Cryo-EM structure of a single-chain β1-adrenoceptor – AmpC β-lactamase fusion protein

Gabriella Collu[1,2], Inayathulla Mohammed[3], Aleix Lafita[4], Tobias Bierig[1,2], Emiliya Poghosyan[1], Spencer Bliven[5], Pavel Afanasyev[6], and Roger M. Benoit[1*]

[1]Laboratory for Multiscale Bioimaging, PSI Center for Life Sciences, 5232 Villigen PSI, Switzerland

[2]Department of Biology, ETH Zurich, 8093 Zurich, Switzerland

[3]BioEM, Biozentrum, University of Basel, Basel, Switzerland

[4]Laboratory of Biomolecular Research, Division of Biology and Chemistry, Paul Scherrer Institute, 5232 Villigen PSI Switzerland

[5]PSI Center for Scientific Computing, Theory and Data, 5232 Villigen PSI, Switzerland

[6]Cryo-EM Knowledge Hub, ETH Zurich, Zurich, Switzerland

*Correspondence to:
Roger M. Benoit
Paul Scherrer Institute PSI
PSI Center for Life Sciences
Laboratory for Multiscale Bioimaging
Switzerland
E-mail: roger.benoit@psi.ch

Abstract

The insertion of fusion proteins has enabled the crystallization of a wide range of G-protein-coupled receptors. Here, we adapted this engineering strategy to cryo-electron microscopy (cryo-EM). We inserted the soluble protein AmpC β-lactamase into the third intracellular loop (ICL3) of ultra-thermostable β1-adrenoceptor (β1AR) via chimeric helix fusions. Biochemical and biophysical characterization showed that the resulting fusion protein after expression, solubilization and purification was monodisperse and able to bind the known β1AR weak partial agonist cyanopindolol, and the antagonist propranolol. The protein particles comprised sufficient mass and discernable structural features to elucidate its cryo-EM structure in complex with cyanopindolol without any natural (G-proteins, arrestins) or artificial (Nanobodies, DARPins) binding partners, to an overall resolution of 4.2 Å. The seven-helix architecture and helix eight, as well as both GPCR - AmpC β-lactamase connections are clearly resolved. β1AR is in its inactive conformation. 3D variability analysis revealed significant flexibility between the two protein domains and within the GPCR helices, offering insights into conformational dynamics. The map contains clear density for the cyanopindolol. The fusion protein geometry theoretically fits a wide range of class A GPCRs, presenting a powerful platform for structure elucidation of a diverse array of class A GPCR – ligand complexes by cryo-EM in the inactive receptor state. The approach furthermore holds potential for structure elucidation of GPCRs in the absence of ligands.

Introduction

G-protein-coupled receptors (GPCRs) interact with extracellular signaling molecules and transduce the information across the cell membrane through conformational changes within the GPCRs, resulting in various intracellular responses.[1] Non-rhodopsin GPCRs were for a long time considered among the most challenging protein drug targets for structural studies. The first structures of class A GPCRs were published in 2007[2,3] and in 2008[4]. Major bottlenecks for the structural biology studies of GPCRs include protein production, for example due to low expression levels or low protein stability, and crystallization, which can be challenging due to high conformational flexibility or a limited area of accessible hydrophilic surface.

Three important strategies that revolutionized GPCR structural studies were the introduction of point mutations to improve the thermostability of the receptors,[4,5] the binding of specific antigen-binding fragments (Fab),[3] and the insertion of fusion proteins to stabilize a specific receptor conformation and/or to increase the soluble surface area for promoting the formation of ordered crystal contacts.[2]

More recently, cryo-electron microscopy (cryo-EM) has emerged as a powerful alternative to X-ray crystallography and represents the next major progress in the field of GPCR structural biology. Protein structure elucidation at near-atomic resolution by single particle cryo-EM has seen an impressive increase in the number of structures published through the "resolution revolution", enabled by important technological advances in specimen preparation, as well as the development of direct electron detectors, improved microscopes and improvements in data analysis.[6] The major advantages of using cryo-EM for structural studies are that only microgram quantities of protein are needed, crystallization is not required, and different conformational states can be revealed. Nonetheless, due to their relatively small molecular mass (~45 kDa), combined with the relatively featureless shape within the detergent micelle, GPCRs remain challenging targets for single particles analysis, owing to the low signal-to-noise ratio and to the absence of prominent structural features that would be needed to achieve high accuracy in the determination of the angular orientation for 3D reconstructions, as well as heterogeneity of the protein complexes.[7,8] For small proteins such as GPCRs, a variety of factors are important to obtain optimal results in cryo-EM.[9] In 2017, the first GPCR structures were solved by cryo-EM, both belonging to class B: the calcitonin receptor (CTR)[10] and the glucagon-like peptide-1 receptor (GLP-1)[11]. The structures of these receptors were solved in complex with the heterotrimeric G-proteins, in the Zhang et al.[11] structure in addition with a specific nanobody. The G-protein complexes had sufficient molecular mass for obtaining a

good signal-to-noise ratio in single-particle cryo-EM data collection and sufficiently clearly identifiable structural features for particle alignment. The complex formation relies on the binding of an agonist molecule that activates the receptor, therefore leading to the elucidation of active conformation structures with G-protein or arrestin coupling (e.g. Refs[10-13]). However, this approach is limited by the necessity of agonist-induced conformational changes that facilitate the binding of these partners, making it challenging to obtain high-resolution structures in the absence of such ligands. Therefore, to date, cryo-EM structures of GPCRs in their inactive conformation and apo structures are underrepresented.

To address this challenge, we optimized a molecular engineering strategy widely used in X-ray crystallography to apply it for cryo-EM studies. By searching the Protein Data Bank[14] (https://www.rcsb.org/) using a BioJava script, we identified AmpC β-lactamase as a relatively large protein (39 kDa) that is well suited for insertion into intracellular loop 3 (ICL3) of class A GPCRs via genetically extended helix fusions.

We designed and cloned single-chain β1AR - AmpC β-lactamase fusion proteins.[15] Here, we present the cryo-EM structure of a β1AR - AmpC β-lactamase fusion protein in complex with the weak partial agonist cyanopindolol at 4.2 Å resolution.

Results

Construct design

To study the structure of cyanopindolol-bound ultra-thermostable β1-adrenergic receptor by cryo-EM, we engineered a single-chain GPCR-fusion protein of sufficient mass and rigidity to allow cryo-EM structural analysis. To identify a suited fusion protein that adds enough mass to the GPCR and that can be connected to the GPCR at ICL3 via rigid, extended α-helices, structures in the PDB were filtered for proteins fulfilling the following criteria: (1) the protein chains are composed of more than 300 residues, (2) N- and C- termini are α-helical, (3) the terminal helices are approximately antiparallel to each other, and (4) the terminal helices are close to 11 Å apart. These constraints were chosen to enable the protein to extend TM5 and TM6 in the ICL3 of class A GPCRs with minimal disruption. Structures from the resulting list of hits were manually inspected, following additional selection criteria: membrane proteins and oligomeric proteins were excluded, and proteins with available high-resolution structures were prioritized.

The AmpC β-lactamase[16] (PDB:1FCO) from *E. coli* fulfilled all the desired criteria and we therefore used it for insertion into β1AR. The N- and C-terminal helices of β-lactamase were directly fused to the transmembrane helices TM5 and TM6 of β1AR.

Nine different constructs were designed, in which the transition between TM5 of the receptor and the N-terminus of the fusion protein, and the connection between the C-terminus of the fusion protein to TM6 of the receptor, were optimized. The initial construct, which we named construct 0/0, was designed based on visual inspection of the PDB structures of a stabilized turkey β1-adrenergic receptor[4] (PDB:2VT4) and of the fusion protein[16] (PDB:1FCO)(Figure 1).

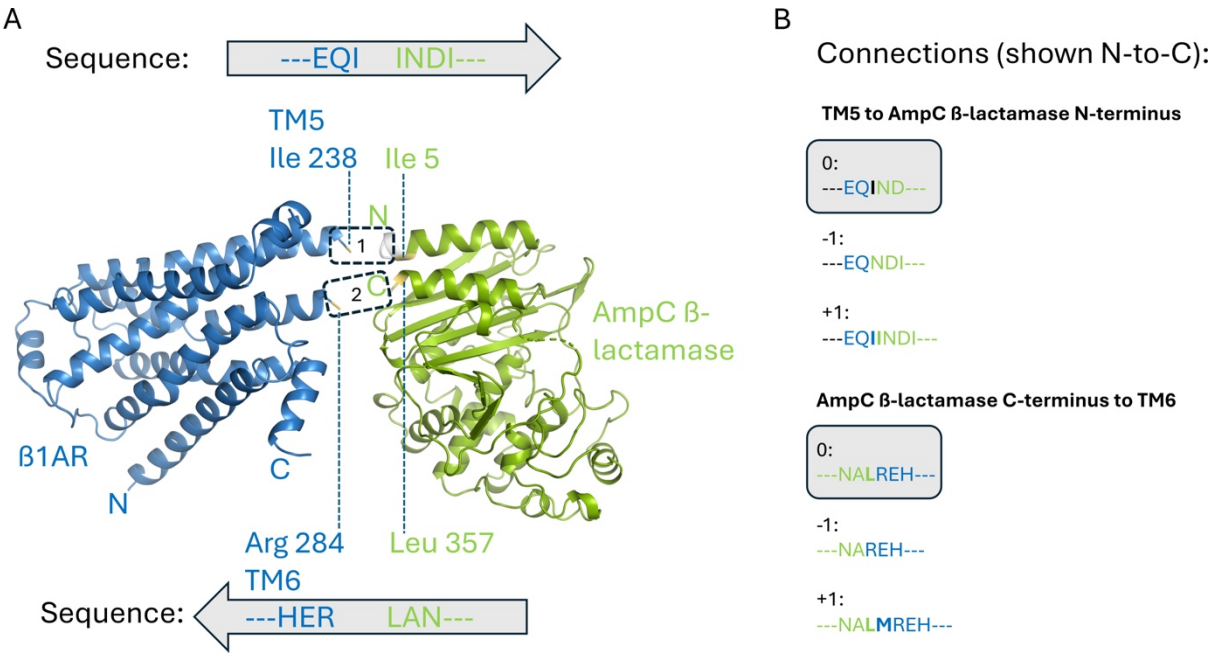

Figure 1) Construct design. A) A structure of a stabilized turkey β1-adrenergic receptor[4] (blue; PDB:2VT4) and a structure of AmpC β-lactamase[16] (green; PDB:1FCO) were manually aligned to design two chimeric helices that connect the proteins. Chimeric helix 1 connects the C-terminal end of the β1AR transmembrane helix 5 to the N-terminal helix of AmpC β-lactamase (dashed box 1, left-to-right). Chimeric helix 2 connects the C-terminal helix of AmpC β-lactamase to the N-terminal end of β1AR transmembrane helix 6 (dashed box 2, right-to-left). The N- and C-termini of β1AR and AmpC β-lactamase are indicated. **B)** Based on this alignment, initial connections were designed (construct 0/0). For the additional constructs, each of the two connections was either kept at 0 or lengthened or shortened by one amino acid, using only original sequences from β1AR and AmpC β-lactamase, in all possible combinations, resulting in 9 constructs with slightly different linker combinations.

The remaining eight constructs were generated by adding or removing one codon at each of the receptor-fusion protein transitions.[15] The sequence of our cryo-EM construct was based on the ultra-thermostable mutant β1AR-JM50.[17,18]

The constructs, in a pcDNA 4/TO mammalian expression vector, included a C-terminal flexible linker comprising a human rhinovirus 3C protease (HRV 3C) cleavage site, followed by an eGFP for detection, and a twin Strep-tag for purification.

Protein expression and purification

The GPCR fusion proteins were produced in stably transfected[19] HEK293S GnTI(-)[20] cells. For expression tests, adherent cell cultures were used; for protein production, the cells were grown as suspension cultures.

To extract the protein of interest, membranes were solubilized using n-Dodecyl-β-D-Maltoside (β-DDM). Purification was performed by Strep-Tactin affinity chromatography followed by size-exclusion chromatography (SEC).

The purification of the proteins for cryo-EM analysis was performed in the presence of the high-affinity β1AR-binder cyanopindolol. The purification of proteins for biophysical characterization was performed in the absence of ligands. For cryo-EM analysis and for the biophysical characterization of the constructs, the C-terminal eGFP, which was fused to the protein of interest via a flexible linker, was removed by cleavage with HRV 3C protease (Figure 2 A and B).

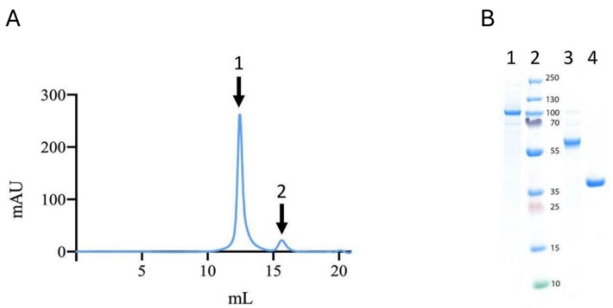

Figure 2) Protein purification. A) Size-exclusion chromatography profile of the β1AR-AmpC β-lactamase fusion protein construct 0/0 (peak 1), with the C-terminal eGFP (peak 2) cleaved off using HRV-3C protease. **B)** SDS-PAGE analysis: (1) uncleaved β1AR-AmpC β-lactamase construct with eGFP, (2) PageRuler Plus Prestained Protein Ladder (Thermo Scientific), (3) cleaved β1AR - AmpC β-lactamase construct, and (4), cleaved-off eGFP.

Ligand-binding assays

N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM) based fluorescent thermal stability[21] assays were performed on the three most diverse constructs, named -1/-1, 0/0 and +1/+1, to assess the stability and correct folding of the receptor in the fusion protein. The thermostability of the proteins was measured upon addition of increasing concentrations of the weak partial β1AR agonist cyanopindolol, or the β1AR antagonist propranolol. The addition of the small molecule cyanopindolol resulted in a strong stabilizing effect on all tested

constructs, while the addition of the small molecule propranolol had a less strong, but still significant stabilizing effect (data not shown). We proceeded with the 0/0 construct for structural studies.

Structure of the β1AR – AmpC β-lactamase fusion protein

Cryo-electron microscopy

We solved the cryo-EM structure of the β1AR fusion protein to an overall resolution of about 4.2 Å. The cryo-EM map revealed the connectivity of all eight helices of the receptor. In contrast to the crystal structure,[18] where a single conformation is embedded in the crystal lattice, the cryo-EM data reveals conformational heterogeneity within the receptor and between the receptor and the fusion protein. 3D variability analysis shows movements between β1AR and AmpC β-lactamase, and variability in the positions of the α-helices of β1AR with respect to each other (Supplementary Movie 1), revealing a conformational flexibility that is typical for GPCRs. The density for the ligand cyanopindolol, which binds deeply in the center of the extracellular region of the receptor, is defined in the cryo-EM map and allows determination of the overall ligand binding mode. The sharpened map of the overall structure, the real-space refined model and a close-up view of the ligand-binding site are shown in Figure 3.

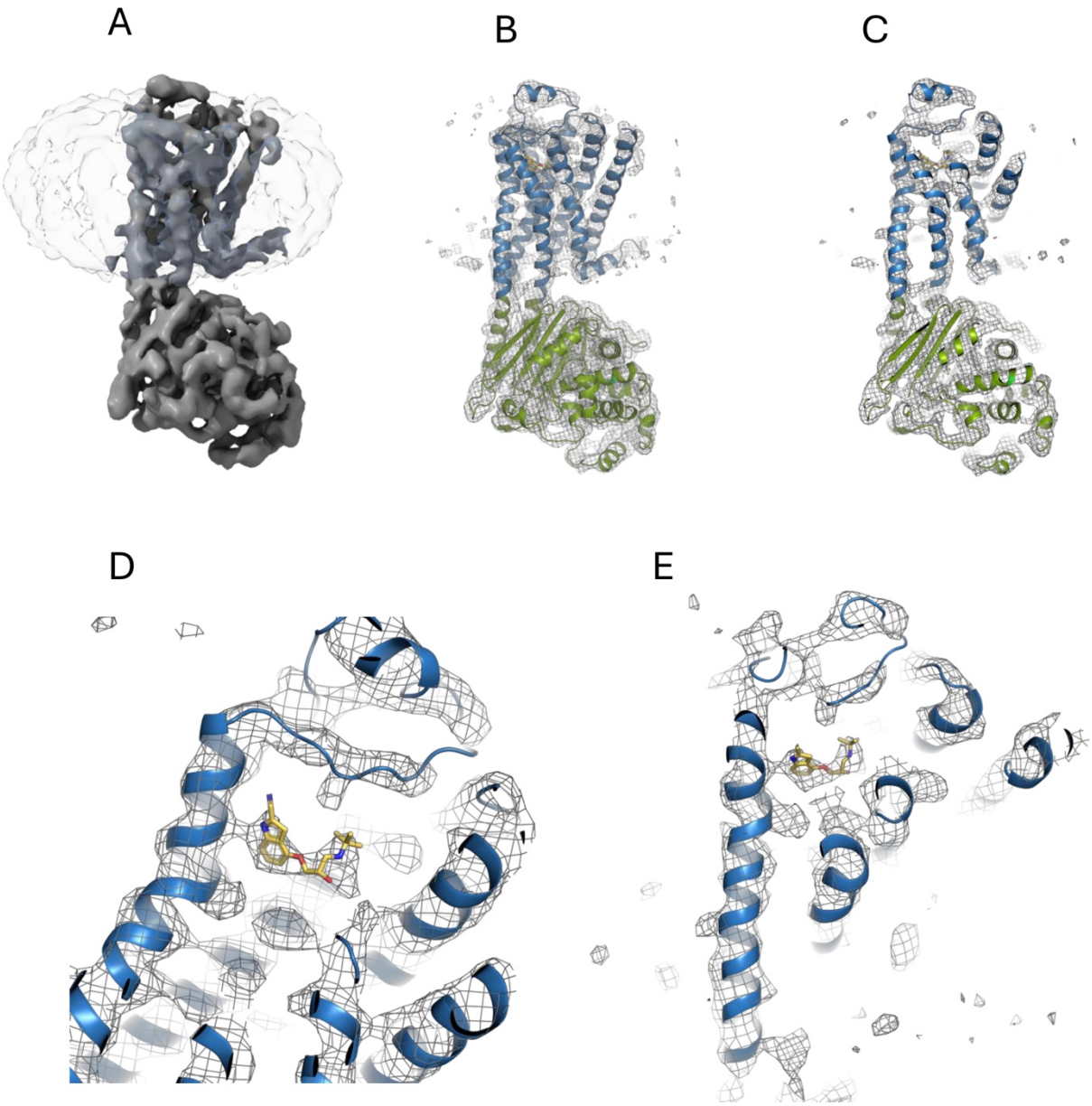

Figure 3) Overall structure and ligand-binding site. A) Cryo-EM map of the β1AR fusion protein at 4.2 Å resolution. **B, C)** Real-space refined model (cartoon representation) of the β1-adrenoceptor – AmpC β-lactamase fusion protein fitted in the cryo-EM map, overall (B) and slab (C) view. **D and E)** Close-up slab views of the ligand binding site. In B-E, the β1-adrenoceptor is shown in blue, AmpC β-lactamase is shown in green, and the ligand cyanopindolol is shown in yellow and atom colors (blue = nitrogen, red = oxygen). The Figure was created using ChimeraX[22] and Pymol[23].

The resolution was limited by a strong conformational heterogeneity (Supplementary Movie 1) and the number of particles. Locally, particularly at the ligand binding site, the resolution and the quality of the map are significantly better than on the periphery (Supplementary Figure 1, see local resolution), which allowed us to build the atomic model especially at the ligand

binding area with strong confidence. The quality of the model was validated with the existing X-ray data.

Helix 8 variability

In a physiological environment, helix 8 of adrenergic receptors is anchored to the cytoplasmic face of the cell membrane through palmitoyl groups.[24] In our construct, the palmitoylation site (C358A) was removed by mutagenesis, and the protein was solubilized in detergent, hence no membrane was present. Despite the lack of membrane anchoring or crystal contacts, helix 8 is defined in the cryo-EM map from Pro 347 up to Arg 361. Due to flexibility, the density is difficult to interpret after Lys 354. Helix 8 adopts a typical position slanted upwards at a shallow angle towards the expected location of a cell membrane. It is likely that the conformation of helix 8 is stabilized by the micelle.

Model building and interpretation

An initial model was generated based on an AlphaFold2[25] model of the fusion protein, and high-resolution crystal structures of both protein components. The model was fitted into the cryo-EM map and refined in iterative cycles of model building and refinement.

Ligand binding site

The cyanopindolol binding site was analyzed and validated by comparison to the crystal structure (PDB:4BVN)[18]. In our model, Asp $121^{3.32}$, Ser $211^{5.42}$ and Tyr $333^{7.53}$ are in hydrogen bonding distance to the ligand, in agreement with the crystal structure. Superscripts indicate Ballesteros-Weinstein nomenclature.[26] Asn $329^{7.39}$, which makes a hydrogen bond with the ligand in the crystal structure, is slightly more distant from the ligand N2 in the cryo-EM model (3.7 vs. 2.8 Å). Asp $121^{3.32}$, Asn $310^{6.55}$ and Tyr$333^{7.53}$ are reasonably well defined in the cryo-EM map, the other hydrogen-bonding residues are located as expected from the overall C_{α}-positions. Overall, we could typically build aromatic residues with strong confidence.

The ligand is furthermore lined by residues that make non-polar contacts: Trp 117, Thr 118, Val 122, Val 125, Thr 126, Phe 201, Thr 203, Tyr 207, Ala 208, Ser 212, Ser 215, Trp 303, Phe 306, Phe 307, and Phe 325. Thr 117, Thr 118, Val 125, and Trp 303 are defined in the cryo-EM map.

Activation state

Superposition of our cryo-EM structure with the crystal structure of the β1AR-JM50 construct[18] (PDB:4BVN) shows that, based on the overall conformation of the transmembrane

helices, the receptor adopts an inactive (R) conformation, in agreement with the X-ray crystal structure of the same receptor without a fusion protein (Figure 4).

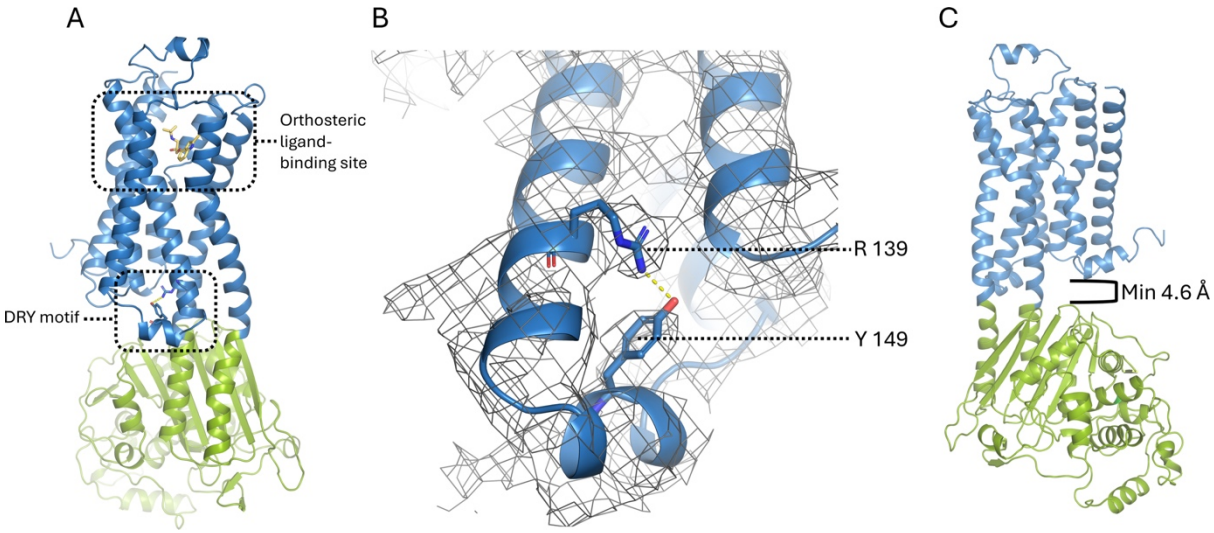

Figure 4) DRY motif and fusion protein geometry. The fusion protein is shown in blue (β1AR) and green (AmpC β-lactamase). **A)** Orthosteric ligand binding site and the DRY motif shown in the context of the overall structure. **B)** Close-up view of the DRY motif. The cryo-EM density map is shown as a mesh. **C)** The minimum distance between the two proteins (except at the covalently linked chimeric helices) is ~4.6 Å, therefore, steric hindrance that could interfere with inter-helix conformational changes (except between the fused TM5 and 6) is unlikely. The Figure was created using Pymol[23].

The DRY motif (Figure 4 A and B) in TM α-helix 3 is composed of Asp $138^{3.49}$, Arg $139^{3.50}$ and Tyr $140^{3.51}$, and has previously been shown to interact with Tyr149 from the short α-helix in ICL2[4] in the inactive state. Arg $139^{3.50}$ and Tyr 149 are defined in the cryo-EM map and are in agreement with this interaction (Figure 4 B).

The minimal distance between the two proteins is 4.6 Å (Figure 4 C). It is therefore unlikely that the AmpC β-lactamase hinders the conformational freedom of β1AR in regions other than the directly fused helices.

Two disulphide bonds hold extracellular loop 2 (ECL2) in place with respect to TM3, and ECL2 has been suggested to control access to the orthosteric ligand binding site.[27] The connection between the loop region and a short helix within ECL is further stabilized by a coordinated Na^+ ion.[4]

The disulphide bond between Cys 114 and Cys 199 is defined in our cryo-EM density map. In the region of the disulphide bond between Cys 192 and Cys 198, on the other hand, the cryo-EM density is difficult to interpret, and the Na^+ ion is not apparent in our structure.

Discussion

The β1-adrenoceptor – AmpC β-lactamase fusion protein was designed to increase the mass of the receptor to improve the signal-to-noise ratio in the acquired cryo-EM data, and to introduce structural features that facilitate angular assignment of the particles in single-particle analysis, with the final aim of establishing a widely applicable engineering strategy for solving GPCR cryo-EM structures without forming active complexes with G-proteins, arrestin or artificial binders.

Many fusion proteins that have been used in GPCR crystallography were connected to the receptor via loops. To achieve rigidity using loops, optimization of the linkers is typically required (e.g. Refs[28,29]). Fusion protein connections that work for crystallization may not necessarily be rigid in solution or even in crystals. For example, in a crystal structure of a nanobody-stabilized β^2AR construct containing a T4 lysozyme (T4L) fusion protein, the connections between the receptor and the fusion protein were flexible, resulting in averaging out of T4L in the structure, producing no interpretable electron density for T4L.[30]

Other connections formed chimeric extended helices that have the potential to be rigid also in solution, outside the crystal lattice (e.g. Ref.[31]).

GPCRs can correctly fold even if the folding flow is interrupted by fusion proteins in intracellular loops. This tolerance towards insertions may be related to the presence of independent folding units. It has been shown in GPCR research that rhodopsin, but also other GPCRs, can be produced as separate fragments, and that these fragments can combine with matching counterparts to form receptors that are functional in receptor-binding.[28,32-34]

Here, we specifically searched for a fusion protein that is larger than thermostabilized apocytochrome b_{562} RIL (BRIL)(e.g. Ref.[31]) but can also be genetically inserted at the location of ICL3 to form rigid, extended helices with TM5 and TM6 of the GPCR.[15]

The resulting cryo-EM map of cyanopindolol-bound ultra-thermostable β1-adrenergic receptor comprises information about the overall architecture of the GPCR, including the seven-helix transmembrane core, helix 8, and the ligand binding mode.

The overall architecture of GPCRs comprises an extracellular part composed of the N-terminus and three extracellular loops, a seven α-helix ligand-binding and information-transducing transmembrane domain core, and an intracellular part including three intracellular loops, the amphipathic helix 8, and the C-terminus.[35]

Thermostability-based ligand binding assays showed that the receptor, in the context of the fusion protein, efficiently binds cyanopindolol (and propranolol), confirming structural integrity.

3D variability analysis of the cryo-EM map revealed conformational heterogeneities within the receptor and between the receptor and the fusion protein. Unlike the crystal structure, the cryo-EM map of the flexible protein complex reflects the dynamics, at the expense of precise side chain positions, which could only be determined for a subset of residues (mostly with aromatic side chains) in the structure. On the other hand, density for the ligand cyanopindolol, which binds deeply in the center of the extracellular region of the receptor, is clearly defined in the cryo-EM map and allows determination of the overall ligand-binding mode. Large scale conformational changes between transmembrane helices 5 and 6 are prevented by the helix connections to the fusion protein.

Helix 8 (H8) of β1AR is critical for receptor stability and activation,[36] and open questions remain about specific helix 8 functions. The base of helix 8 is wedged between helices 1 and 7, and appears like a lever arm, similar to a single α-helical domain.[37] It seems plausible that this placement of H8 may stabilize the helix 1-7 core transmembrane bundle, and that leverage from H8 could potentially affect the conformation of the core and have a function in mechanosensing. Involvement in such a function of helix 8 has been suggested for GPCRs which are known to be mechanosensitive.[38]

In a physiological environment, helix 8 is anchored to the cytoplasmic face of the cell membrane through palmitoyl groups.[24] In our construct sequence, the palmitoylation site (C358A) was removed by mutagenesis, and the protein was solubilized in detergent, hence no cellular membrane but only a micelle was present. Despite the absence of membrane anchoring or crystal contacts, helix 8 is defined in the cryo-EM map, in a typical position slanted upwards at a shallow angle towards the expected location of the cell membrane. Based on the localization of helix 8 relative to the lower rim of the DDM micelle, it is likely that the micelle stabilizes the helix through hydrophobic interactions.

The development of protein engineering approaches to enable the precise alignment of GPCR single particles in the absence of activating G proteins or arrestins is an active field of research. There have been recent reports of other strategies with similar aims, in most cases involving binding proteins such as nanobodies. For example, a newly developed nanobody that binds to an intracellular GPCR loop requires mutation of only a short stretch in the amino acid sequence of different receptors for binding.[39] Furthermore, Mukherjee et al. generated synthetic

antibodies that recognize apocytochrome b562 (BRIL), a frequently used fusion protein in GPCR structural biology, for cryo-EM applications.[40] Zhang et al. explored fusion protein connections with one or two continuous single helices, often in combination with a nanobody,[41] indicating that for structure elucidation at high resolution, either two stable helix connections, or one stable helix connection in combination with an additional, optimized linker or directed domain-to-domain contacts, are typically required.[41] Xu et al. designed an approach comprising a third receptor-to-fusion protein connection by fusing one chain of the calcineurin heterodimer into ICL3, and the other chain to the C-terminus of TM 7 of the receptor.[42] A similar three-point fusion using a GFP in combination with a GFP-binding nanobody has also been described for a transporter.[43]

In comparison to these and similar approaches, the AmpC β-lactamase we explored here is relatively large (much larger than for example BRIL), contributing strongly to a good signal-to-noise ratio even without binders, and makes no contacts to the receptor except for the direct helix fusions, allowing for the highest possible conformational freedom between receptor helices, except between TM5 and TM6, where movement is locked through the helix fusions.

AmpC β-lactamase has an ideal geometry for insertion into the third intracellular loop of typical class A GPCRs and is a promising tool for the elucidation of GPCR structures in the absence of natural or artificial binding proteins. Because fusion proteins do not require complex formation, this approach holds great promise as a tool for the elucidation of cryo-EM structures with different ligands at an increased throughput for drug discovery research. Other potential applications include the elucidation of unliganded (apo) cryo-EM structures, and the elucidation of cryo-EM structures of GPCRs with low expression levels.

Our approach demonstrates the feasibility of creating a single construct where the fusion partner AmpC β-lactamase is genetically fused to transmembrane helices 5 and 6 (TM5 and TM6) of the receptor, eliminating the need for additional binding partners such as nanobodies. Further studies may confirm that this strategy can be widely applied to engineer other class A GPCRs similarly, thereby increasing the number of structures available for future research.

Acknowledgements

We thank Mohamed Chami, Lubomir Kovacik and Carola Alampi at C-CINA Basel for support in cryo-EM data collection, the EM Facility at the PSI for technical support, and Takashi Ishikawa, Gebhard F.X. Schertler, Jan Pieter Abrahams, G.V. Shivashankar, Michel Steinmetz and Dmitry Veprintsev for supporting the project. We thank Hans Widmer, Sandra Jacob, Christian Wiesmann, Andreas Schenk and Gregor Cicchetti for interesting discussions about the project, Mara Wieser for support in protein production, Xavier Deupi for GPCR-related advice and Greta Assmann for computational support. We thank Daniel Böhringer for the coordination of the collaboration with the ScopeM Cryo-EM Knowledge Hub, and Julius Rabl for discussions and advice on the model building. This work was supported by grants from Novartis FreeNovation, the Promedica Stiftung Chur (1401/M) and the Swiss National Science Foundation (SNF SPARK, CRSK-3_190414) to R.M.B., and by grant 192760 from the Swiss National Science Foundation.

Author contributions

The project was initiated and coordinated by R.M.B. The fusion proteins were designed by R.M.B.. The script for parsing the PDB was written by A.L. and S.B.. The protein was expressed by G.C. and T.B and purified by G.C.. Biophysical characterization was performed by G.C.. Grids for cryo-EM data collection were prepared and screened by G.C., I.M. and E.P.. Initial data analysis was performed by G.C. and E.P.. The cryo-EM structure was elucidated by P.A., G.C., I.M and R.M.B.. The final structure was analyzed and interpreted by R.M.B., P.A., I.M. and G.C.. The manuscript was written by R.M.B., G.C. and P.A., with important contributions from all authors.

Materials and Methods

Receptor and fusion protein selection

For the design of the chimeric protein, we used a β1AR construct with established expression and purification protocols. The sequence was based on the ultra-thermostable mutant β1AR-JM50,[17] which contains the same mutations and truncations as the previous thermostabilized mutant β36-m23,[5] but also comprises three additional thermostabilizing mutations: I129V, D322K, Y343L. The mutations from the β36-m23 construct served the following purposes: R68S, M90V, Y227A, A282L, F327A, F338M contributed to the increased thermostability and

to the antagonist state,[4,44] C358A eliminated the palmitoylation site, C116L resulted in improved expression levels, and the truncations at the N-terminus, C-terminus (and at intracellular loop 3) reduced flexible regions of the protein.[4,5,18] The construct furthermore contains a D200E mutation, located in extracellular loop 2.

Fusion proteins were identified using a Java program. BioJava[45] was used for the analysis. All structures in the Protein Data Bank as of February 2017 were considered. Secondary structure was assigned using the DSSP algorithm,[46] implemented in BioJava. Candidates were considered if they had 15 residues annotated as helical near both termini. The terminal helices were required to fall within 11±1Å apart, as measured at the helix centroid. Finally, helices were required to be antiparallel within 15°. Structures which matched the filter criteria were then manually inspected, with 1FCO chosen as the most promising candidate.

The source code and documentation to run the program can be found in the following open repository: https://github.com/lafita/motif-search

Transfection and stable cell line generation

HEK293S GnTI(-) cells[20] were grown to ~90% confluence in a cell culture dish at 37 °C, 5% CO_2. Transfection was performed using Lipofectamine 3000 reagent (Invitrogen) according to the standard protocol. 100 µg/ml Zeocin were used for stable cell line selection.[19]

Protein expression

The stably transfected cells were first upscaled in adherent cultures and then transferred to suspension cultures. Suspension cultures were grown in PEM (Protein Expression Medium, Gibco) supplemented with 10% FBS, 1% GlutaMAX (GIBCO), Zeocin (Invivogen) to a final concentration of 100 µg/ml and Penicillin-Streptomycin (PAN Biotech) to a final concentration of 100 U/ml. The suspension cultures were grown in a shaker-incubator at 120 rpm, 37 °C, 5% CO_2. Cell densities were maintained between 800'000 and 1'500'000 cells/ml. Once the desired amount of cell culture volume was reached, the overexpression of the fusion protein was induced through the addition of tetracycline to a final concentration of 3 µg/ml. The induced cells were incubated for another 72 hours and then collected by centrifugation at 4'000 rpm, 15 min. at 4 °C and stored at -80 °C until further use.

Membrane preparation and protein solubilization

All subsequent steps were performed at 4 °C or on ice. The thawed cell pellets were resuspended in a pellet-to-buffer ratio of 1:5 in 50 mM HEPES, pH 7.5, 1 mM $MgCl_2$, 150 mM

NaCl, supplemented with DNAse I and complete protease inhibitor cocktail (Roche). Cell membranes were disrupted using a dounce homogenizer. The homogenized material was ultracentrifuged at 230'000 rcf for 45 min., 4 °C. The supernatant was discarded. The membranes were gently resuspended using an ultra-turrax dispersing machine (IKA, USA) in solubilization buffer (50 mM HEPES, pH 7.5, 150 mM NaCl), supplemented with complete protease inhibitor cocktail (Roche). For solubilization, β-DDM was added to a final concentration of 1% (w/v). The mixture was stirred at 150 rpm for 2 hours at 4 °C. Another ultracentrifugation step was carried out to clarify the solution containing the solubilized receptor.

Protein purification

The solubilized protein was incubated for 1 hour with washed and pre-equilibrated Strep-Tactin Superflow plus resin (Qiagen). The resin was then loaded into an Econo-Column Chromatography Column (Biorad) for gravity flow purification. The resin was washed 10 x with 2 volumes of wash buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.03 % (w/v) DDM). Stepwise protein elution was performed in wash buffer supplemented with 5 mM desthiobiotin (IBA-Lifesciences). The eluted protein was then treated with His-tagged human rhinovirus 3C protease (HRV 3C) to cleave off the C-terminal eGFP fusion and the Strep-tag. The protease was removed by Ni-NTA IMAC. The protein was then concentrated to 0.5 ml using a 50 kDa cutoff concentrator (Vivaspin, Sartorius, MWCO 50 kDa).

Size-exclusion chromatography was carried out on a Superdex 200 increase 10/300 GL column (GE Healthcare) in 50 mM HEPES, pH 7.5, 100 mM NaCl, 0.03% (w/v) DDM, 1 µM cyanopindolol. The protein was concentrated to 3 mg/ml using a 50 kDa cutoff concentrator (Vivaspin, Sartorius). The purified protein was frozen in liquid nitrogen and stored at -80 ºC.

Cryo-EM specimen preparation and data collection

For the final data collection, 3.5 µl of protein solution were applied onto the glow discharged Quantifoil Cu R1.2/1.3 200 grids (Electron Microscopy Sciences). The grids were blotted for 3s and plunge frozen into liquid ethane using a Vitrobot (Thermo Fisher Scientific) that was operated at 10 °C and 100% relative humidity. The data were collected on a 300 KeV FEI Titan Krios transmission electron microscope with a GATAN post-column Quantum-LS energy filter (20 eV zero-loss filter). 8,009 movies (40 frames each) were acquired in counting mode using SerialEM[47] with a total dose of 64 e-/Å^2 and pixel size 0.82 Å/pix. Data screening and selection

of good micrographs were performed using in house scripts.[48] Gain reference was estimated *a posteriori*[49] using the *relion_estimate_gain* tool of the Relion5 software.[50]

Single-particle analysis

Single-particle analysis was performed mostly in CryoSPARC v4.4.1 (Ref.[51]), however, many steps of the data processing were validated and optimized using Relion[50] and cisTEM[52] software packages. The processing pipeline is summarized in Supplementary Figure 1. Blob picking was initially used for the reference-free picking of 2,729,813 particles, out of which over a few iterations of 2D-classification 423,253 were selected and subjected for an *ab initio* initial model building based on randomly selected 20,000 particles. The particle set was then subjected to iterative rounds of 2D-classification and hetero-refinement for reducing the number of "bad" particles. Out of 164,992 "good" particles, 2D class-averages were obtained and were used as references for particle picking. The picked particles were combined again with those from the previous iteration of blob-picking, the duplicates were removed, and a stack of 2,865,326 particles was again subjected to iterative rounds of 2D-classification and hetero-refinements, resulting in the end in a stack of 183,537 "good" particles, which was further subjected to 3D-refinements using (Homo-/NU- and Local-refinement jobs). Various parameters and types of iterative jobs were tested to find an optimal workflow (including different manually created and dynamic masks, resolution limits and strength of low-pass filters), which resulted in a 3D-reconstruction at 4.3 Å resolution. Further 3D-classification jobs were performed to identify the largest 3D class-average out of 37,653 particles, reconstructed at 4.2Å resolution. The resulting 3D class-averages were analyzed to identify flexibility of the complex.

Model building

Initial models of the fusion protein were generated using AlphaFold 2 and 3[25,53], to include the overall architecture of the complete fusion protein in a single chain. Only AlphaFold2 predicted the geometry between the GPCR and the fusion protein as it was observed in the cryo-EM density. The model from AlphaFold2 was rigid body fitted into the cryo-EM map.

After initial real-space refinement, the GPCR part of the AlphaFold2 model was replaced by the model of a superposed high-resolution crystal structure[18] (PDB:4BVN), to obtain a more precise starting model. After additional real-space refinement, the AmpC β-lactamase part of the AlphaFold2 model was in the same way replaced by a superposed high-resolution structure of this protein[54] (PDB:6T3D) to create the model for further refinement. The overall structure

of the complete fusion protein could be built into the cryo-EM map, from Serine 37 to Arginine 361 of the receptor. The model was refined in iterative steps of model building in Coot[55] version 0.9.8.3 and real space refinement in phenix[56] version 1.20-4459, using PDB:4BVN and PDB:6T3D for reference model restraints. MolProbity[57] was used for validation of the geometry and stereochemistry of the model (Supplementary Table 1).

The model was refined as a continuously numbered single chain. After refinement, the GPCR part was renumbered to match the correct receptor numbering also in TM6 and 7. AmpC β-lactamase was renumbered to match the numbering in the crystal structure originally used in the construct design[16] (PDB:1FCO) plus 1000, to keep it separate from the receptor numbering.

References

1. Weis, W.I., and Kobilka, B.K. (2018). The Molecular Basis of G Protein-Coupled Receptor Activation. Annu Rev Biochem *87*, 897-919. 10.1146/annurev-biochem-060614-033910.

2. Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Kuhn, P., Weis, W.I., Kobilka, B.K., and Stevens, R.C. (2007). High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science *318*, 1258-1265. 10.1126/science.1150577.

3. Rasmussen, S.G., Choi, H.J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R., Sanishvili, R., Fischetti, R.F., et al. (2007). Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature *450*, 383-387. 10.1038/nature06325.

4. Warne, T., Serrano-Vega, M.J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G.W., Tate, C.G., and Schertler, G.F.X. (2008). Structure of a β-adrenergic G-protein-coupled receptor. Nature *454*, 486-491. 10.1038/nature07101.

5. Warne, T., Serrano-Vega, M.J., Tate, C.G., and Schertler, G.F.X. (2009). Development and crystallization of a minimal thermostabilised G protein-coupled receptor. Protein Expres Purif *65*, 204-213. 10.1016/j.pep.2009.01.014.

6. Kuhlbrandt, W. (2014). Biochemistry. The resolution revolution. Science *343*, 1443-1444. 10.1126/science.1251652.

7. Afanasyev, P., Seer-Linnemayr, C., Ravelli, R.B.G., Matadeen, R., De Carlo, S., Alewijnse, B., Portugal, R.V., Pannu, N.S., Schatz, M., and van Heel, M. (2017). Single-particle cryo-EM using alignment by classification (ABC): the structure of Lumbricus terrestris haemoglobin. IUCrJ *4*, 678-694. 10.1107/S2052252517010922.

8. Nygaard, R., Kim, J., and Mancia, F. (2020). Cryo-electron microscopy analysis of small membrane proteins. Curr Opin Struct Biol *64*, 26-33. 10.1016/j.sbi.2020.05.009.

9. Danev, R., Belousoff, M., Liang, Y.L., Zhang, X., Eisenstein, F., Wootten, D., and Sexton, P.M. (2021). Routine sub-2.5 A cryo-EM structure determination of GPCRs. Nat Commun *12*, 4333. 10.1038/s41467-021-24650-3.

10. Liang, Y.L., Khoshouei, M., Radjainia, M., Zhang, Y., Glukhova, A., Tarrasch, J., Thal, D.M., Furness, S.G.B., Christopoulos, G., Coudrat, T., et al. (2017). Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. Nature *546*, 118-123. 10.1038/nature22327.

11. Zhang, Y., Sun, B., Feng, D., Hu, H., Chu, M., Qu, Q., Tarrasch, J.T., Li, S., Sun Kobilka, T., Kobilka, B.K., and Skiniotis, G. (2017). Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. Nature *546*, 248-253. 10.1038/nature22394.

12. Huang, W., Masureel, M., Qu, Q., Janetzko, J., Inoue, A., Kato, H.E., Robertson, M.J., Nguyen, K.C., Glenn, J.S., Skiniotis, G., and Kobilka, B.K. (2020). Structure of the neurotensin receptor 1 in complex with beta-arrestin 1. Nature *579*, 303-308. 10.1038/s41586-020-1953-1.

13. Staus, D.P., Hu, H., Robertson, M.J., Kleinhenz, A.L.W., Wingler, L.M., Capel, W.D., Latorraca, N.R., Lefkowitz, R.J., and Skiniotis, G. (2020). Structure of the M2

muscarinic receptor-beta-arrestin complex in a lipid nanodisc. Nature *579*, 297-302. 10.1038/s41586-020-1954-0.

14. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235-242. 10.1093/nar/28.1.235.

15. Krebs, A.S., Bierig, T., Collu, G., and Benoit, R.M. (2019). Seamless insert-plasmid assembly at sub-terminal homologous sequences. Plasmid *106*, 102445. 10.1016/j.plasmid.2019.102445.

16. Patera, A., Blaszczak, L.C., Shoichet, B.K (2000). Crystal structures of substrate and inhibitor complexes with AmpC β-lactamase: possible implications for substrate-assisted catalysis. J. Am. Chem. Soc. *122*, 10504-10512.

17. Miller, J.L., and Tate, C.G. (2011). Engineering an ultra-thermostable beta(1)-adrenoceptor. J Mol Biol *413*, 628-638. 10.1016/j.jmb.2011.08.057.

18. Miller-Gallacher, J.L., Nehme, R., Warne, T., Edwards, P.C., Schertler, G.F., Leslie, A.G., and Tate, C.G. (2014). The 2.1 A resolution structure of cyanopindolol-bound beta1-adrenoceptor identifies an intramembrane Na+ ion that stabilises the ligand-free receptor. PLoS One *9*, e92727. 10.1371/journal.pone.0092727.

19. Chaudhary, S., Pak, J.E., Gruswitz, F., Sharma, V., and Stroud, R.M. (2012). Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. Nat Protoc *7*, 453-466. 10.1038/nprot.2011.453.

20. Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002). 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 U S A *99*, 13419-13424. 10.1073/pnas.212519299.

21. Alexandrov, A.I., Mileni, M., Chien, E.Y., Hanson, M.A., and Stevens, R.C. (2008). Microscale fluorescent thermal stability assay for membrane proteins. Structure *16*, 351-359. 10.1016/j.str.2008.02.004.

22. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci *27*, 14-25. 10.1002/pro.3235.

23. The PyMOL Molecular Graphics System, Version 3.1.0 Schrödinger, LLC.

24. O'Dowd, B.F., Hnatowich, M., Caron, M.G., Lefkowitz, R.J., and Bouvier, M. (1989). Palmitoylation of the human beta 2-adrenergic receptor. Mutation of Cys341 in the carboxyl tail leads to an uncoupled nonpalmitoylated form of the receptor. J Biol Chem *264*, 7564-7569.

25. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-03819-2.

26. Ballesteros A., W.H. (1995). 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. 10.1016/S1043-9471(05)80049-7.

27. Gonzalez, A., Perez-Acle, T., Pardo, L., and Deupi, X. (2011). Molecular basis of ligand dissociation in beta-adrenergic receptors. PLoS One *6*, e23815. 10.1371/journal.pone.0023815.

28. Rosenbaum, D.M., Cherezov, V., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Yao, X.J., Weis, W.I., Stevens, R.C., and Kobilka, B.K. (2007). GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. Science *318*, 1266-1273. 10.1126/science.1150609.

29. Smyth, D.R., Mrozkiewicz, M.K., McGrath, W.J., Listwan, P., and Kobe, B. (2003). Crystal structures of fusion proteins with large-affinity tags. Protein Sci *12*, 1313-1322. 10.1110/ps.0243403.

30. Rasmussen, S.G., Choi, H.J., Fung, J.J., Pardon, E., Casarosa, P., Chae, P.S., Devree, B.T., Rosenbaum, D.M., Thian, F.S., Kobilka, T.S., et al. (2011). Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. Nature *469*, 175-180. 10.1038/nature09648.

31. Wacker, D., Wang, S., McCorvy, J.D., Betz, R.M., Venkatakrishnan, A.J., Levit, A., Lansu, K., Schools, Z.L., Che, T., Nichols, D.E., et al. (2017). Crystal Structure of an LSD-Bound Human Serotonin Receptor. Cell *168*, 377-389 e312. 10.1016/j.cell.2016.12.033.

32. Huang, K.S., Bayley, H., Liao, M.J., London, E., and Khorana, H.G. (1981). Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. J Biol Chem *256*, 3802-3809.

33. Karnik, S.S. (2002). Analysis of structure-function from expression of G protein-coupled receptor fragments. Methods Enzymol *343*, 248-259. 10.1016/s0076-6879(02)43140-0.

34. Kobilka, B.K., Kobilka, T.S., Daniel, K., Regan, J.W., Caron, M.G., and Lefkowitz, R.J. (1988). Chimeric alpha 2-,beta 2-adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. Science *240*, 1310-1316. 10.1126/science.2836950.

35. Venkatakrishnan, A.J., Deupi, X., Lebon, G., Tate, C.G., Schertler, G.F., and Babu, M.M. (2013). Molecular signatures of G-protein-coupled receptors. Nature *494*, 185-194. 10.1038/nature11896.

36. Delos Santos, N.M., Gardner, L.A., White, S.W., and Bahouth, S.W. (2006). Characterization of the residues in helix 8 of the human beta1-adrenergic receptor that are involved in coupling the receptor to G proteins. J Biol Chem *281*, 12896-12907. 10.1074/jbc.M508500200.

37. Collu, G., Bierig, T., Krebs, A.S., Engilberge, S., Varma, N., Guixa-Gonzalez, R., Sharpe, T., Deupi, X., Olieric, V., Poghosyan, E., and Benoit, R.M. (2022). Chimeric single alpha-helical domains as rigid fusion protein connections for protein nanotechnology and structural biology. Structure *30*, 95-106 e107. 10.1016/j.str.2021.09.002.

38. Erdogmus, S., Storch, U., Danner, L., Becker, J., Winter, M., Ziegler, N., Wirth, A., Offermanns, S., Hoffmann, C., Gudermann, T., and Mederos, Y.S.M. (2019). Helix 8 is the essential structural motif of mechanosensitive GPCRs. Nat Commun *10*, 5784. 10.1038/s41467-019-13722-0.

39. Robertson, M.J., Papasergi-Scott, M.M., He, F., Seven, A.B., Meyerowitz, J.G., Panova, O., Peroto, M.C., Che, T., and Skiniotis, G. (2022). Structure determination of inactive-state GPCRs with a universal nanobody. Nat Struct Mol Biol *29*, 1188-1195. 10.1038/s41594-022-00859-8.

40. Mukherjee, S., Erramilli, S.K., Ammirati, M., Alvarez, F.J.D., Fennell, K.F., Purdy, M.D., Skrobek, B.M., Radziwon, K., Coukos, J., Kang, Y., et al. (2020). Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. Nat Commun *11*, 1598. 10.1038/s41467-020-15363-0.

41. Zhang, K., Wu, H., Hoppe, N., Manglik, A., and Cheng, Y. (2022). Fusion protein strategies for cryo-EM study of G protein-coupled receptors. Nat Commun *13*, 4366. 10.1038/s41467-022-32125-2.

42. Xu, J., Chen, G., Wang, H., Cao, S., Heng, J., Deupi, X., Du, Y., and Kobilka, B.K. (2024). Calcineurin-fusion facilitates cryo-EM structure determination of a Family A GPCR. Proc Natl Acad Sci U S A *121*, e2414544121. 10.1073/pnas.2414544121.

43. Niu, Y., Liu, R., Guan, C., Zhang, Y., Chen, Z., Hoerer, S., Nar, H., and Chen, L. (2022). Structural basis of inhibition of the human SGLT2-MAP17 glucose transporter. Nature *601*, 280-284. 10.1038/s41586-021-04212-9.

44. Serrano-Vega, M.J., Magnani, F., Shibata, Y., and Tate, C.G. (2008). Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. Proc Natl Acad Sci U S A *105*, 877-882. 10.1073/pnas.0711253105.

45. Lafita, A., Bliven, S., Prlic, A., Guzenko, D., Rose, P.W., Bradley, A., Pavan, P., Myers-Turnbull, D., Valasatava, Y., Heuer, M., et al. (2019). BioJava 5: A community driven open-source bioinformatics library. PLoS Comput Biol *15*, e1006791. 10.1371/journal.pcbi.1006791.

46. Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577-2637. 10.1002/bip.360221211.

47. Mastronarde, D.N. (2003). SerialEM: A program for automated tilt series acquisition on Tecnai microscopes using predition of specimen position. Microscopy and Microanalysis *9*, 1182-1183. 10.1017/S1431927603445911.

48. Afanasyev, P. (2023). CryoEM tools. https://doi.org/10.5281/zenodo.8187481.

49. Afanasyev, P., Ravelli, R.B., Matadeen, R., De Carlo, S., van Duinen, G., Alewijnse, B., Peters, P.J., Abrahams, J.P., Portugal, R.V., Schatz, M., and van Heel, M. (2015). A posteriori correction of camera characteristics from large image data sets. Sci Rep *5*, 10317. 10.1038/srep10317.

50. Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530. 10.1016/j.jsb.2012.09.006.

51. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290-296. 10.1038/nmeth.4169.

52. Grant, T., Rohou, A., and Grigorieff, N. (2018). cisTEM, user-friendly software for single-particle image processing. Elife *7*. 10.7554/eLife.35383.

53. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493-500. 10.1038/s41586-024-07487-w.

54. Lang, P.A., Parkova, A., Leissing, T.M., Calvopina, K., Cain, R., Krajnc, A., Panduwawala, T.D., Philippe, J., Fishwick, C.W.G., Trapencieris, P., et al. (2020). Bicyclic Boronates as Potent Inhibitors of AmpC, the Class C beta-Lactamase from Escherichia coli. Biomolecules *10*. 10.3390/biom10060899.

55. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501. 10.1107/S0907444910007493.

56. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol *75*, 861-877. 10.1107/S2059798319011471.

57. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci *27*, 293-315. 10.1002/pro.3330.