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [6] Department of Obstetrics & Gynecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [7] The National Resource for Automated Molecular Microscopy, New York Structural Biology Center, New York, New York 10027, USA. [8] Department of Molecular Therapeutics, The Scripps Research Institute, Scripps Florida, Jupiter, Florida 33458, USA. [9] Jules Stein Eye Institute and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA. [10] Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [11] Department of Pharmacology, Vanderbilt University, Nashville, Tennessee 37232, USA. [12] Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [13] BioXFEL, NSF Science and Technology Center, 700 Ellicott Street, Buffalo, New York 14203, USA. [14] Department of Physics, Arizona State University, Tempe, Arizona 85287, USA. [15] Beijing Computational Science Research Center, Haidian District, Beijing 10084, China. [16] Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53211, USA. [17] State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. [18] Swiss Light Source at Paul Scherrer Institute, CH-5232 Villigen, Switzerland. [19] Department of Biological Sciences, Bridge Institute, University of Southern California, Los Angeles, California 90089, USA. [20] School of Medicine and School of Biochemistry and Immunology, Trinity College, Dublin 2, Ireland. [21] Ben May Department for Cancer Research, University of Chicago, Chicago, Illinois 60637, USA. [22] Laboratory of Biomolecular Research at Paul Scherrer Institute, CH-5232 Villigen, Switzerland. [23] Department of Biology, Universität Konstanz, 78457 Konstanz, Germany. [24] Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China. [25] Centre for Ultrafast Imaging, 22761 Hamburg, Germany. [26] Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [27] iHuman Institute, ShanghaiTech University, 2F Building 6, 99 Haike Road, Pudong New District, Shanghai 201210, China. [28] VARI-SIMM Center, Center for Structure and Function of Drug Targets, CAS-Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. †Present address: Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

*These authors contributed equally to this work.

of the TM domain[8,11–14], thus 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[14,15].

Rhodopsin is a prototypical GPCR responsible for light perception[7]. Along with the β_2-adrenergic receptor (β_2AR), rhodopsin has served as a model system for studying GPCR signalling[16]. 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[17,18]. Light-activated rhodopsin is then phosphorylated by rhodopsin kinase (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 E113$^{3.28}$Q/M257$^{6.40}$Y mutation, which yields a constitutively active rhodopsin[19] (superscripts in residues refer to the Ballesteros–Weinstein numbering[20]). The crystal structure of bovine rhodopsin has been determined in the inactive, resting state[7], the ligand-free state[21,22], and the ligand-activated state in complex with a G-protein peptide[23]. Arrestin structures have been determined in the inactive[24,25] and pre-activated form[1,2]. Recent electron microscopy analysis has revealed the assembly and conformational dynamics of the β_2AR–β-arrestin complex[26]. Here we report the crystal structure of an active form of human 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.

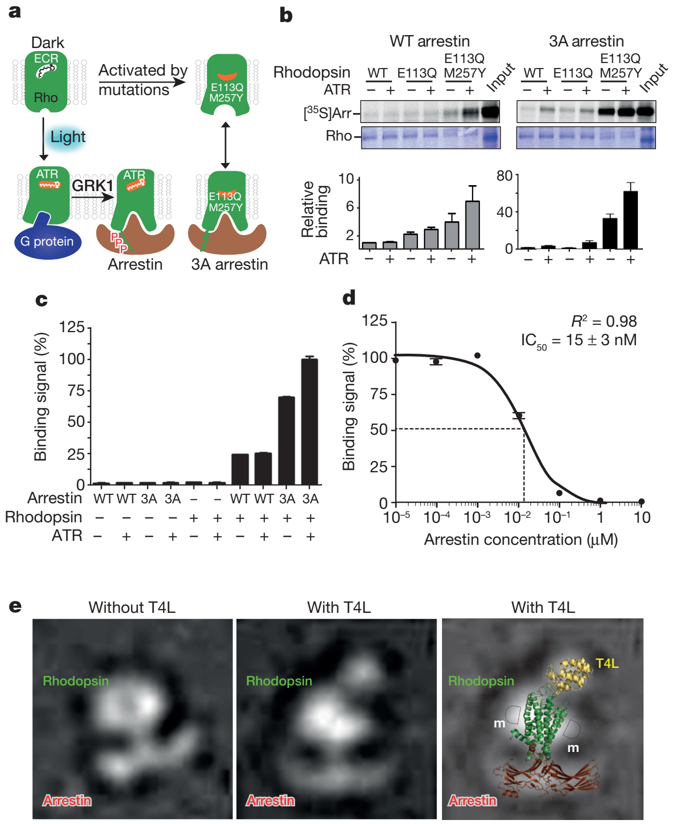


Figure 1 | Rhodopsin–arrestin interactions and complex assembly. **a**, Diagram of the binding of rhodopsin (Rho) with G protein and arrestin as described in the main text. Labels are 11-*cis* retinal (ECR), all-*trans*-retinal (ATR), and rhodopsin kinase (GRK1). **b**, Rhodopsin (Rho) and arrestin (Arr) interaction determined by pull-down assay in the absence and presence of ATR (top panel). Middle panel, rhodopsin loading controls. Bottom panel, relative binding of ^{35}S-labelled arrestin was determined by densitometry ($n = 3$, error bars, s.d.). **c**, Binding of His$_8$–MBP–rhodopsin (E113$^{3.28}$Q/M257$^{6.40}$Y) protein to biotin–MBP–arrestin (wild type (WT) and 3A) measured by AlphaScreen in the absence or presence of 5 μM ATR. The first six columns are controls (luminescence signals in the presence of only one of the binding partners; $n = 3$, error bars, s.d.). **d**, Competition of arrestin binding to rhodopsin was determined by a homologous AlphaScreen assay and the IC$_{50}$ value was derived from repeat experiments ($n = 3$, error bars, s.d.). **e**, Negative stain electron microscopy images of rhodopsin–arrestin complexes without or with T4L at the N terminus; right panel, overlay of the electron microscopy image with the structures of T4L, rhodopsin and arrestin. m, detergent micelle.

Characterization and crystallization

To characterize the rhodopsin–arrestin interaction, we expressed and purified E113$^{3.28}$Q and E113$^{3.28}$Q/ M257$^{6.40}$Y mutant receptors (Extended Data Fig. 1a). These mutations were introduced in the context of the N2$^{\mathrm{Nterm}}$C/N282$^{\mathrm{ECL3}}$C mutant that is known to create a disulfide bond that increases rhodopsin stability without affecting its activity[27–29]. The N2$^{\mathrm{Nterm}}$C/N282$^{\mathrm{ECL3}}$C mutant is referred to as our wild-type control. To determine the interaction between rhodopsin and arrestin, we developed a bead binding pull-down assay. In this assay, rhodopsin expressed as a fusion with a maltose-binding protein (MBP) at its N terminus was bound to amylose beads, which were then used to pull down *in vitro* translated arrestin labelled with ^{35}S. Wild-type arrestin has weak background binding to wild-type rhodopsin (Fig. 1b). The E113$^{3.28}$Q mutation increased wild-type arrestin binding by twofold to threefold, and the E113$^{3.28}$Q/ M257$^{6.40}$Y 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 30-fold increase of 3A arrestin binding to the E113$^{3.28}$Q/M257$^{6.40}$Y receptor. All-*trans*-retinal further increased 3A arrestin binding to the E113$^{3.28}$Q/M257$^{6.40}$Y 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$_8$-tagged rhodopsin and biotin-tagged arrestin. Wild-type arrestin interacted weakly with the E113$^{3.28}$Q/M257$^{6.40}$Y rhodopsin, regardless of the presence of all-*trans*-retinal (Fig. 1c). As a positive control, the GαCT-HA peptide, a high affinity peptide variant of the C terminus of G-transducin (G$_t$)[30], readily interacted with the E113$^{3.28}$ Q/M257$^{6.40}$Y 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 GαCT-HA with the E113$^{3.28}$Q/M257$^{6.40}$Y receptors revealed an IC$_{50}$ value of 15 nM and 700 nM for the binding of 3A arrestin and the GαCT-HA peptide, respectively (Fig. 1d and Extended Data Fig. 1d). The strength of the interaction between the E113$^{3.28}$Q/M257$^{6.40}$Y rhodopsin and the 3A arrestin is in a similar range as the estimated K_d value of 30–80 nM for the binding of arrestin to the fully activated phosphorylated rhodopsin[31].

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[32,33]. Therefore, we engineered a fusion protein in which 3A arrestin is linked by a 15-residue linker to the C terminus of E113$^{3.28}$ Q/M257$^{6.40}$Y 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. 2a). The T4L–rhodopsin–arrestin fusion protein is monomeric and relatively stable with a T_m of 59 °C (Extended Data Fig. 2b, c). Negative stain electron microscopy images revealed that E113$^{3.28}$Q/M257$^{6.40}$Y rhodopsin and 3A arrestin form a stable complex with arrestin bound to the cytoplasmic side of rhodopsin (Fig. 1e). The T4L–rhodopsin–arrestin fusion protein formed crystals with sizes in the range of 5 to 15 μm under various lipid cubic phase (LCP) crystallization conditions (Extended Data Fig. 2d, e). Despite extensive optimization, the crystals diffracted only to 6–8 Å at synchrotron sources (Extended Data Fig. 2f). We thus turned our attention to the emerging method of SFX[34] with an LCP injector (LCP-SFX)[35,36].

Structure determination by SFX

Because of the small size of crystals, we hypothesized that diffraction by X-ray free electron laser (XFEL) at the Linac Coherent Light Source (LCLS) would improve data quality given the advantages of intense and very short XFEL pulses for micrometre-size crystals. In the LCP-SFX method, a stream of gel-like LCP with fully hydrated microcrystals runs continuously in vacuum across the 1.5-μm-diameter XFEL beam, which delivers 120 X-ray pulses per second with less than 50 fs pulse duration and sufficient intensity to capture crystal diffraction patterns with a single pulse. Within ~12 h of run time, we collected over 5 million detector frames, of which 22,262 had more than 40 diffraction spots as determined by the Cheetah hit-finding software[37].

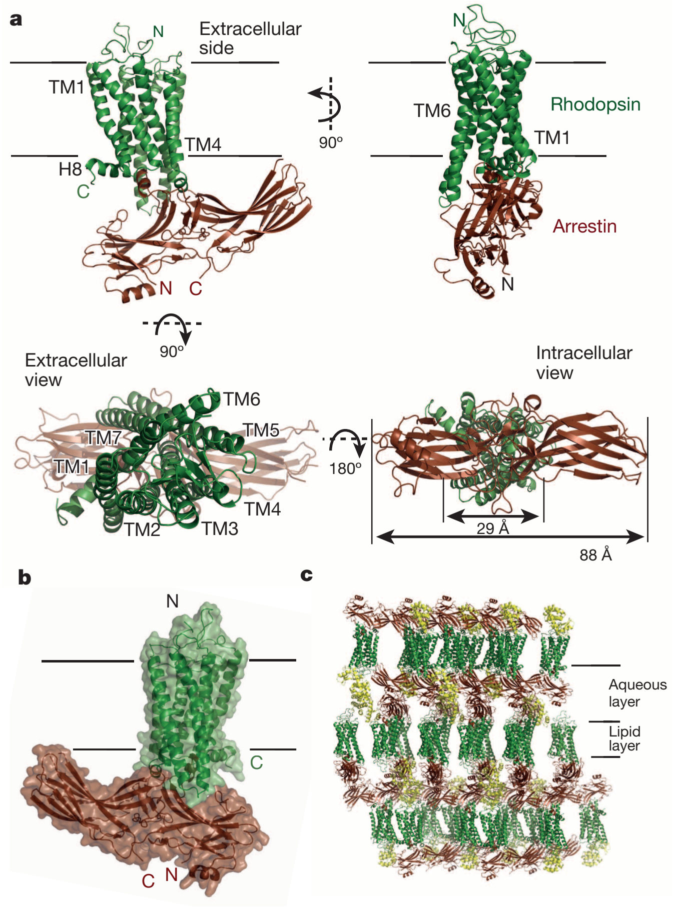


Figure 2 | The structure of the rhodopsin–arrestin complex. a, The structure of the rhodopsin–arrestin complex in four orientations. The relative dimensions of rhodopsin and arrestin are shown in the intracellular view. TM1–TM7 indicates rhodopsin transmembrane helices 1–7; H8 is intracellular helix 8. **b**, An overall view of the rhodopsin–arrestin complex shown with transparent solid surface. T4 lysozyme (T4L) is omitted from this view. **c**, Crystal packing diagram of the rhodopsin–arrestin complex with T4L as yellow ribbon model.

Diffraction patterns from 18,874 crystals could be indexed and integrated using CrystFEL[38]. 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 $P2_12_12_1$ (Supplementary Table 2) and the structure was solved by molecular replacement using known structures of active rhodopsin[39] and pre-activated arrestin[1] (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 R_work and R_free 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[40]. The DEER distances from residue Y74$^{2.41}$ of 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[41], 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 Å^2, which is substantially smaller than the area (2,576 Å^2) buried in the β_2AR–G$_\mathrm{s}$ complex[8]. Unlike the continuous interface observed in the β_2AR–G$_\mathrm{s}$ complex, the rhodopsin–arrestin complex has four distinct

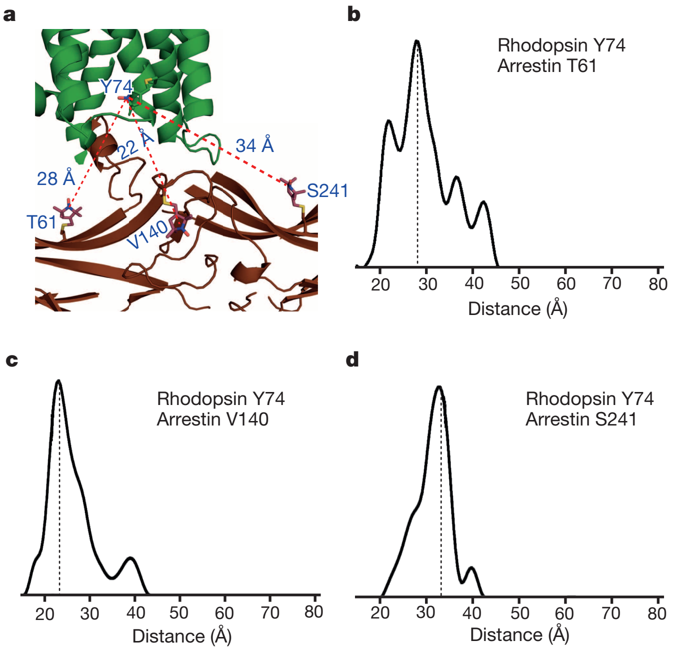


Figure 3 | DEER validation of rhodopsin–arrestin complex assembly. a, An overall view of rhodopsin–arrestin assembly showing the three intermolecular distances based on the models of the R1 nitroxide pairs at rhodopsin residue Y74$^{2.41}$ and three arrestin residues T61, V140, and S241 based on the crystal structure. **b**–**d**, The experimental distance distributions between the nitroxide spin labelled R1 pairs of rhodopsin Y74$^{2.41}$ 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.

arrestin interface patches (Fig. 4a, b and Extended Data Fig. 4b). 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-biased signalling[14,15]. Moreover, the close interactions between rhodopsin's helix 8 and arrestin have been shown to be essential for high-affinity binding of arrestin to the activated rhodopsin[42]. 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 $2F_o - F_c$ 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[43,44].

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[45]. 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[46] 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).

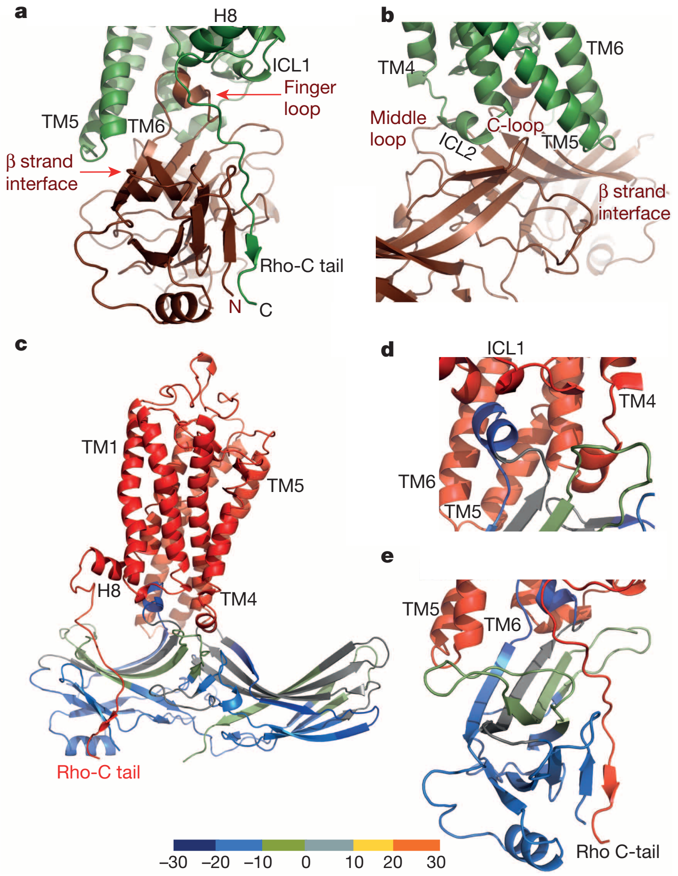


Figure 4 | The rhodopsin–arrestin interface and its validation by HDX. a, **b**, Two overall views showing the four interface patches of the rhodopsin–arrestin complex. **c**–**e**, Mapping of HDX on the rhodopsin-bound arrestin structure. Rhodopsin is coloured in red and arrestin is coloured based on 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.

The second set consisted of Tango assays[47], 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$^{3.28}$Q/M257$^{6.40}$Y 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

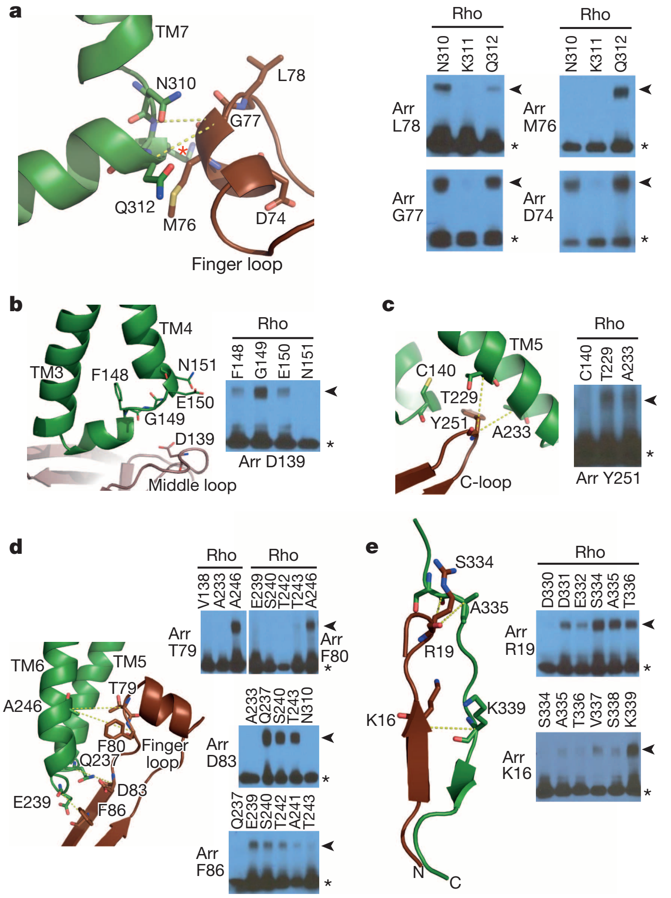


Figure 5 | Validation of the rhodopsin–arrestin interface by disulfide bond cross-linking. **a**–**e**, Structure and cross-linking of arrestin with rhodopsin. Panels are arrestin finger loop with rhodopsin TM7 and helix 8 (**a**); arrestin middle loop with rhodopsin ICL2 (**b**); arrestin C-loop residue 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 Cα atom of the finger loop residue G77 of arrestin to the Cα atoms of N310$^{7.57}$, K311$^{8.48}$ and Q312$^{8.49}$ in rhodopsin fit the requirement for disulfide bond formation (Fig. 5a). G77C cross-linked efficiently with N310$^{7.57}$C and Q312$^{8.49}$C, but not with K311$^{8.48}$ C because the Cβ of K311$^{8.48}$ 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 Q312$^{8.49}$C 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 T70$^{\text{ICL1}}$C and K67$^{\text{ICL1}}$C, respectively (Extended Data Fig. 7d).

We also observed cross-linking of the arrestin middle loop (D139) with rhodopsin ICL2 (G149$^{\text{ICL2}}$) (Fig. 5b), of the arrestin C-loop (Y251) with rhodopsin TM5 (T229$^{5.64}$ and A233$^{5.68}$) (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 with Q237$^{\text{ICL3}}$C of TM5 in rhodopsin (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 S334$^{\text{Cterm}}$ of rhodopsin (Fig. 5e), K16 of arrestin with S338$^{\text{Cterm}}$ and K339$^{\text{Cterm}}$ 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 GαCT and analogue peptides have been determined[11,22,23,39] and reveal that the arrangement of TM helices in light-activated rhodopsin is similar to that in the G-protein-bound β$_2$AR complex, with the exception of TM6, whose outward movement in β$_2$AR is much more pronounced upon binding to G protein[8]. The arrestin-bound rhodopsin has its intracellular end of TM6 moved outward by approximately 10 Å relative to its inactive structure (Fig. 6a, b and Extended Data Fig. 9). This is in contrast to the 14 Å outward movement of TM6 reported in the G-protein-bound β$_2$AR complex[8]. Compared to the active conformation of rhodopsin bound to GαCT peptides[11,22,23,39], 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 crosslinking 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 arrestin[48], 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 arrestin[43,48]. 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

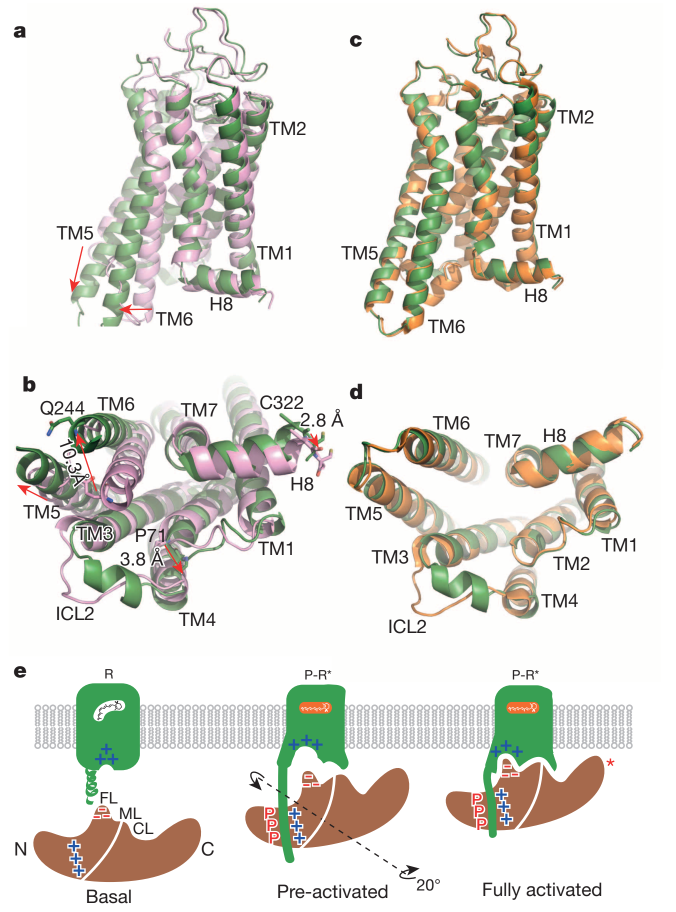


Figure 6 | Structural basis of arrestin-biased signalling and arrestin recruitment. a, **b**, Two views of structural overlays of arrestin-bound rhodopsin (green) with inactive rhodopsin (pink). **c, d**, Two views of structural overlays of arrestin-bound rhodopsin (green) with GαCT peptide-bound rhodopsin (orange). **e**, A cartoon model of arrestin recruitment by a phosphorylated and active 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).

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 (Fig. 2a, b). Such asymmetric assembly brings the arrestin C-domain towards the membrane, with the C-edge either being touched or embedded in the membrane layer (Extended Data Fig. 13). The C-edge is comprised of conserved hydrophobic residues (F197, F198, M199, F339, and L343). It has been puzzling why single alanine mutations at these residues would affect arrestin binding to rhodopsin given how far away they are from the receptor[43]. The close proximity of these hydrophobic residues to the membrane surface may provide an explanation for the effects of these mutations on rhodopsin binding. GPCR signalling regulator proteins are normally membrane-associated through lipid modifications, as is the case for GPCR kinase 1 (GRK1) and the G-protein subunits Gα and Gβγ. Yet, there is no known lipid modification for any of the arrestins. We speculate that the conserved hydrophobic patch at the C-tip of arrestin may function as a lipid-interacting module that helps to stabilize its interaction with the receptor. Furthermore, one primary function of arrestin is to mediate endocytosis of ligand-activated GPCRs and the highly asymmetric nature of the rhodopsin–arrestin assembly may facilitate the membrane curvature for subsequent endocytotic processes. Alternatively, the remote C-tip could serve as the binding site of a second rhodopsin, which has been proposed to form dimers in the rod outer-segment disc membrane[49].

LCP-SFX is a new technology that has been used to determine several crystal structures[35,36,50]. Rhodopsin–arrestin is a challenging membrane protein complex and obtaining a structure of this complex at a sufficiently high resolution was an intractable task using existing methods that include synchrotron-based crystallography and cryo-electron microscopy[26]. The rhodopsin–arrestin complex structure reported here demonstrates the utility of X-ray lasers when combined with SFX and an LCP crystal delivery system[35]. The SFX method is relatively new and under continuous development. Given its success in solving the rhodopsin–arrestin structure, we expect that X-ray lasers, with further method development, will continue to provide breakthrough insights into biology and chemistry.

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-kDa 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 signalling 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 β_2 adrenergic receptor-Gs protein complex. *Nature* **477,** 549–555 (2011).
9. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2-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 β_2-adrenergic receptor characterized by ^{19}F-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, X. F., 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. *Biochem. Soc. Trans.* **40,** 389–393 (2012).
19. Han, M., Smith, S. O. & Sakmar, T. P. Constitutive activation of opsin 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).

22. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455,** 497–502 (2008).
23. Choe, H. W. *et al.* Crystal structure of metarhodopsin II. *Nature* **471,** 651–655 (2011).
24. Hirsch, J. A., Schubert, C., Gurevich, V. V. & Sigler, P. B. The 2.8 Å crystal structure of visual arrestin: a model for arrestin's regulation. *Cell* **97,** 257–269 (1999).
25. Granzin, J. *et al.* X-ray crystal structure of arrestin from bovine rod outer segments. *Nature* **391,** 918–921 (1998).
26. Shukla, A. K. *et al.* Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature* **512,** 218–222 (2014).
27. Standfuss, J., Zaitseva, E., Mahalingam, M. & Vogel, R. Structural impact of the E113Q counterion mutation on the activation and deactivation pathways of the G protein-coupled receptor rhodopsin. *J. Mol. Biol.* **380,** 145–157 (2008).
28. Xie, G., Gross, A. K. & Oprian, D. D. An opsin mutant with increased thermal stability. *Biochemistry* **42,** 1995–2001 (2003).
29. Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.* **372,** 1179–1188 (2007).
30. Martin, E. L., Rens-Domiano, S., Schatz, P. J. & Hamm, H. E. Potent peptide analogues of a G protein receptor-binding region obtained with a combinatorial library. *J. Biol. Chem.* **271,** 361–366 (1996).
31. Zhuang, T. *et al.* Involvement of distinct arrestin-1 elements in binding to different functional forms of rhodopsin. *Proc. Natl Acad. Sci. USA* **110,** 942–947 (2013).
32. Bayburt, T. H. *et al.* Monomeric rhodopsin is sufficient for normal rhodopsin kinase (GRK1) phosphorylation and arrestin-1 binding. *J. Biol. Chem.* **286,** 1420–1428 (2011).
33. Hanson, S. M. *et al.* Each rhodopsin molecule binds its own arrestin. *Proc. Natl Acad. Sci. USA* **104,** 3125–3128 (2007).
34. Boutet, S. *et al.* High-resolution protein structure determination by serial femtosecond crystallography. *Science* **337,** 362–364 (2012).
35. Weierstall, U. *et al.* Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature Commun.* **5,** 3309 (2014).
36. Liu, W. *et al.* Serial femtosecond crystallography of G protein-coupled receptors. *Science* **342,** 1521–1524 (2013).
37. Barty, A. *et al.* software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47,** 1118–1131 (2014).
38. White, T. A. *et al.* CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45,** 335–341 (2012).
39. Deupi, X. *et al.* Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc. Natl Acad. Sci. USA* **109,** 119–124 (2012).
40. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl Acad. Sci. USA* **105,** 7439–7444 (2008).
41. Kim, M. *et al.* Conformation of receptor-bound visual arrestin. *Proc. Natl Acad. Sci. USA* **109,** 18407–18412 (2012).
42. Kirchberg, K. *et al.* Conformational dynamics of helix 8 in the GPCR rhodopsin controls arrestin activation in the desensitization process. *Proc. Natl Acad. Sci. USA* **108,** 18690–18695 (2011).
43. Ostermaier, M. K., Peterhans, C., Jaussi, R., Deupi, X. & Standfuss, J. Functional map of arrestin-1 at single amino acid resolution. *Proc. Natl Acad. Sci. USA* **111,** 1825–1830 (2014).
44. Sommer, M. E., Hofmann, K. P. & Heck, M. Distinct loops in arrestin differentially regulate ligand binding within the GPCR opsin. *Nature Commun.* **3,** 995 (2012).
45. West, G. M. *et al.* Protein conformation ensembles monitored by HDX reveal a structural rationale for abscisic acid signaling protein affinities and activities. *Structure* **21,** 229–235 (2013).
46. Ohguro, H., Palczewski, K., Walsh, K. A. & Johnson, R. S. Topographic study of arrestin using differential chemical modifications and hydrogen-deuterium exchange. *Protein Sci.* **3,** 2428–2434 (1994).
47. Barnea, G. *et al.* The genetic design of signaling cascades to record receptor activation. *Proc. Natl Acad. Sci. USA* **105,** 64–69 (2008).
48. Gurevich, V. V. & Benovic, J. L. Visual arrestin interaction with rhodopsin. Sequential multisite binding ensures strict selectivity toward light-activated phosphorylated rhodopsin. *J. Biol. Chem.* **268,** 11628–11638 (1993).
49. Fotiadis, D. *et al.* Atomic-force microscopy: rhodopsin dimers in native disc membranes. *Nature* **421,** 127–128 (2003).
50. Zhang, H. *et al.* Structure of the angiotensin receptor revealed by serial femtosecond crystallography. *Cell* **161,** 833–844 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements Portions of this research were carried out at the Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory. Use of the LCLS at the SLAC National Accelerator Laboratory is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. Parts of the sample injector used at LCLS for this research was funded by the National Institutes of Health, P41GM103393, formerly P41RR001209. We thank staff members of the Life Science Collaborative Access Team (ID-21) of the Advanced Photon Source (APS) for assistance in data collection at the beam lines of sector 21, which is in part funded by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817), and the General Medicine Collaborative Access Team for assistance in data collection at the beam lines of sector 23 (ID-23), funded in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). Use of APS was supported by the Office of Science of the US Department of Energy, under contract no. DE-AC02-06CH11357. This work was supported in part by the Jay and Betty Van Andel Foundation, Ministry of Science and Technology (China) grants 2012ZX09301001 and 2012CB910403, 2013CB910600, XDB08020303, 2013ZX09507001, Amway (China), National Institute of Health grants, DK071662 (H.E.X.); GM073197 and GM103310 (C.S.P. and B.C.); GM102545 and GM104212 (K.M.); EY011500 and GM077561 (V.V.G.), EY005216 and P30 EY000331 (W.L.H.), the National Institutes of Health Common Fund in Structural Biology grants P50 GM073197 (V.C. and R.C.S.), P50 GM073210 (M.C.), and GM095583 (P.F.); National Institute of General Medical Sciences PSI: Biology grants U54 GM094618 (V.C., V.K., and R.C.S.), GM108635 (V.C.), U54 GM094599 (P.F.), GM097463 (J.S.), and U54 GM094586 (JCSG); NSF Science and Technology Center award 1231306 (J.C.H.S., P.F. and U.W.); Swiss National Science Foundation grant 31003A_141235 (J.S.); the Canada Excellence Research Chair program and the Anne & Max Tanenbaum Chair in Neuroscience at the University of Toronto (O.P.E.); and Science Foundation Ireland, grant 12/IA/1255 (M.C.). Parts of this work were also supported by the Helmholtz Gemeinschaft, the DFG Cluster of Excellence Center for Ultrafast Imaging, and the BMBF project FKZ 05K12CH1 (H.N.C., A.B., C.G., O.Y., T.W.); the Irene and Eric Simon Brain Research Foundation (R.L.). We thank A. Brunger and O. Zeldin for analysing the XFEL data and for advising on refinement; B. Weis for advice on twin refinement and structure validation; J. Rini for advice on the piggyBac expression system; A. Lebedev for his advice regarding the Zanuda program and the choice of the space group; and A. Walker for final editing of the manuscript. C.G. kindly thanks the PIER Helmholtz-Graduate School and the Helmholtz Association for financial support. We also thank the TianHe research and development team of National University of Defense Technology (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 synchrotron data, helped with the SFX data collection, processed the data, and solved the structures; X. Gao expressed and purified rhodopsin–arrestin complexes, characterized their binding and thermal stability, discovered the initial crystallization conditions with 9.7 MAG (1-(9Z-hexadecenoyl)-*rac*-glycerol), prepared most crystals for synchrotron data collection, prepared all crystals for the final data collection by SFX, helped with SFX data collection, and established the initial cross-linking method for the rhodopsin–arrestin complex; Y.H. designed and performed Tango assays and disulfide bond cross-linking experiments; C.Z. developed the mammalian expression methods; P.W.d.W. helped with XFEL data processing and performed computational experiments; J.K., M.H.E.T., K.M.S.-P., K.P., J.M., Y.J., X.Z., and X. Gu performed cell culture, mutagenesis, protein purification, rhodopsin–arrestin binding experiments; W.L. and A.I. grew crystals and collected synchrotron data at APS and SFX data at LCLS, G.W.H. and Q.X. determined and validated the structure. Z.Z. and V.K. constructed the full model, the phosphorylated rhodopsin–arrestin model, and helped writing the paper; D.W., S.L., D.J., C.K., Sh.B., and N.A.Z. helped with XFEL data collection and initial data analysis; Sé.B., M.M., and G.J.W. set up the XFEL experiment, performed the data collection, and commented on the paper. A.B., T.A.W., C.G., O.Y., and H.N.C. helped with XFEL data collection and data analysis, processed the data and helped with structure validation. G.M. W., B.D.P., and P.R.G. performed HDX experiments and helped with manuscript writing. J.L. helped initiate this collaborative project and with writing the paper. M.W. collected the 7.7 Å dataset at the Swiss Light Source. A.M., C.S.P., and B.C. were responsible for electron microscopy images of rhodopsin–arrestin complexes. M.T. and Y.Z. performed mass spectrometry experiments to validate the protein contents in the crystals; D.L., N. H., and M.C. provided the 9.7 MAG phase diagram and helped with SFX data collection and with writing the paper. J.S. provided a computational model of the rhodopsin–arrestin complex and helped with discussion and writing; K.D., H.L., and Y.D. helped with data analysis and twinning problems; R.J.L. constructed single-Cys arrestin-1 mutants for DEER and tested their binding to rhodopsin; S.A.V. expressed these mutants in *Escherichia coli* and purified them; V.V.G. provided arrestin genes, designed single-Cys arrestin-1 mutants for DEER, and helped analysing the data and writing the paper. H.Y. and H.J. performed computational modelling, figure preparation, and helped with writing the paper; J.C.H.S. and U.W. designed the LCP injector and helped with data collection. Sh.B., S.R.-C., C.E.C., J.C., C.K., I.G., P.F., and R.F. helped with data collection, on-site crystal characterization as well as data analysis, and validation of the structure. L.N.C. and O.P.E. generated the Y74C/C140S/C316S stable cell line, characterized and provided the rhodopsin mutant sample for DEER measurements. N.V.E. and W.L.H. incorporated rhodopsin into nanodiscs, spin-labelled rhodopsin and arrestin, performed DEER experiments and helped with manuscript writing. R.C.S. supervised crystal growth, data collection, structure solution and validation, and helped with manuscript writing. V.C. was the Principal Investigator of the LCLS data collection, supervised crystal growth, data collection at APS and LCLS, structure solution and validation, and helped with manuscript writing; K.M. supervised research, analysed data, and helped with writing the paper. H.E.X. conceived the project, designed the research, performed synchrotron and LCLS data collection and structure solution, and wrote the paper with contributions from all authors.

Author Information The coordinates of the rhodopsin–arrestin complex and diffraction data have been deposited in the Protein Data Bank under accession number 4ZWJ. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.E.X. (Eric.Xu@vai.org).

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-inducible expression cassette encoding a fusion protein with His$_8$–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: N2$^{\mathrm{Nterm}}$C and N282$^{\mathrm{ECL3}}$C to form a disulfide bond, and E113$^{3.28}$Q and M257$^{6.40}$Y for constitutive receptor activity. The C terminus of rhodopsin was fused to 3A arrestin (L374A, V375A, F376A, residues 10–392) with a 15 amino acid linker (AAAGSAGSAGSAGSA).

The fusion protein constructs were expressed in HEK293S cells (Invitrogen) transiently transfected using Lipofectamine 2000 (Invitrogen). Cells were transfected at a density of 2×10^6 cells per ml at a 100 ml scale in SFM4TransFx-293 (HyClone). Six hours post-transfection, cells were diluted tenfold with CDM4HEK293 medium (HyClone). When the cell density reached approximately 4×10^6 cells per ml, protein expression was induced by the addition of doxycycline to a final concentration of 1 µg ml^{-1}. After 24 h induction, cells were harvested, resuspended in hypotonic buffer (20 mM HEPES-Na, pH 7.5, 10 mM NaCl, 10 mM MgCl$_2$) supplemented with EDTA-free protease inhibitor cocktail (Roche), followed by Dounce homogenization. The lysate was centrifuged at 45,000 r.p.m. at 4 °C for 1 h. The membranes were solubilized in 20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10% glycerol, 0.5% (w/v) *n*-dodecyl-β-D-maltopyranoside (DDM, Anatrace), 0.1% (w/v) cholesteryl hemisuccinate (CHS, Anatrace), and protease inhibitor cocktail for two hours at 4 °C. The supernatant was isolated by centrifugation at 45,000 r.p.m. for one hour, and incubated with amylose beads (New England Biolabs) at 4 °C overnight. Typically 10 ml of resin were used for supernatant from one litre of the original cell culture. The resin was washed with 200 ml of washing buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 0.005% (w/v) lauryl maltose neopentyl glycol (MNG-3, Anatrace), 0.001% (w/v) CHS) and the protein was eluted with washing buffer containing 10 mM maltose. The eluted fusion protein was concentrated to 40–50 mg per 1 ml and the His$_8$–MBP–MBP tag was cleaved by overnight incubation with 3C protease at 4 °C. The cleaved His$_8$–MBP–MBP tag was removed by incubating with 300 µl Ni^{2+}-NTA beads (Qiagen) for three hours at 4 °C. The purified T4L–rhodopsin–arrestin was collected and concentrated to 30 mg ml^{-1} for crystallization.

Wild-type rhodopsin and rhodopsin mutants E113$^{3.28}$Q and E113$^{3.28}$Q/M257$^{6.40}$Y used for MBP pull-down and AlphaScreen assays were expressed from the same vector with a His$_8$–MBP tag at the N terminus. Protein expression and purification were similar as for the rhodopsin–arrestin fusion protein used for crystallization, with the difference that 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 tag into the first expression cassette of a modified pET-Duet expression vector (Novagen), and the biotin ligase gene *BirA* 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[51]. BL21 (DE3) cells transformed with the expression plasmid were grown in LB broth at 16 °C to an OD$_{600}$ of ~1.0 and induced with 0.1 mM IPTG in the presence of 40 µM biotin for 16 h. Cells were harvested, resuspended in 50 ml extraction buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 10% glycerol) per two litres of cells, and passed three times through a French Press with pressure set at 1,000 Pa. The lysate was centrifuged at 16,000 r.p.m. in a Sorvall SS-34 rotor for 30 min, and the supernatant was loaded on a 5 ml amylose HP column (GE Healthcare). The column was washed with 100 ml of wash buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 10% glycerol) and eluted in buffer containing 20 mM Tris-HCl, pH 8.0, 150 mM NaCl and 20 mM maltose. The eluted biotin–MBP–arrestin was concentrated and further purified by size-exclusion chromatography through a HiLoad 16/60 Superdex 200 column (GE Healthcare) in 20 mM Tris-HCl, pH 8.0, and 150 mM NaCl. Monomeric protein was collected for further assay.

Untagged 3A arrestin (residues 10–392) was expressed as a His$_6$–SUMO fusion protein from the expression vector pSUMO (LifeSensors). The expression and purification of untagged arrestin followed the same method as for biotin–MBP–arrestin, except that the His$_6$–SUMO tag was cleaved overnight with SUMO protease[52] at a protease/protein ratio of 1:1,000 in the cold room.

MBP pull-down assay. Arrestin was cloned into the pCITE-4a vector (Novagen) 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 [^{35}S]methionine-labelled wild-type arrestin and 3A arrestin (L374A, V376A, and F376A). Radiolabelled wild-type arrestin and mutant arrestin proteins were incubated with His$_8$–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 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 peptide. Interactions between rhodopsin and arrestin were assessed by luminescence-based AlphaScreen assay (Perkin Elmer), which our group has used extensively to determine ligand-dependent protein–protein interactions of nuclear receptors. The AlphaScreen principle is illustrated in Extended Data Fig. 1c. Briefly, biotinylated arrestin was bound to streptavidin-coated donor beads and His8-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%. GαCT-HA (TGGRVLEDLKSCGLF) and biotinylated GαCT-HA were synthesized by Peptide 2.0. For the competition assay, different amounts of untagged arrestin or GαCT-HA were added to the reaction to compete with tagged arrestin or GαCT-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-coumariny)phenyl]maleimide (CPM) as described previously[53]. 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, while 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 immediately 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$_{\mathrm{m}}$) as inflection point of the melting curves.

Crystallization. T4L–rhodopsin–arrestin crystals were grown in lipid cubic phase (LCP)[54]. Protein solution (~30 mg ml^{-1}) was mixed with monopalmitolein (9.7 MAG, from Nu-Chek Prep, Inc.) containing 10% cholesterol at a 1:1 ratio (w/w) using a coupled syringe mixer[55] and 50 nl boluses of protein-laden LCP were dispensed on 96-well glass sandwich plates (Molecular Dimensions or Marienfeld-Superior) and overlaid with 0.8 µl precipitant solutions using a Gryphon LCP robot (Art Robbins Instruments) or an NT8-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)[56]. Crystals that reached full size (around 10–20 µm) within four days at 20 °C were harvested from the mesophase and were flash frozen in liquid nitrogen without additional cryoprotectant. Crystals used for synchrotron 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[36,57]. 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[55]. Crystals were grown in several syringes at 20 °C, consolidated and transferred into the LCP injector[35] for XFEL diffraction data collection. The average crystal size was 5–10 µm as determined under a polarized light microscope. The phase diagram for 9.7 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].

Data collection (synchrotron). A partial 8.0 Å synchrotron data set was collected at 100 K using an X-ray beam at the wavelength of 1.0 Å at the 21 ID-D beam line of LS-CAT and at 23-ID-D of GM-CAT at the Advanced Photon Source at

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[58] 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[59] at the Linac Coherent Light Source (LCLS) in the SLAC National Accelerator Laboratory (Menlo Park, California, USA). X-ray pulses of 50 fs duration at a wavelength of 1.3 Å (9.5 keV) were attenuated to ~3% (3×10^{10} photons per pulse) and focused to ~1.5 μm diameter at the interaction point using Kirkpatrick–Baez mirrors[60]. Rhodopsin–arrestin complex crystals in LCP were injected across the XFEL beam using an LCP injector[35] 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[61] (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[38] 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 merged using a Monte Carlo integration across the crystal rocking curve of each reflection[64]. 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 truncated at 3.8 Å/3.8 Å/3.3 Å using the get_hkl program of CrystFEL[38] 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 diffraction[65]. The resolution cutoff for several published XFEL structures follows this criterion, including the $5HT_{2B}$ GPCR XFEL structure[36], which has a CC* of 0.74. The statistics of the final data used in structure refinement are shown in Supplementary Table 1b.

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 $P4_3$. Based on analysis of the Zanuda program[66], we determined that the most likely space group was $P2_12_12_1$ and that the crystals were physically twinned. The data were reprocessed with the Laue group of mmm and molecular replacement searches were performed in $P2_12_12_1$. This space group assignment resulted in the best statistics and map quality out of several possible space groups (Supplementary Table 2).

The crystals appeared to be pseudo-merohedrally twinned based on L-test analysis[67]. Despite the challenge of twinned data, the rhodopsin–arrestin complex structure was solved by the molecular replacement method implemented in Phaser[68] using the models of constitutively active rhodopsin, pre-activated arrestin, and T4L (PDB codes: 4A4M[39], 4J2Q[1], and 3SN6[8], 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 residue in a position to form a covalent bond with the first residue of rhodopsin, supporting the correct positioning of T4Ls by molecular replacement.

The structure was initially refined against the XFEL data without twin law to an R_{free}-factor of ~36% and the model maps from the data were of sufficient quality to interpret the overall structure of the rhodopsin–arrestin complex (Extended Data Fig. 3). The model then underwent iterated cycles of manual building into $2F_o - F_c$ maps with Coot[69] and refinement with REFMAC[70] and the PHENIX[71], where rigid body, individual position, group B-factor, and TLS refinements were used along with NCS restraints and twin law (k, h, -l). The arrestin residues 70–78 and 165–175, which were not included in the molecular replacement model, were manually placed into the density map. These two regions of arrestin became visible because they are either engaged in direct interaction with rhodopsin or involved in crystal packing. Regions with poor density were removed from the final model, including T4L from the B complex and the N-domain of T4L (residues 13–57) from the C complex. We did not observe clear electron density for the all-*trans*-retinal, which was thus not included in the structure. The structure has been carefully refined to the final state that has excellent geometry and refinement statistics (Supplementary Table 1b). Ramachandran plot analysis indicates that 100% of the residues are in favourable or allowed regions (no outliers). The final structure was validated with MolProbity, which revealed an all-atom-clash score of 1.47 and MolProbity score of 1.13[72].

The real-space correlation coefficients against a $2mF_o - DF_c$ map for each chain of the structure on a per residue basis using the CCP4 EDSTATS program[73] or the MolProbity program in Phenix indicated an overall good fit between the structure and the electron density map. The density fit correlation in Coot was low with the structure from Phenix twinned refinement because the map from the MTZ file with map coefficients produced by phenix.refine was not in absolute scale. This problem was overcome by using the $2F_o - F_c$ CNS format map from phenix.refine, which generated a normal correlation in Coot, similar to those from EDSTATS and MolProbity. All structural figures were prepared using PyMOL[74].

Cell-based assays for rhodopsin–arrestin interactions (Tango assays). For the cell-based Tango assay[47], we generated fusion constructs consisting of rhodopsin (1-321), a tobacco etch virus (TEV) protease cleavage site (TEV site), and the transcriptional activator tTA (Rho–TEV site–rTA) as well as of arrestin (10–392) and TEV protease (Arr–TEV protease). HTL cells were a gift from G. Barnea and R. Axel (Brown University and Columbia University). 10 ng Rho–TEV site–tTA construct were transfected together with 10 ng Arr–TEV protease plasmid and 1 ng of phRG-tk *Renilla* luciferase expression vector into HTL cells using Xtremgene (Roche). One day after transfection, cells were induced by vehicle (DMSO) or all-*trans*-retinal (10 μM) overnight. Cells were harvested and lysed in Passive Lysis Buffer (Promega). Luciferase activity was measured using the Dual Luciferase Kit (Promega) according to manufacturer's instructions.

Mutagenesis. Site-directed mutagenesis was carried out using the QuikChange method (Agilent). Mutations and all plasmid constructs were confirmed by sequencing before protein expression, MBP-pull down assay, Tango assay, and AlphaScreen assay.

Electron microscopy studies of the rhodopsin–arrestin complex. Samples were prepared as previously described[75]. Briefly, the sample was applied to a freshly glow-discharged carbon coated copper grid and allowed to adhere for 10 s before being reduced to a thin film by blotting. Immediately after blotting 3 μl of a 1% solution of uranyl formate was applied to the grid and blotted off directly. This was repeated three times. Data were acquired using a Tecnai F20 Twin transmission electron microscope operating at 200 kV, with a dose of ~40 e^- Å^{-2} and nominal underfocus ranging from 2 to 3 μm. Images were automatically collected at a nominal magnification of 62,000 × and pixel size of 0.273 nm. All images were recorded with a Tietz F416 4k × 4k pixel CCD camera using Leginon data collection software[76]. Experimental data were processed by the Appion software package[77], which interfaces with the Leginon database infrastructure. ~6,000 particles were automatically extracted from 54 electron microscopy micrographs[78] and the particle stack was then aligned and sorted using the XMIPP reference-free maximum likelihood alignment[79]. Several exemplary 2D averages are shown in Fig. 1e, which were derived from class averages for the complex with or without T4L, computed from ~14,000 particles selected from 155 electron microscopy-micrographs.

In-cell disulfide bond cross-linking. The open reading frames of full-length arrestin with 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 H_2O_2, which was freshly diluted in the cell culture medium to a final concentration of 1 mM. After 5 min treatment with H_2O_2, the medium was aspirated and 100 μl of CelLytic M (Sigma C2978) were 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 spun at 16,000*g* at 4 °C for 5 min. The supernatants (10 μl) were 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[80], 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) decyl maltose neopentyl glycol (DMNG) was used in place of DDM in the exchange buffer. Briefly, all stock solutions and dilutions were made using the 7TM 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 H_2O or in D_2O for on-exchange). All HDX protein stock solutions were prepared at 15 μM in the 7TM

HDX H_2O 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 D_2O on-exchange buffer. Exchange was quenched by adding 25 μl of quench solution (100 mM NaH_2PO_4, 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 pepsin column. Peptides were captured and desalted on a C8 trap. Peptides were then separated across a 5μ 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 the Workbench program[81] and converted to % deuterium values. Back-exchange correction was based on an estimated 70% deuterium recovery and accounting for the known 80% deuterium content of the on-exchange buffer. The Workbench software used *P* values lower than 0.05 for two consecutive time points to determine significance. Sequence coverage experiments for these proteins were carried out in the 7TM HDX buffer and LC system described above, but with a longer 60 min gradient. 75 pmol of protein were loaded onto the column. For sequencing, tandem mass spectra were obtained using data-dependent acquisition with 30 s dynamic exclusion, where the top five most abundant ions in each scan were selected and subjected to CID fragmentation. Each scan was the average of 3 microscans under normal scan mode in both MS and MS/MS. Peptides were identified by searching spectra against an in-house database using the mascot search engine as described previously[45].

Generation of a stable cell line expressing bovine rhodopsin mutant Y74C/C140S/C316S and rhodopsin preparation. The rhodopsin mutant Y74C/C140S/C316S gene[82] was PCR-amplified from the pMT vector by using two flanking primers containing NheI and NotI restriction sites and subcloned into vector PB-T-PAF[83]. The insert in the resulting plasmid termed PB-Rho-Y74C was verified by automated DNA sequencing.

Generation of the stable cell line as well as rhodopsin production and purification was performed as described[84]. HEK293S GnTI^{-} cells[85] were cultured in DMEM/F12 (Wisent) supplemented with 10% heat inactivated FBS (Life Technologies) and 100 U ml^{-1} penicillin/streptomycin (Life Technologies) and incubated at 37 °C, 5% CO_2. For transfection, cells were plated in 6-well plates at a density of ∼600,000 cells per well. The following day, cells were co-transfected with 4 μg of PB-Rho-Y74C, 0.5 μg of PB-RB[83] and 0.5 μg of pCyL43[83,86] using JetPrime (PolyPlus, France) reagent following the manufacturer's instructions and medium was replaced by fresh medium 4 h later. (Plasmids PB-T-PAF and PB-RB were provided by James Rini, University of Toronto, Canada, and pCyL43 by the Wellcome Trust Sanger Institute, UK). Two days after transfection, cells were transferred to a 100-mm tissue culture plate. Dual drug selection was started the following day, using 10 μg ml^{-1} puromycin (Bioshop, Canada) and 5 μg ml^{-1} blasticidin (Bioshop, Canada) and lasted for 2 weeks. Cells were then transferred to larger flasks and finally to pleated roller bottles (Thermo Scientific, USA) containing 250 ml of DMEM/F12 supplemented with 10% heat-inactivated FBS and 100 U ml^{-1} of penicillin and 100 μg ml^{-1} streptomycin. Roller bottles were incubated at 37 °C, 5% CO_2 and rotated at 0.1 r.p.m. Expression was induced after 5–6 days by replacing medium with 250 ml of fresh medium containing 1 μg ml^{-1} doxycycline (Bio Basic) and 1 μg ml^{-1} aprotinin (Bioshop Canada). On day 3 after induction, cells were harvested, using PBS-EDTA with protease inhibitor tablet (Roche) to detach cells. Cell pellets were flash frozen in liquid nitrogen and stored at −80 °C until purification.

For purification, cell pellets with rhodopsin mutant Y74C/C140S/C316S were incubated with 11-*cis* retinal to reconstitute rhodopsin which was purified and spin labelled with 1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl methanethiosulfonate (MTSSL) as previously described[41].

Incorporation of spin labelled rhodopsin into nanodiscs. Nanodiscs were prepared by mixing cholate solubilized lipid (70% POPC + 30% POPS), scaffolding protein MSP1E3D1[87], and rhodopsin (in 90 mM β-OG) in a molar ratio of 140:1:0.1. The OG and cholate were removed from the mixture by dialysis against buffer D (20 mM MOPS, 150 mM NaCl, pH 6.8). The nanodiscs were further purified by passing them over a nickel NTA column to remove any receptor not incorporated into the discs. Upon elution from the nickel column with imidazole, the nanodiscs were buffer exchanged into buffer D containing 10% glycerol.

Spin labelling of arrestin mutants. C-terminally truncated mutants of arrestin, lacking endogenous cysteines, have been shown to bind to non-phosphorylated rhodopsin[88]. Single-cysteine mutants in this background were expressed in *E. coli*, as described[89], spin labelled in buffer D overnight using a tenfold molar excess of MTSSL. The base mutant also contained two alanine substitutions (F85A and F197A) which have been shown to disrupt arrestin-1 oligomerization[89]. Non-covalently bound spin label was removed from the sample by extensive washes with buffer D using Amicon 10 kDa concentrators.

DEER spectroscopy of the rhodopsin–arrestin complex. Preparations of rhodopsin (Y74$^{2.41}$C) and bovine arrestin (S60C, V139C, and L240C, which correspond to the mouse arrestin T61C, V140C, and S241C) and their spin labelling were performed as described previously[41]. 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 cutoff filter. Immediately after irradiation, the samples were flash frozen in liquid nitrogen, and loaded into an EN 5107D2 resonator for Q band DEER measurements. Measurements were performed at 80 K on a Bruker Elexsys 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 swept spectrum, and the observer $\pi/2$ (16 ns) and π (32 ns) pulses were positioned 50 MHz (17.8 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/.

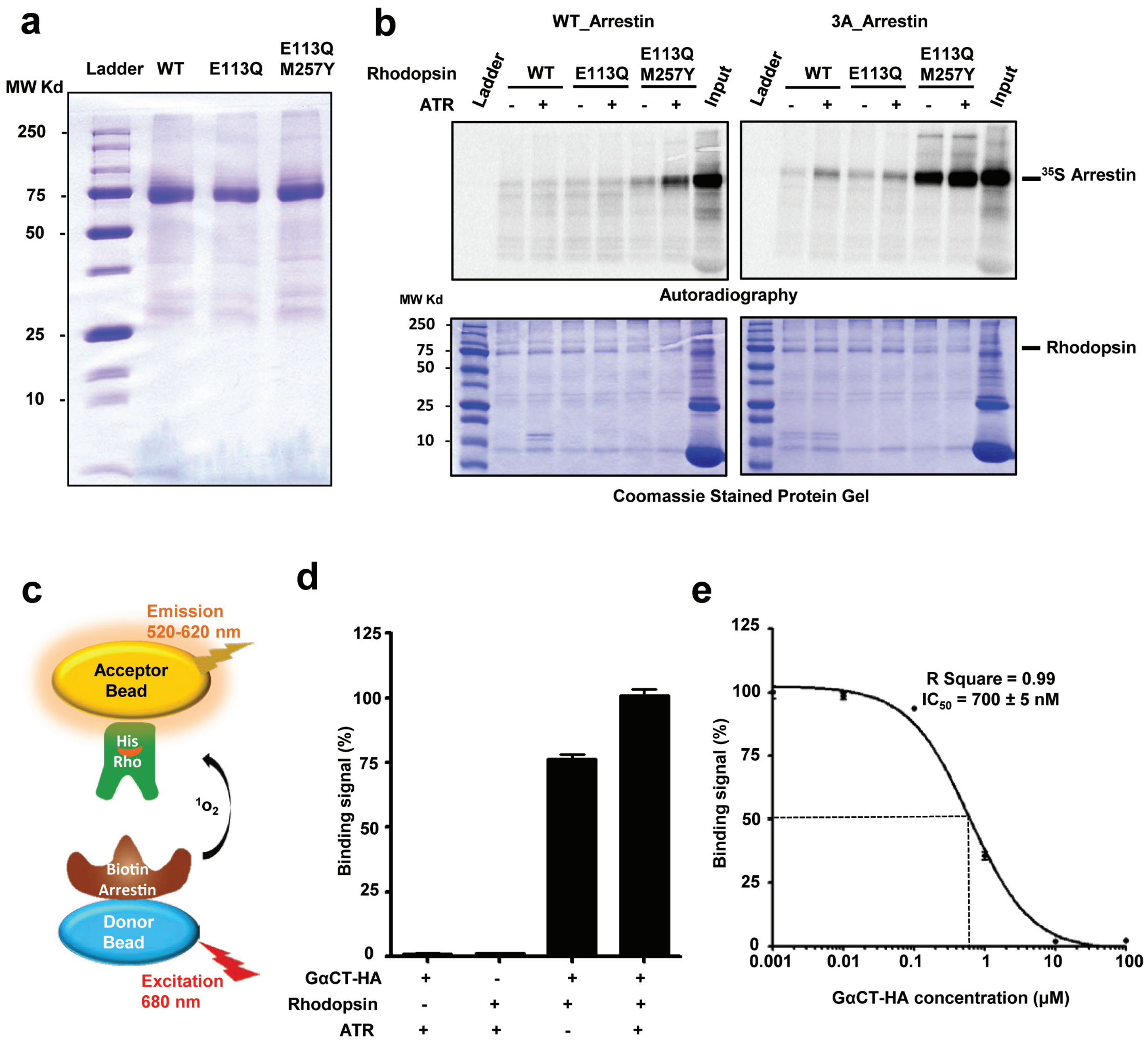
To estimate the median distances, the distance distributions were integrated and normalized to the maximum amplitude. The median distance was estimated as that corresponding to 0.5 of the integrated intensity. The modelled distances between nitroxide spin labels are based on the crystal structure of the rhodopsin–arrestin complex. R1 nitroxide side chains were modelled into the structure using common R1 rotamers[90,91].

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 described[55,92]. 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 (L_α), 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 (monoolein)[92] and is similar to what was found with 7.9 MAG[93]. 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 MAG[35]. 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 MAG[92] and 7.7 MAG[94], 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 β_2AR–Gs complex crystallization and structure determination[8].

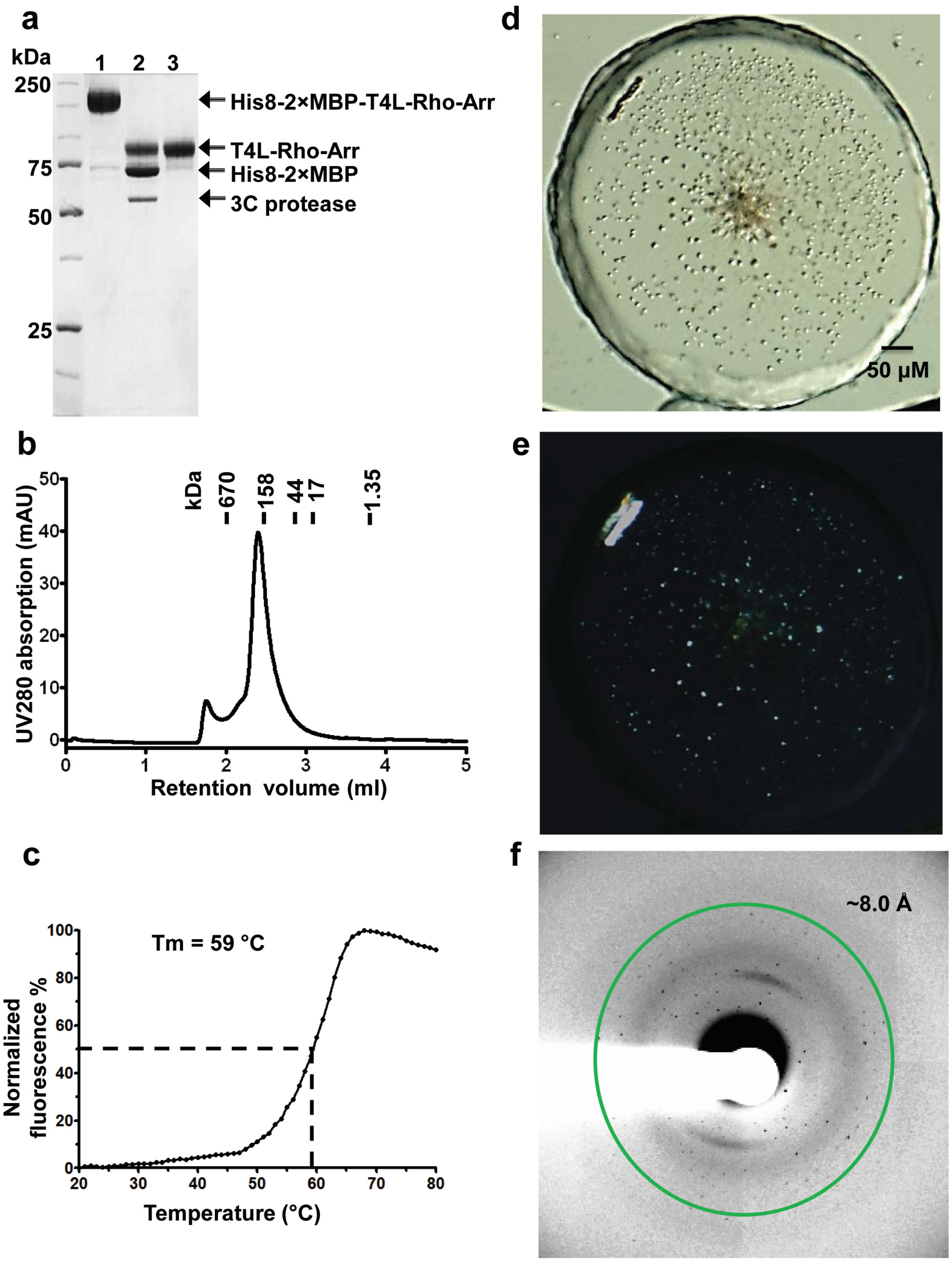
Molecular modelling of the full-length rhodopsin–arrestin complex. Energy-based conformational modelling of the rhodopsin–arrestin complex was performed with the ICM-Pro molecular suite[95], using a global energy optimization procedure similar to the one described recently for modelling of the full-length complex of CRFR1[96]. 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 327–345) of the human rhodopsin structure, as well as missing residues in the arrestin 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 crosslinking 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 5 Å distance, to medium at 7 Å distance and very weak at 12 Å. The C-terminal peptide conformation and conformation of the contact side chains of the arrestin were optimized to convergence (3 independent simulations of 10^6 steps) using global optimization procedure in internal coordinates with improved conformational energy terms for protein and peptides[97]. 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 D331, 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 terminus 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).

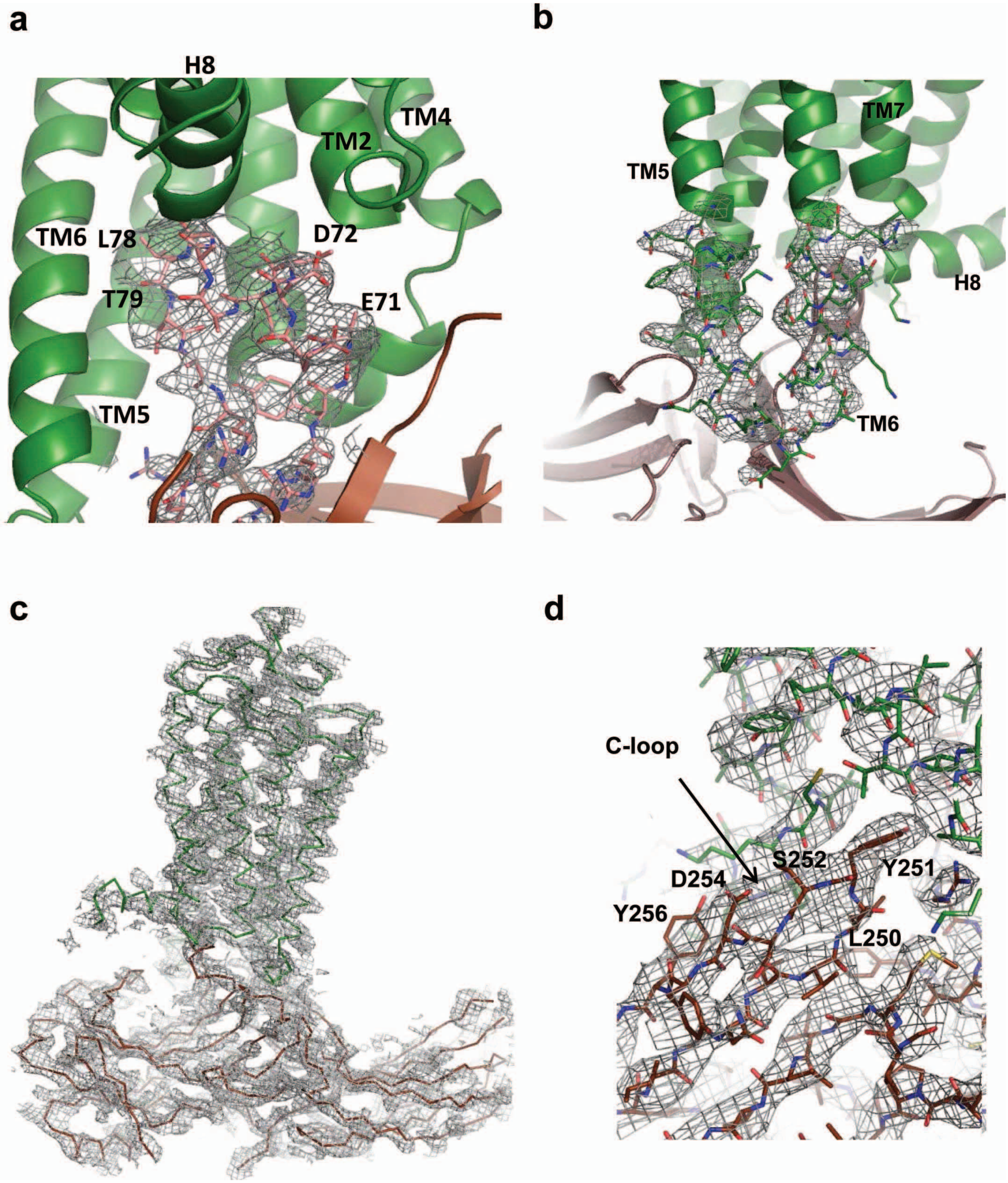
51. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* **8,** 921–929 (1999).
52. Mossessova, E. & Lima, C. D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element 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. Crystallizing membrane proteins using lipidic mesophases. *Nature Protocols* **4,** 706–731 (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 microcrystals 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,** 1231–1240 (2013).
62. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: 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 obstinate 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. Karplus, 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 origin 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,** 486–501 (2010).
70. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67,** 355–367 (2011).
71. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
72. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66,** 12–21 (2010).
73. Tickle, I. J. Statistical quality indicators for electron-density maps. *Acta Crystallogr. D* **68,** 454–467 (2012).
74. DeLano, W. L. & Lam, J. W. PyMOL: a communications tool for computational models. *Abstr. Pap. Am. Chem. Soc.* **230,** U1371–U1372 (2005).
75. Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B. & Wilson, I. A. Organization of the influenza virus replication machinery. *Science* **338,** 1631–1634 (2012).
76. Suloway, C. *et al.* Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151,** 41–60 (2005).
77. Lander, G. C. *et al.* Appion: An integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166,** 95–102 (2009).
78. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. DoG Picker and TiltPicker: Software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166,** 205–213 (2009).
79. Scheres, S. H. W., Nunez-Ramirez, R., Sorzano, C. O. S., Carazo, J. M. & Marabini, R. Image processing for electron microscopy single-particle analysis using XMIPP. *Nature Protocols* **3,** 977–990 (2008).
80. Goswami, D. *et al.* Time window expansion for HDX analysis of an intrinsically disordered protein. *J. Am. Soc. Mass Spectrom.* **24,** 1584–1592 (2013).
81. Pascal, B. D. *et al.* HDX Workbench: software for the analysis of H/D exchange MS data. *J. Am. Soc. Mass Spectrom.* **23,** 1512–1521 (2012).
82. Klein-Seetharaman, J. *et al.* Single-cysteine substitution mutants at amino acid positions 55-75, the sequence connecting the cytoplasmic ends of helices I and II in rhodopsin: reactivity of the sulfhydryl groups and their derivatives identifies a tertiary structure that changes upon light-activation. *Biochemistry* **38,** 7938–7944 (1999).
83. Li, Z., Michael, I. P., Zhou, D., Nagy, A. & Rini, J. M. Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. *Proc. Natl Acad. Sci. USA* **110,** 5004–5009 (2013).
84. Caro, L. N. *et al.* Rapid and facile recombinant expression of bovine rhodopsin in HEK293S GnTI(-) cells using a PiggyBac inducible system. *Methods Enzymol.* **556,** 307–330 (2015).
85. Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* **99,** 13419–13424 (2002).
86. Wang, W. *et al.* Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **105,** 9290–9295 (2008).
87. Bayburt, T. H., Leitz, A. J., Xie, G., Oprian, D. D. & Sligar, S. G. Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J. Biol. Chem.* **282,** 14875–14881 (2007).
88. Hanson, S. M. *et al.* Differential interaction of spin-labeled arrestin with inactive and active phosphorhodopsin. *Proc. Natl Acad. Sci. USA* **103,** 4900–4905 (2006).
89. Hanson, S. M. *et al.* Structure and function of the visual arrestin oligomer. *EMBO J.* **26,** 1726–1736 (2007).
90. Fleissner, M. R., Cascio, D. & Hubbell, W. L. Structural origin of weakly ordered nitroxide motion in spin-labeled proteins. *Protein Sci.* **18,** 893–908 (2009).
91. Lietzow, M. A. & Hubbell, W. L. Motion of spin label side chains in cellular retinol-binding protein: correlation with structure and nearest-neighbor interactions in an antiparallel beta-sheet. *Biochemistry* **43,** 3137–3151 (2004).
92. Qiu, H. & Caffrey, M. The phase diagram of the monoolein/water system: metastability and equilibrium aspects. *Biomaterials* **21,** 223–234 (2000).
93. Misquitta, Y. *et al.* Rational design of lipid for membrane protein crystallization. *J. Struct. Biol.* **148,** 169–175 (2004).
94. Misquitta, L. V. *et al.* Membrane protein crystallization in lipidic mesophases with tailored bilayers. *Structure* **12,** 2113–2124 (2004).
95. ICM. Manual v. 3.8 (MolSoft LLC, 2014).
96. Coin, I. *et al.* Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR. *Cell* **155,** 1258–1269 (2013).
97. Arnautova, Y. A., Abagyan, R. A. & Totrov, M. Development of a new physics-based internal coordinate mechanics force field and its application to protein loop modeling. *Proteins* **79,** 477–498 (2011).



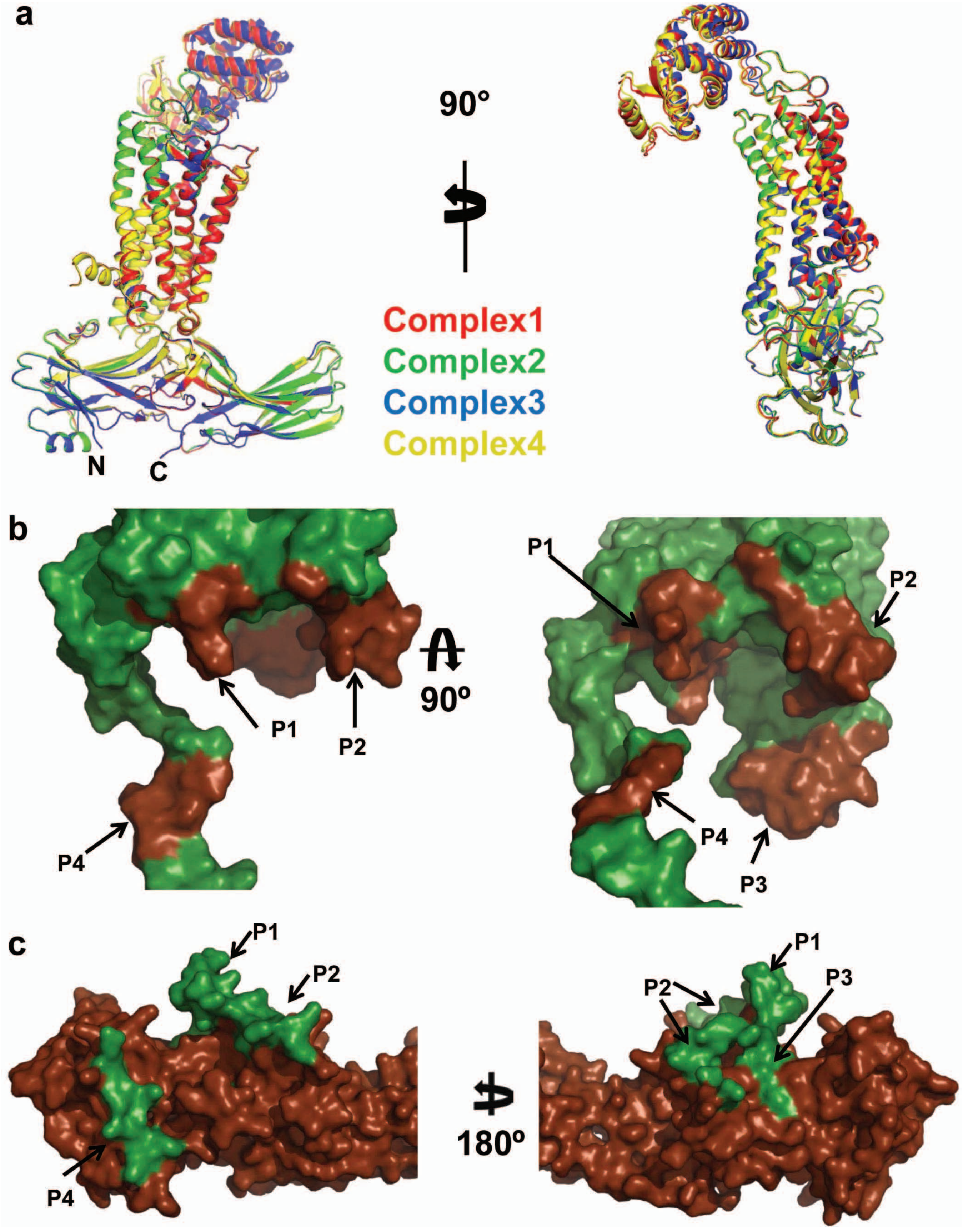
Extended Data Figure 1 | Constitutively active rhodopsin interacts with arrestin and GαCT-HA. **a**, SDS–PAGE of N-terminally MBP-tagged wild-type and mutant rhodopsin. **b**, Non-cropped versions of the pull-down assay gels shown in Fig. 1b. The interactions between mouse wild-type arrestin and human wild-type or E113$^{3.28}$Q rhodopsin are very weak. In contrast, the interaction between constitutively active rhodopsin (E113$^{3.28}$Q/M257$^{6.40}$Y) 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 E113$^{3.28}$Q/M257$^{6.40}$Y rhodopsin and GαCT-HA (TGGRVLEDLKSCGLF) 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 E113$^{3.28}$Q/M257$^{6.40}$Y and GαCT-HA by homologous competition. His$_6$–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 IC$_{50}$ of ~700 nM ($n = 3$, error bars, s.d.).


a
kDa
1
2
3
250
75
50
25
His8-2×MBP-T4L-Rho-Arr
T4L-Rho-Arr
His8-2×MBP
3C protease
b
UV280 absorption (mAU)
50
40
30
20
10
0
kDa
670
158
44
17
1.35
0
1
2
3
4
5
Retention volume (ml)
c
Normalized fluorescence %
100
80
60
40
20
0
Tm = 59 °C
20
30
40
50
60
70
80
Temperature (°C)
d
50 μM
e
f
~8.0 Å

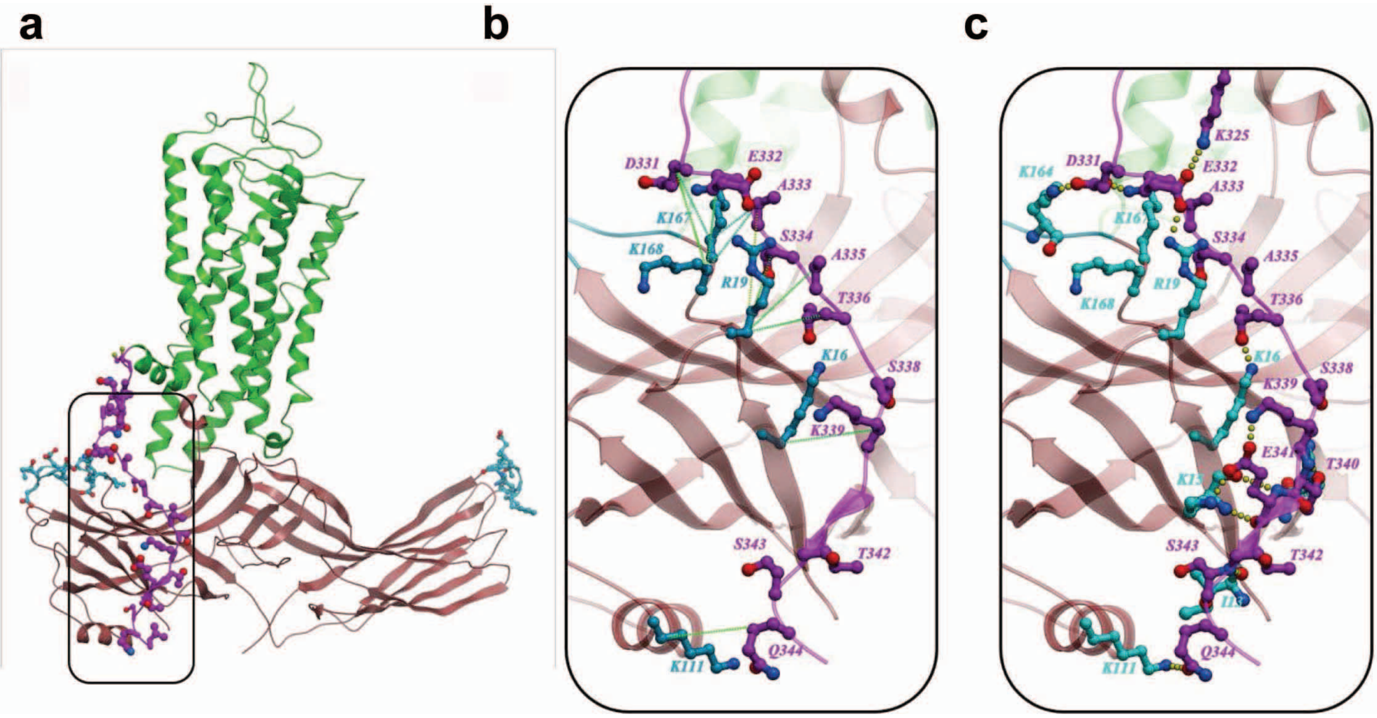
Extended Data Figure 2 | Purification and crystallization of T4L–rhodopsin–arrestin. **a**, Purification of the T4L–rhodopsin–arrestin (T4L–Rho–Arr) complex. $\mathrm{His_8}$–MBP–MBP–T4L–Rho–Arr complex was first purified by amylose column chromatography (lane 1). The $\mathrm{His_8}$–MBP–MBP tandem 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.



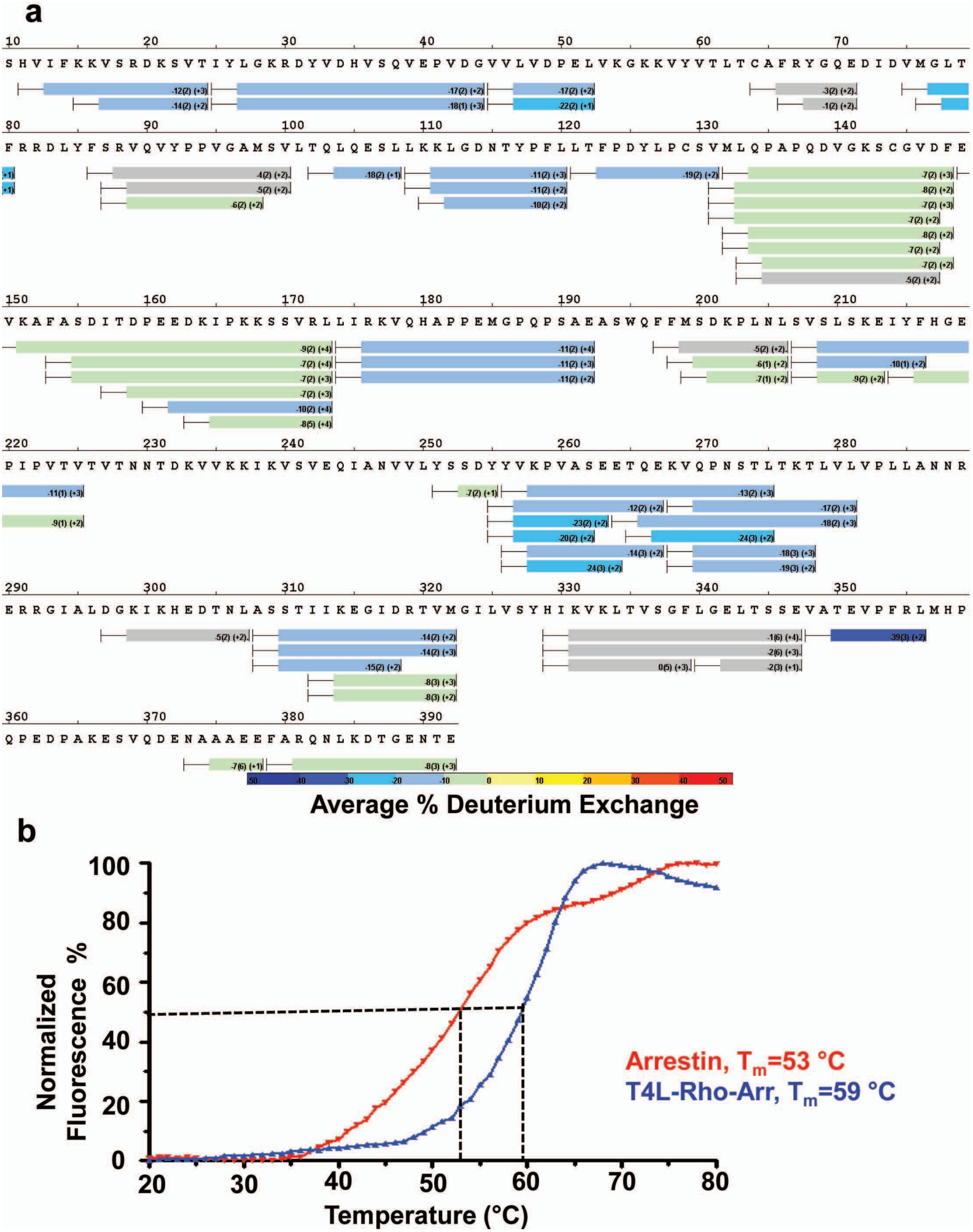
Extended Data Figure 3 | Electron density map for the overall complex and the key interfaces based on the XFEL data. **a**, A $2F_o - F_c$ electron density map contoured at 1σ of the arrestin finger loop, which forms the key interface with TM7 and helix 8. **b**, A $2F_o - F_c$ electron density map contoured at 1σ of the loop between TM5 and TM6, which forms the key interface with the β-strand following the finger loop. **c**, A 3,000 K simulated annealing omit map ($2F_o - F_c$ electron density map contoured at 1σ) 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 1σ calculated from the 3.8 Å/3.8 Å/3.3 Å truncated XFEL data. Key residues are labelled.


a
90°
Complex1
Complex2
Complex3
Complex4
N
C
b
P1
P2
90º
P4
P1
P2
P4
P3
c
P1
P2
P4
180º
P1
P2
P3

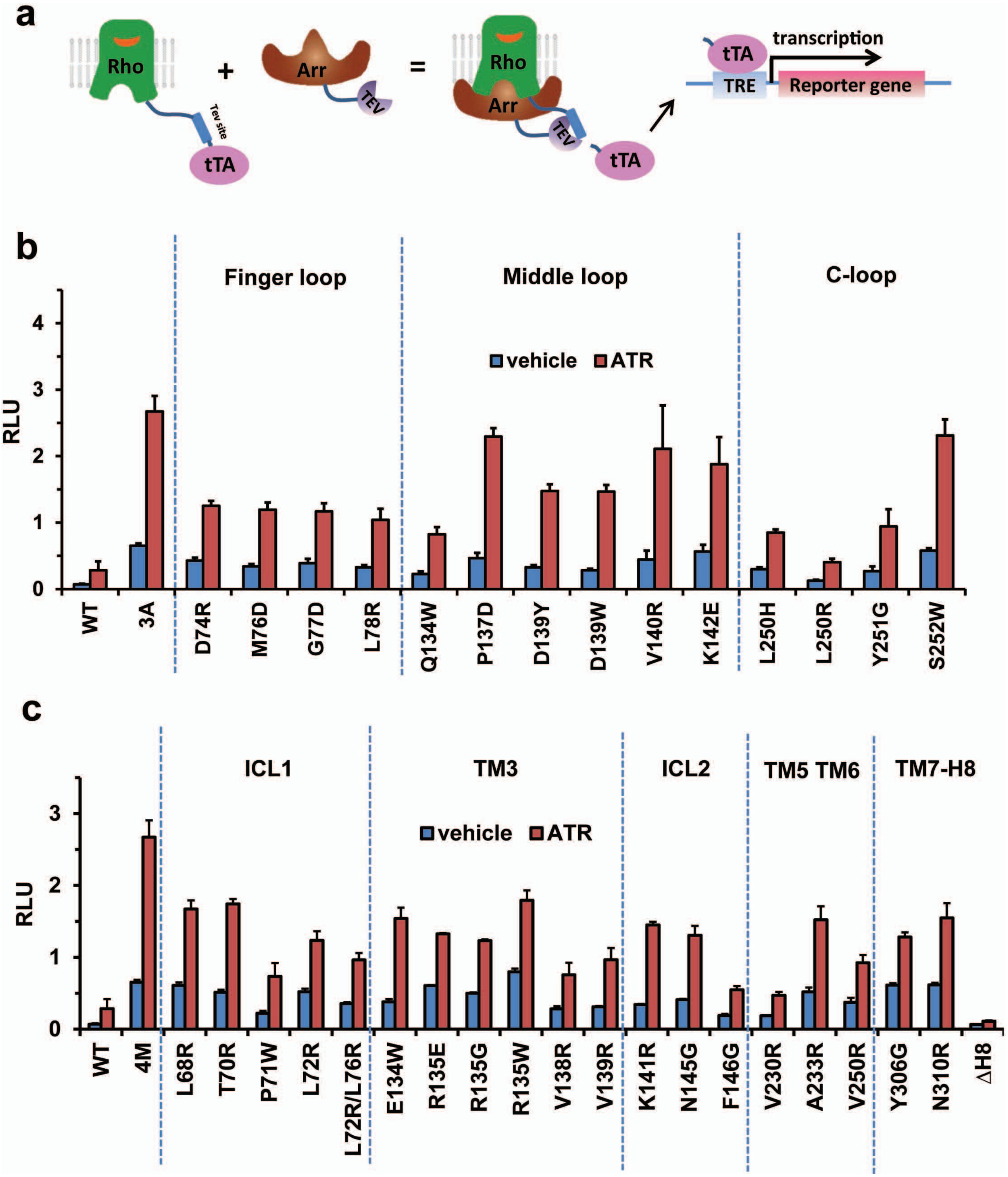
Extended Data Figure 4 | Structure similarity of the four rhodopsin–arrestin complexes in the asymmetric units and the interface between rhodopsin and arrestin. **a**, Two 90° views of the superposition of the four rhodopsin–arrestin complexes are shown as cartoon representation. The four complexes have 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.


a
b
D331
E332
A333
K167
K168
S334
R19
A335
T336
K16
S338
K339
S343
T342
K111
Q344
c
K325
K164
D331
E332
A333
K167
S334
R19
A335
K168
T336
K16
S338
K339
E341
T340
S343
T342
K111
Q344

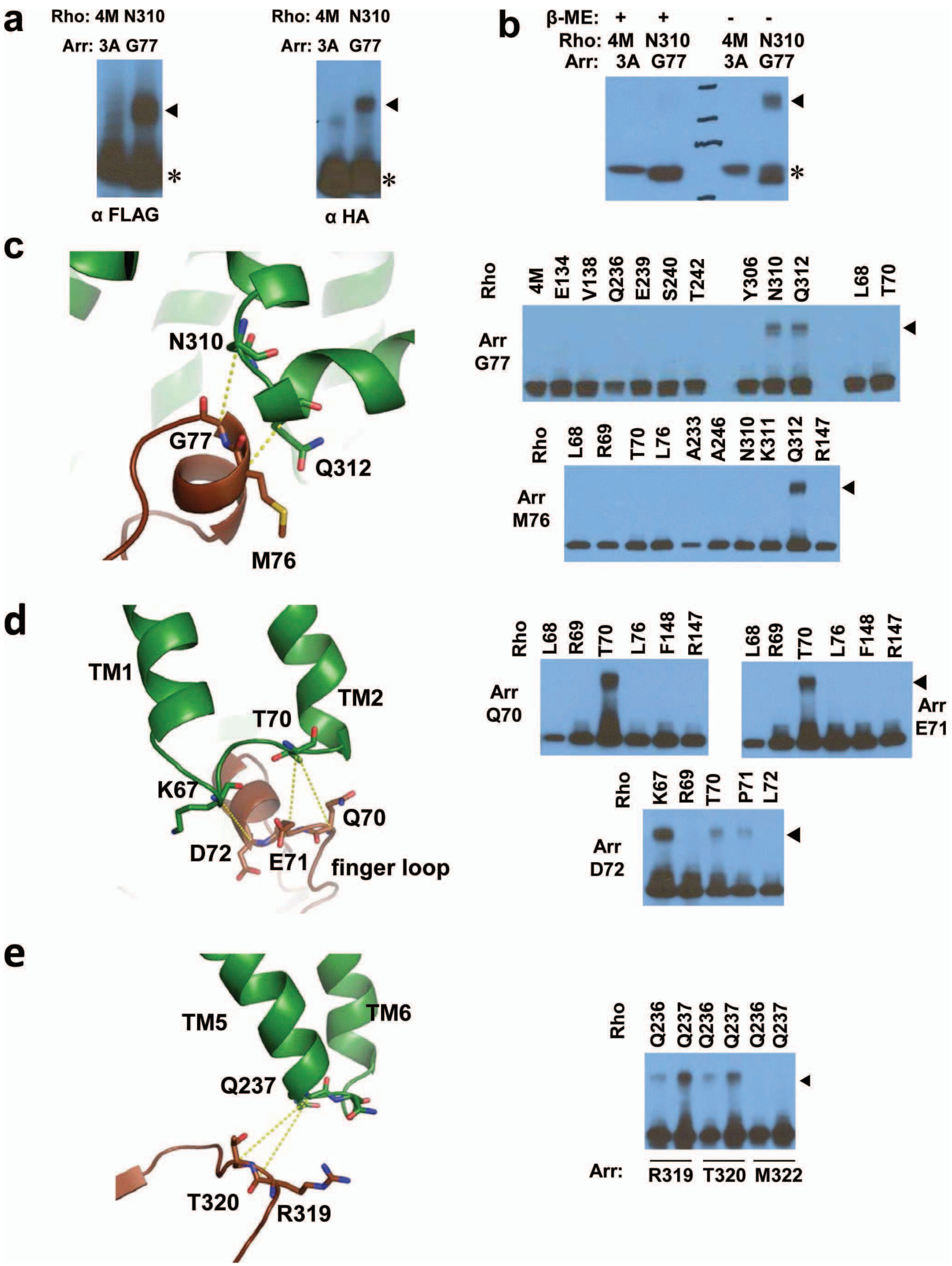
Extended Data Figure 5 | Conformational modelling of the rhodopsin–arrestin full length complex. **a**, An overview of the computational model. **b**, Predicted interactions of the rhodopsin C terminus with arrestin, showing strong to medium pairwise restraints between Cβ atoms of rhodopsin and arrestin 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
10 20 30 40 50 60 70
S H V I F K K V S R D K S V T I Y L G K R D Y V D H V S Q V E P V D G V V L V D P E L V K G K K V Y V T L T C A F R Y G Q E D I D V M G L T
-12(2) (+3) -17(2) (+2) -17(2) (+2) -3(2) (+2)
-14(2) (+2) -18(1) (+3) -22(2) (+1) -1(2) (+2)
80 90 100 110 120 130 140
F R R D L Y F S R V Q V Y P P V G A M S V L T Q L Q E S L L K K L G D N T Y P F L L T F P D Y L P C S V M L Q P A P Q D V G K S C G V D F E
+1) -4(2) (+2) -18(2) (+1) -11(2) (+3) -19(2) (+2) -7(2) (+3)
+1) -5(2) (+2) -11(2) (+2) -8(2) (+2)
-6(2) (+2) -10(2) (+2) -7(2) (+3)
-7(2) (+2)
-8(2) (+2)
-7(2) (+2)
-7(2) (+2)
-5(2) (+2)
150 160 170 180 190 200 210
V K A F A S D I T D P E E D K I P K K S S V R L L I R K V Q H A P P E M G P Q P S A E A S W Q F F M S D K P L N L S V S L S K E I Y F H G E
-9(2) (+4) -11(2) (+4) -5(2) (+2)
-7(2) (+4) -11(2) (+3) -6(1) (+2) -10(1) (+2)
-7(2) (+3) -11(2) (+2) -7(1) (+2) -9(2) (+2)
-7(2) (+3)
-10(2) (+4)
-8(5) (+4)
220 230 240 250 260 270 280
P I P V T V T V T N N T D K V V K K I K V S V E Q I A N V V L Y S S D Y Y V K P V A S E E T Q E K V Q P N S T L T K T L V L V P L L A N N R
-11(1) (+3) -7(2) (+1) -13(2) (+3)
-12(2) (+2) -17(2) (+3)
-9(1) (+2) -23(2) (+2) -18(2) (+3)
-20(2) (+2) -24(3) (+2)
-14(3) (+2) -18(3) (+3)
-24(3) (+2) -19(3) (+2)
290 300 310 320 330 340 350
E R R G I A L D G K I K H E D T N L A S S T I I K E G I D R T V M G I L V S Y H I K V K L T V S G F L G E L T S S E V A T E V P F R L M H P
-5(2) (+2) -14(2) (+2) -1(6) (+4) -39(3) (+2)
-14(2) (+3) -2(6) (+3)
-15(2) (+2) 0(5) (+3) -2(3) (+1)
-8(3) (+3)
-8(3) (+2)
360 370 380 390
Q P E D P A K E S V Q D E N A A A E E F A R Q N L K D T G E N T E
-7(6) (+1) -8(3) (+3)
-50 -40 -30 -20 -10 0 10 20 30 40 50
Average % Deuterium Exchange
b
Normalized Fluorescence %
100
80
60
40
20
0
20 30 40 50 60 70 80
Temperature (°C)
Arrestin, T$_m$=53 °C
T4L-Rho-Arr, T$_m$=59 °C

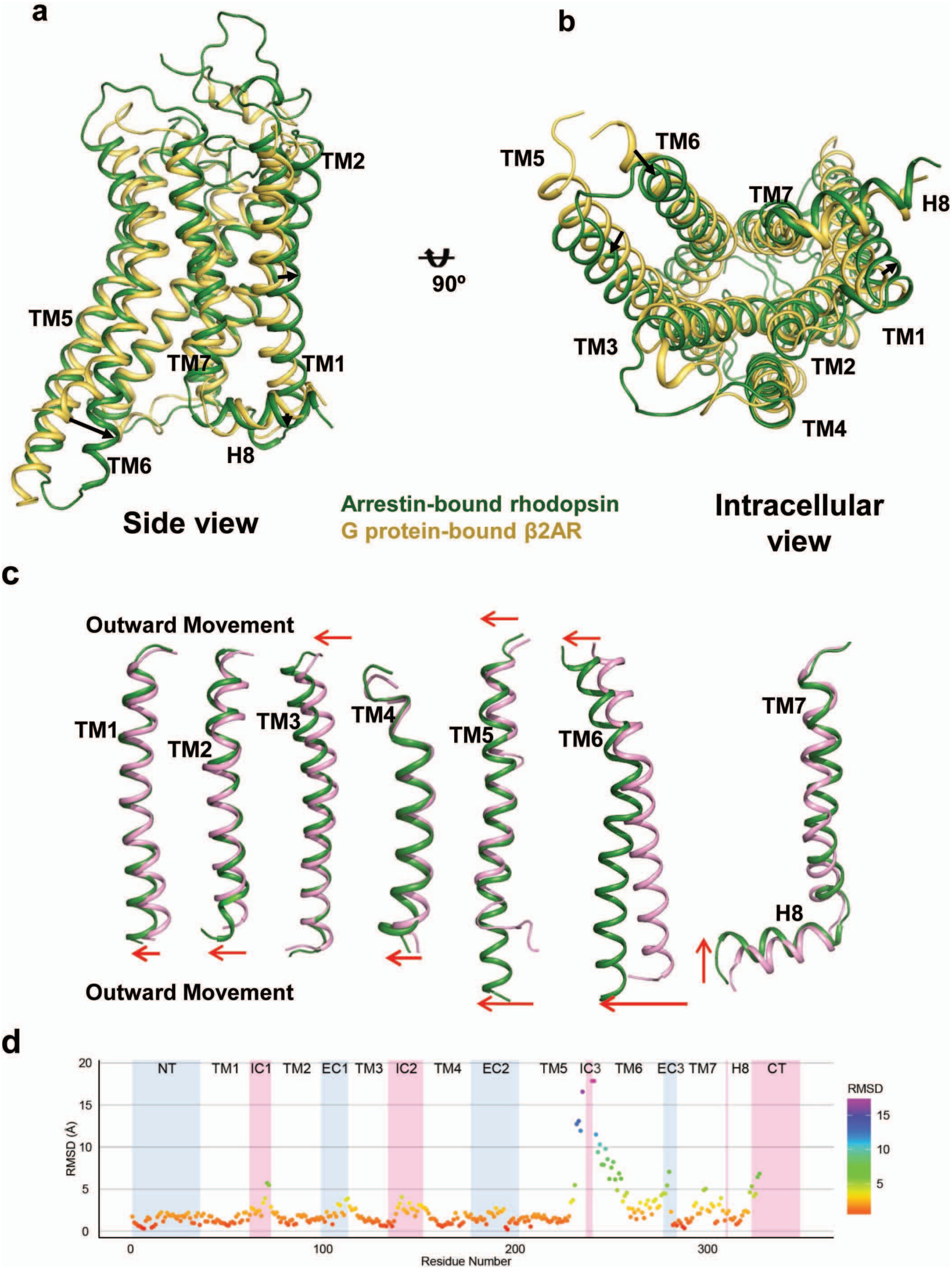
Extended Data Figure 6 | Dynamics of free 3A arrestin and rhodopsin-bound arrestin determined by HDX. a, HDX perturbation map between rhodopsin-bound arrestin and free arrestin, which is derived from the difference in the HDX rate between rhodopsin-bound arrestin and free arrestin. The bars below the 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.



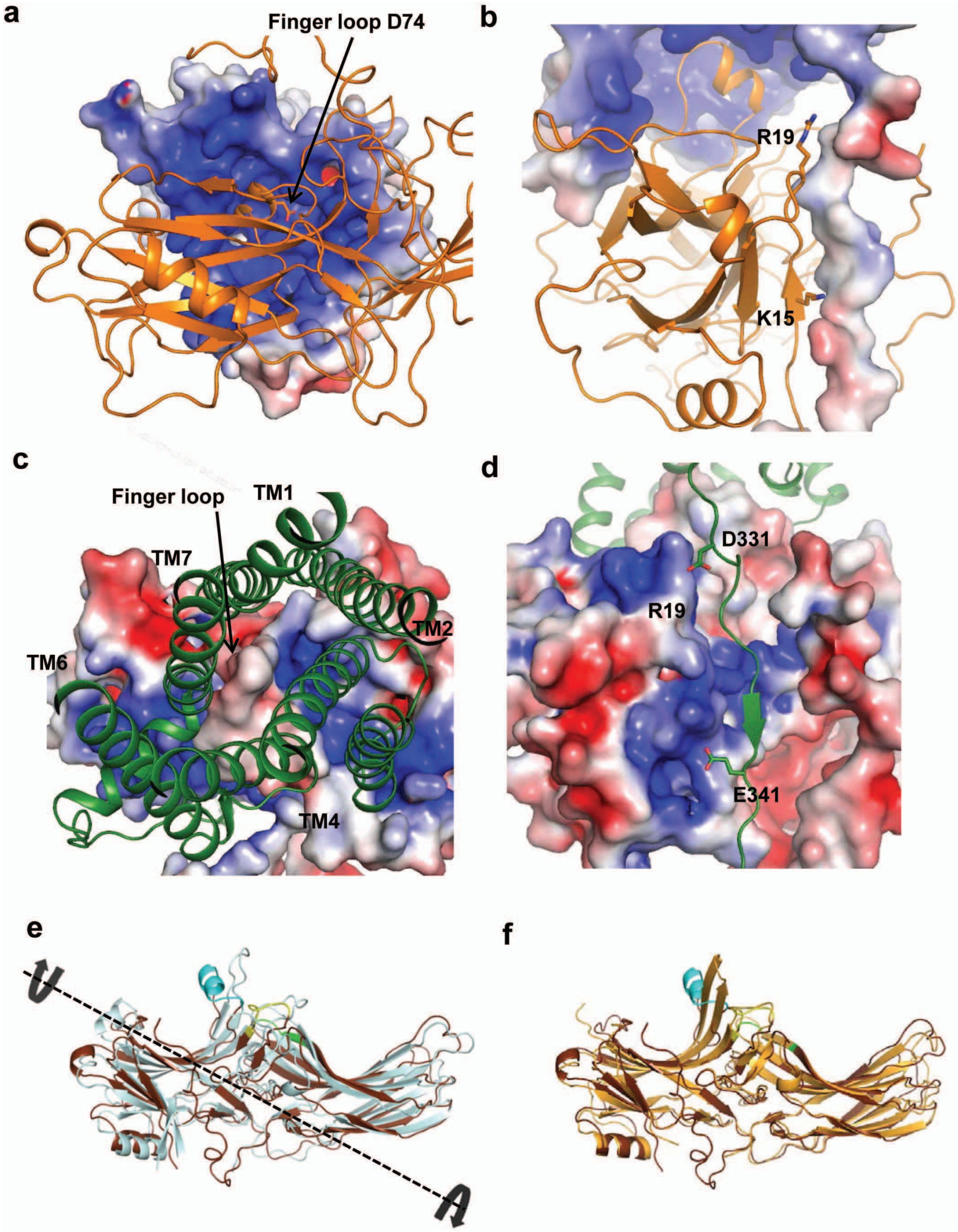
Extended Data Figure 7 | Cell-based Tango assays to validate the rhodopsin–arrestin interface. a, Cartoon illustration of the Tango assay for rhodopsin–arrestin interactions in cells. **b**, **c**, Mutations of key arrestin (**b**) and rhodopsin (**c**) residues that mediate the rhodopsin–arrestin interactions. Tango assay were performed in the absence or presence of 10 µM all-*trans*-retinal (ATR). ($n = 3$, error bars, s.d.).


a
Rho: 4M N310
Arr: 3A G77
*
α FLAG
Rho: 4M N310
Arr: 3A G77
*
α HA
b
β-ME: + + - -
Rho: 4M N310 4M N310
Arr: 3A G77 3A G77
*
c
N310
G77
Q312
M76
Rho 4M E134 V138 Q236 E239 S240 T242 Y306 N310 Q312 L68 T70
Arr G77
Rho L68 R69 T70 L76 A233 A246 N310 K311 Q312 R147
Arr M76
d
TM1
TM2
T70
K67
Q70
D72
E71
finger loop
Rho L68 R69 T70 L76 F148 R147
Arr Q70
L68 R69 T70 L76 F148 R147
Arr E71
Rho K67 R69 T70 P71 L72
Arr D72
e
TM5
TM6
Q237
T320
R319
Rho Q236 Q237 Q236 Q237 Q236 Q237
Arr: R319 T320 M322

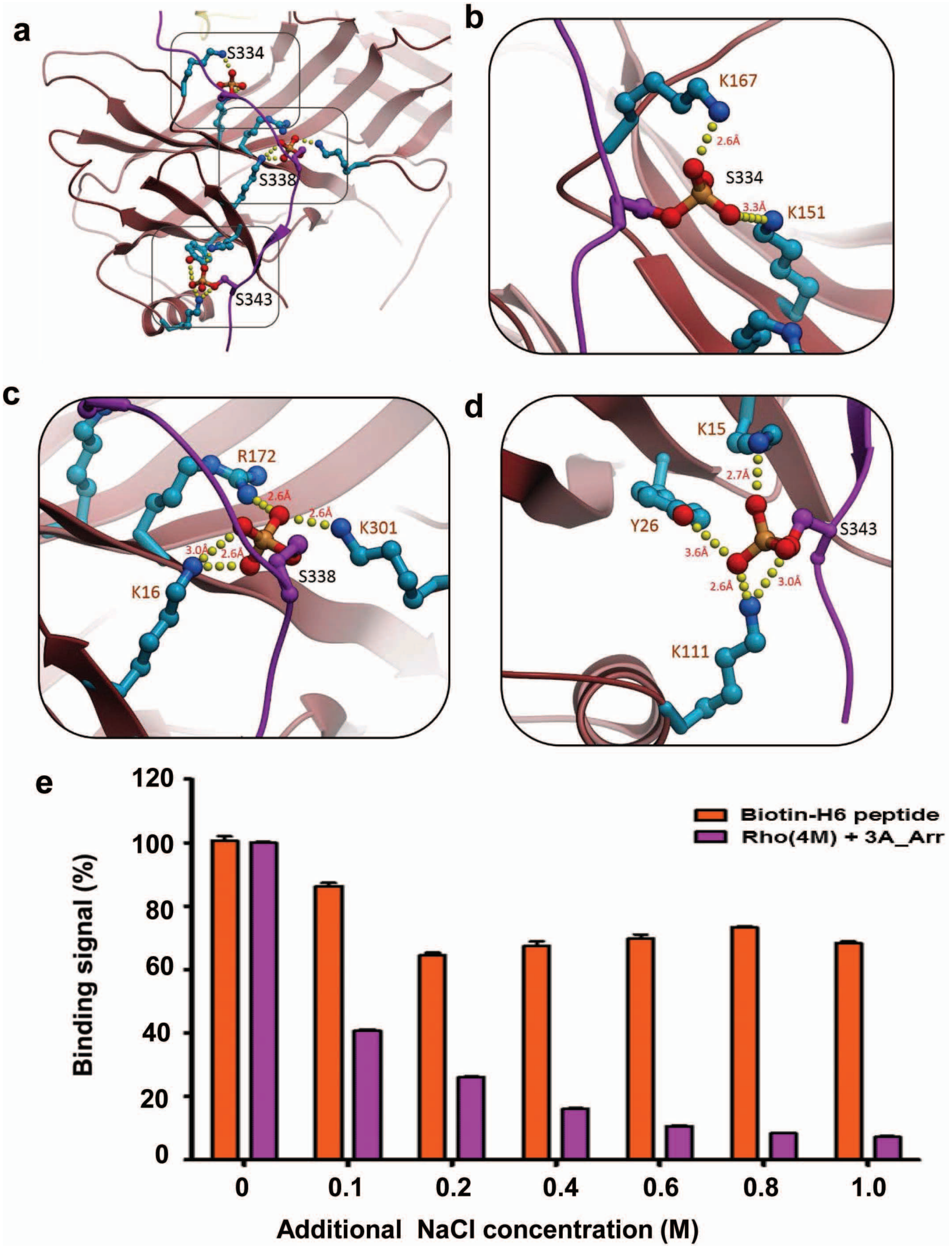
Extended Data Figure 8 | Control experiments for disulfide bond cross-linking specificity. **a**, The product of the cross-linking reaction of finger loop residue G77C with N310$^{7.57}$C of TM7 was confirmed by western blots using anti-Flag antibody (which detects arrestin–Flag fusion) and anti-HA antibody (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$^{7.57}$C 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$^{7.57}$C of TM7 and Q312$^{8.49}$ of helix 8, and M76C was cross-linked to N310$^{7.57}$C of TM7 and Q312$^{8.49}$C 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^{ICL3}C from TM5 of rhodopsin.


a
TM2
TM5
TM7
TM1
TM6
H8
Side view
90º
b
TM5
TM6
TM7
H8
TM3
TM1
TM2
TM4
Intracellular view
Arrestin-bound rhodopsin
G protein-bound β2AR
c
Outward Movement
TM1
TM2
TM3
TM4
TM5
TM6
TM7
H8
Outward Movement
d
NT
TM1
IC1
TM2
EC1
TM3
IC2
TM4
EC2
TM5
IC3
TM6
EC3
TM7
H8
CT
RMSD (Å)
20
15
10
5
0
0
100
200
300
Residue Number
RMSD
15
10
5

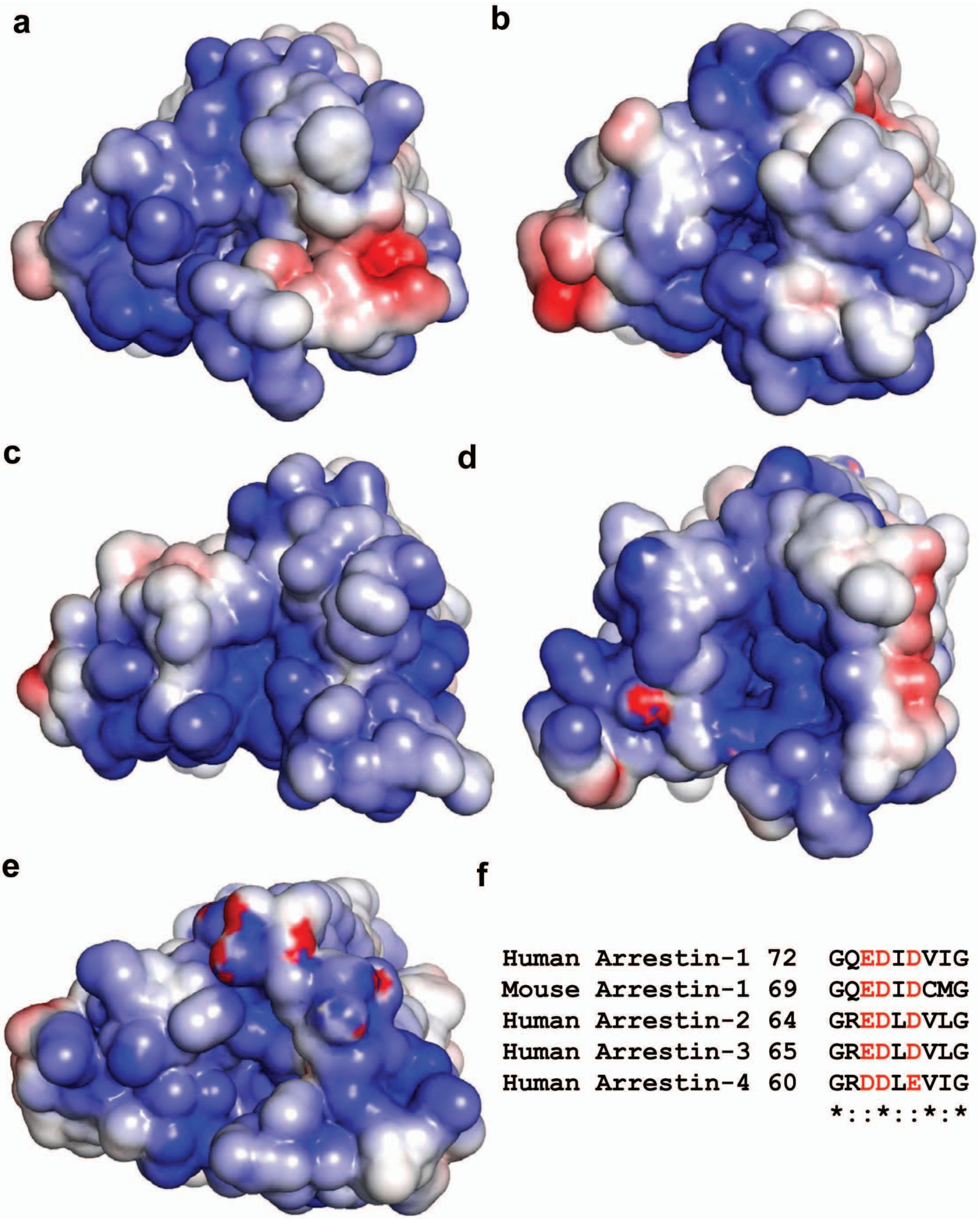
Extended Data Figure 9 | Structure comparison of the arrestin-bound rhodopsin with the β_2-adrenergic receptor in complex with G_s protein (PDB code 3SN6) and the inactive rhodopsin (PDB code 1F88). **a**, Superposition of arrestin-bound rhodopsin (green) with G_s protein-bound β_2 adrenergic receptor (light yellow). The major conformational changes are indicated by arrows. **b**, An intracellular view of a superposition of arrestin-bound rhodopsin (green) and G_s protein-bound β_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.


a
Finger loop D74
b
R19
K15
c
Finger loop
TM1
TM7
TM2
TM6
TM4
d
D331
R19
E341
e
f

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**.

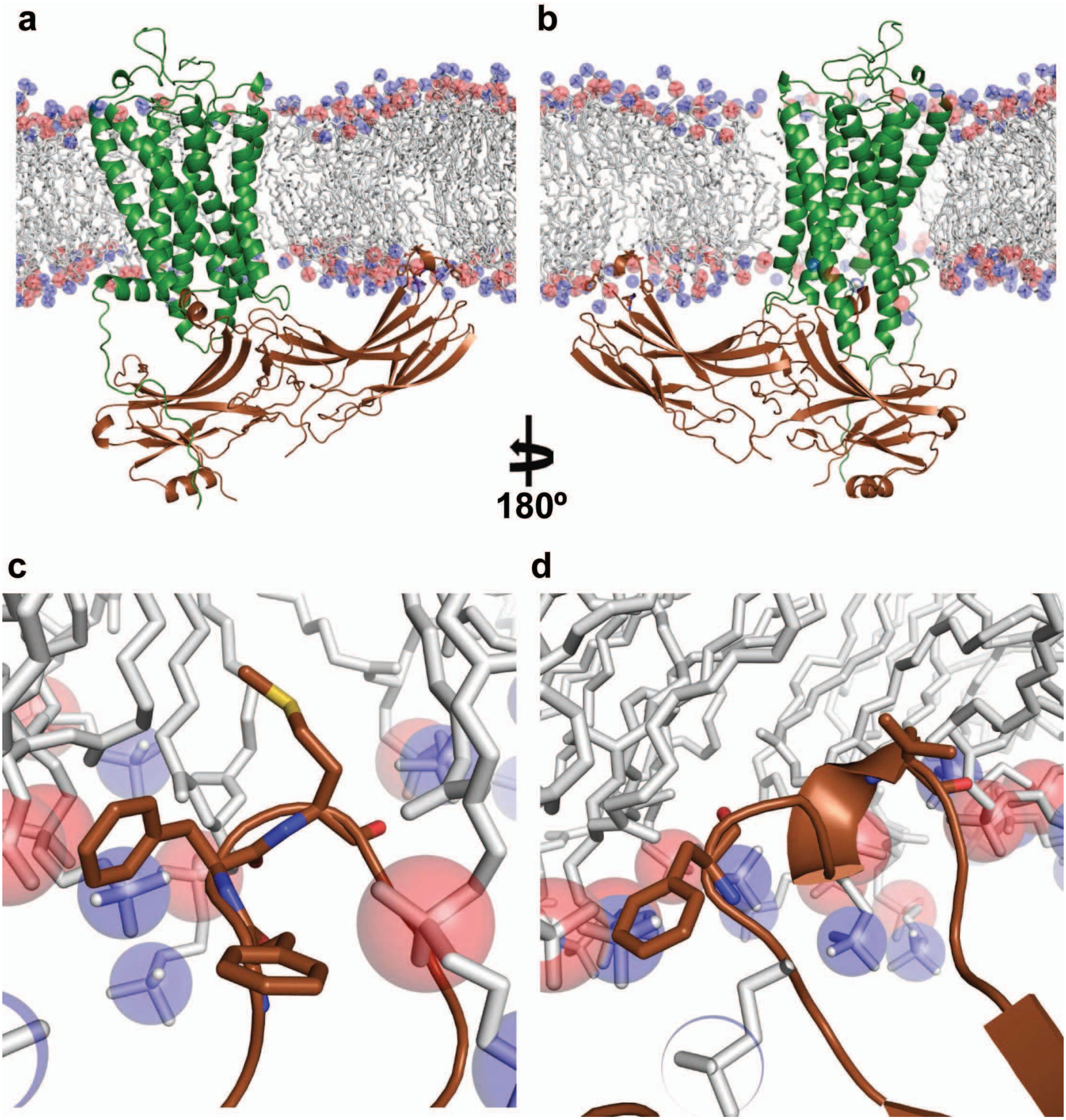

a
S334
S338
S343
b
K167
2.6Å
S334
3.3Å
K151
c
R172
2.6Å
2.6Å
K301
3.0Å
2.6Å
S338
K16
d
K15
2.7Å
Y26
S343
3.6Å
2.6Å
3.0Å
K111
e
Binding signal (%)
0
20
40
60
80
100
120
Biotin-H6 peptide
Rho(4M) + 3A_Arr
0
0.1
0.2
0.4
0.6
0.8
1.0
Additional NaCl concentration (M)

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–His$_6$) shows much less salt sensitivity than the interaction between His-tag–rhodopsin and biotin arrestin, which is very sensitive to salt, with an IC$_{50}$ of around 200 mM NaCl (100 mM NaCl added to 100 mM salt of the original assay buffer) (*n* = 3, error bars, s.d.).



```
Human Arrestin-1 72   GQEDIDVIG
Mouse Arrestin-1 69   GQEDIDCMG
Human Arrestin-2 64   GREDLDVLG
Human Arrestin-3 65   GREDLDVLG
Human Arrestin-4 60   GRDDLEVIG
                      *::*::*:*
```

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: β_1AR, PDB code 2Y02 (**a**); β_2AR, PDB code 3PDS (**b**); A_{2A} adenosine receptor, PDB code 3QAK (**c**); serotonin receptor $5HT_{1B}$, PDB code 4IAR (**d**); serotonin receptor $5HT_{2B}$, 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.



Extended Data Figure 13 | 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.