

## List of Publications

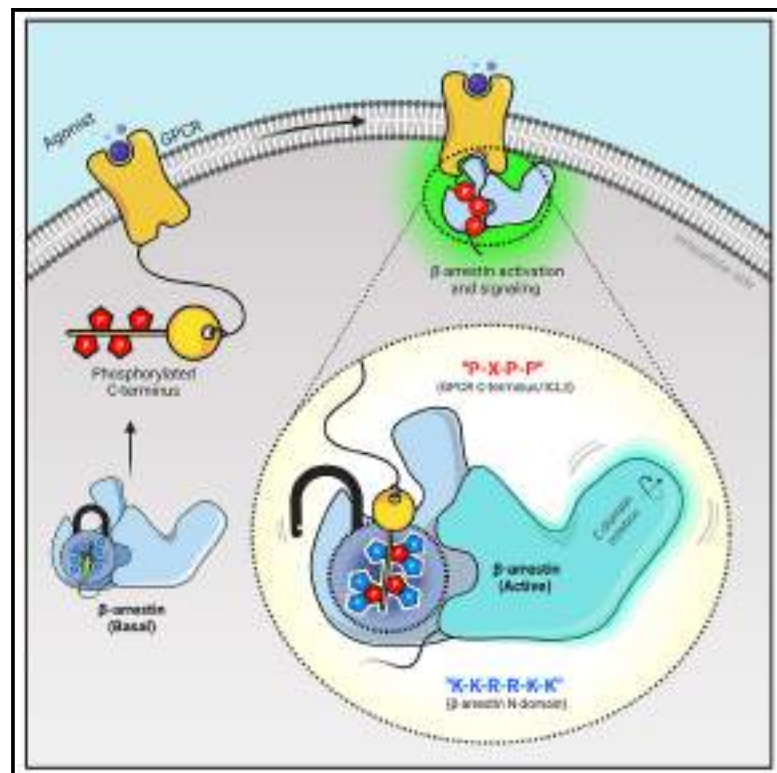
1. **Sarma P**, Carino CMC, Seetharama D, Pandey S, Dwivedi-Agnihotri H, Rui X, Cao Y, Kawakami K, Kumari P, Chen YC, Luker KE, Yadav PN, Luker GD, Laporte SA, Chen X, Inoue A, Shukla AK. Molecular insights into intrinsic transducer-coupling bias in the CXCR4-CXCR7 system. **Nature Communications**. 2023 Aug 9;14(1):4808.
2. Maharana J\*, **Sarma P\***, Yadav MK, Saha S, Singh V, Saha S, Chami M, Banerjee R, Shukla AK. Structural snapshots uncover a key phosphorylation motif in GPCRs driving  $\beta$ -arrestin activation. **Molecular Cell**. 2023 Jun 15;83(12):2091-2107.e7. (**Joint 1<sup>st</sup> author**).  
This article was featured on the cover page of the journal.
3. Yadav MK\*, Maharana J\*, Yadav R\*, Saha S\*, **Sarma P\***, Soni C, Singh V, Saha S, Ganguly M, Li XX, Mahapatra S, Mishra S, Khant HA, Chami M, Woodruff TM, Banerjee R, Shukla AK and Gati C. Molecular basis of anaphylatoxin recognition, activation, and signaling-bias at complement receptors. **Cell** (in press). 2023. (**Joint 1<sup>st</sup> author**).
4. Yadav MK\*, **Sarma P\***, Ganguly M, Mishra S, Maharana J, Zaidi N, Dalal A, Singh V, Saha S, Mahajan G, Sharma S, Chami M, Banerjee R, Shukla AK. Structure-guided engineering of biased-agonism in the human niacin receptor via single amino acid substitution. 2023. (**Manuscript under review**). (**Joint 1<sup>st</sup> author**).
5. Maharana J\*, Sano FK\*, **Sarma P\***, Yadav MK, Duan L, Stepniwski TM, Chaturvedi M, Ranjan A, Singh V, Saha S, Mahajan G, Chami M, Shihoya W, Selent J, Chung KY, Banerjee R, Nureki O, Shukla AK. Molecular insights into atypical modes of beta-arrestin interaction with seven transmembrane receptors. 2023. (**Manuscript under review**). (**Joint 1<sup>st</sup> author**).
6. Isaikina P, Petrovic I, Jakob RP, **Sarma P**, Ranjan A, Baruah M, Panwalkar V, Maier T, Shukla AK, Grzesiek S. A key GPCR phosphorylation motif discovered in arrestin2-CCR5 phosphopeptide complexes. **Molecular Cell**. 2023 Jun 15;83(12):2108-2121.e7.
7. **Sarma P**, Shukla AK. Resonating with the signaling bias of CXCR7. **Molecular Cell**. 2022 Sep 15;82(18):3318-3320.
8. **Sarma P**, Banerjee R, Shukla AK. Structural snapshot of a  $\beta$ -arrestin-biased receptor. **Trends in Pharmacological Sciences**. 2023 Jan;44(1):1-3.
9. Baidya M, Chaturvedi M, Dwivedi-Agnihotri H, Ranjan A, Devost D, Namkung Y, Stepniwski TM, Pandey S, Baruah M, Panigrahi B, **Sarma P**, Yadav MK, Maharana J, Banerjee R, Kawakami K, Inoue A, Selent J, Laporte SA, Hébert TE, Shukla AK. Allosteric modulation of GPCR-induced  $\beta$ -arrestin trafficking and signaling by a synthetic intrabody. **Nature Communications**. 2022 Aug 8;13(1):4634.
10. Maharana J, Banerjee R, Yadav MK, **Sarma P**, Shukla AK. Emerging structural insights into

GPCR- $\beta$ -arrestin interaction and functional outcomes. **Current Opinion in Structural Biology**. 2022 Aug;75:102406.

11. **Sarma P**, Saha S, Shukla AK. Making the switch: The role of Gq in driving GRK selectivity at GPCRs. **Science Signaling**. 2022 Mar 22;15(726):eabo4949.
12. Maharana J, **Sarma P**, Shukla AK. Scratching the itch with cryo-EM. **Nature Chemical Biology**. 2022 Mar;18(3):242-243.

# Structural snapshots uncover a key phosphorylation motif in GPCRs driving $\beta$ -arrestin activation

## Graphical abstract



## Authors

Jagannath Maharana,  
Parishmita Sarma, Manish K. Yadav, ...,  
Mohamed Chami, Ramanuj Banerjee,  
Arun K. Shukla

## Correspondence

ramanujb@iitk.ac.in (R.B.),  
arshukla@iitk.ac.in (A.K.S.)

## In brief

Maharana et al. determine multiple structures of activated  $\beta$ -arrestins in complex with the carboxyl terminus phosphopeptides of different GPCRs using cryo-EM. Based on these structural snapshots, they discover and experimentally validate a significantly conserved phosphorylation motif in GPCRs that drives  $\beta$ -arrestin interaction and activation.

## Highlights

- Cryo-EM structure determination of  $\beta$ -arrestins in complex with GPCR phosphopeptides
- Identification of P-X-P-P motif in GPCRs for  $\beta$ -arrestin interaction and activation
- Experimental validation of P-X-P-P motif using mutagenesis and conformational sensor
- Discovery of a significantly conserved  $\beta$ -arrestin activation mechanism by GPCRs



## Article

# Structural snapshots uncover a key phosphorylation motif in GPCRs driving $\beta$ -arrestin activation

Jagannath Maharana,<sup>1,3</sup> Parishmita Sarma,<sup>1,3</sup> Manish K. Yadav,<sup>1</sup> Sayantan Saha,<sup>1</sup> Vinay Singh,<sup>1</sup> Shirsha Saha,<sup>1</sup> Mohamed Chami,<sup>2</sup> Ramanuj Banerjee,<sup>1,\*</sup> and Arun K. Shukla<sup>1,4,\*</sup>

<sup>1</sup>Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur 208016, India

<sup>2</sup>BioEM Lab, Biozentrum, University of Basel, 4056 Basel, Switzerland

<sup>3</sup>These authors contributed equally

<sup>4</sup>Lead contact

\*Correspondence: [ramanujb@iitk.ac.in](mailto:ramanujb@iitk.ac.in) (R.B.), [arshukla@iitk.ac.in](mailto:arshukla@iitk.ac.in) (A.K.S.)

<https://doi.org/10.1016/j.molcel.2023.04.025>

## SUMMARY

Agonist-induced GPCR phosphorylation is a key determinant for the binding and activation of  $\beta$ -arrestins ( $\beta$ arrs). However, it is not entirely clear how different GPCRs harboring divergent phosphorylation patterns impart converging active conformation on  $\beta$ arrs leading to broadly conserved functional responses such as desensitization, endocytosis, and signaling. Here, we present multiple cryo-EM structures of activated  $\beta$ arrs in complex with distinct phosphorylation patterns derived from the carboxyl terminus of different GPCRs. These structures help identify a P-X-P-P type phosphorylation motif in GPCRs that interacts with a spatially organized K-K-R-R-K-K sequence in the N-domain of  $\beta$ arrs. Sequence analysis of the human GPCRome reveals the presence of this phosphorylation pattern in a large number of receptors, and its contribution in  $\beta$ arr activation is demonstrated by targeted mutagenesis experiments combined with an intra-body-based conformational sensor. Taken together, our findings provide important structural insights into the ability of distinct GPCRs to activate  $\beta$ arrs through a significantly conserved mechanism.

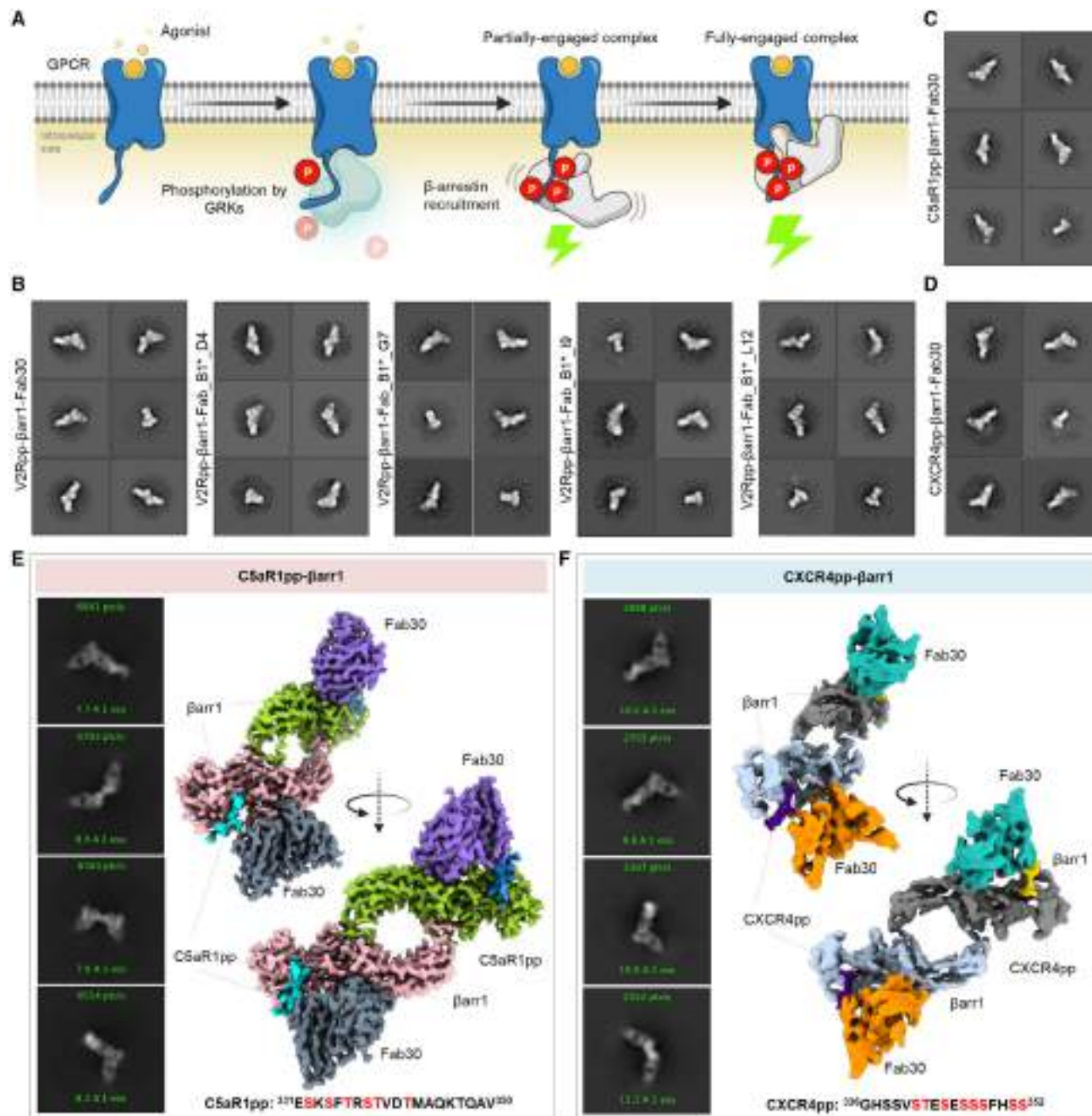
## INTRODUCTION

G protein-coupled receptors (GPCRs) are typically characterized by their conserved seven transmembrane (7TM) architecture and agonist-induced coupling to heterotrimeric G-proteins and  $\beta$ -arrestins ( $\beta$ arrs).<sup>1</sup> Of these,  $\beta$ arrs are multifunctional cytosolic proteins critically involved in regulating the signaling and trafficking of GPCRs.<sup>2–4</sup> Their interaction with GPCRs involves a major contribution from receptor phosphorylation, which not only drives the affinity of receptor- $\beta$ arr interaction but also imparts functionally competent active conformation in  $\beta$ arrs driving downstream functional outcomes<sup>5–12</sup> (Figure 1A). Additional interaction of  $\beta$ arrs with the receptor transmembrane core and membrane lipid bilayer induces further structural changes as visualized using biophysical and direct structural approaches,<sup>13–20</sup> which also fine-tune the functional capabilities of  $\beta$ arrs and possibly spatiotemporal aspects of their regulatory mechanisms.<sup>21,22</sup> Despite a poorly conserved primary sequence of GPCRs in terms of the number and spatial positioning of the putative phosphorylation sites, the near-universal nature of  $\beta$ arr interaction, ensuing signaling and regulatory responses, remains to be fully understood at the molecular level. Moreover, differential receptor phosphorylation by different subtypes of

GPCR kinases (GRKs), and possibly other kinases, in cell type-specific manner adds further complexity to GPCR- $\beta$ arr binding modalities and context-specific functional specialization.<sup>23–29</sup>

There are two isoforms of  $\beta$ arrs namely,  $\beta$ arr1 and 2, also known as Arrestin2 and 3, respectively, and despite a high degree of sequence and structural similarity, they often exhibit a significant diversity in their functional contribution toward GPCR signaling and regulation.<sup>30</sup> There are only a very few structures of  $\beta$ arr1 in active conformation, either in complex with receptor-derived phosphopeptides<sup>31,32</sup> or in complex with agonist-bound, phosphorylated GPCRs.<sup>15–20</sup> The structural coverage for  $\beta$ arr2 is even more sparse with the structural snapshots limited to either a complex with CXC chemokine receptor subtype 7 (CXCR7)-derived phosphopeptide<sup>33</sup> or in IP6-bound state.<sup>34</sup> Although these structures provide useful information about  $\beta$ arrs' interaction with the receptors, there is only limited information available about the phosphorylation patterns either due to chimeric constructs used in these studies or the lack of visualization of multiple phosphorylation sites resulting from insufficient structural resolution. Thus, the quest to decipher precise molecular details of convergent  $\beta$ arr activation by GPCRs harboring different phosphorylation patterns remains open. This represents a major knowledge gap in our current





**Figure 1. Reconstitution and structure determination of C5aR1pp/CXCR4pp- $\beta$ arr1 complexes**

(A) Agonist stimulation of GPCRs leads to receptor phosphorylation by GPCR kinases (GRKs) followed by the recruitment and activation of  $\beta$ arrestins governed through the phosphorylated residues and activated receptor core.

(B) Negative-staining EM-based 2D class averages of V2Rpp- $\beta$ arr1 complexes stabilized by Fab30, Fab\_B1\*\_D4, Fab\_B1\*\_G7, Fab\_B1\*\_I9, and Fab\_B1\*\_L12, respectively.

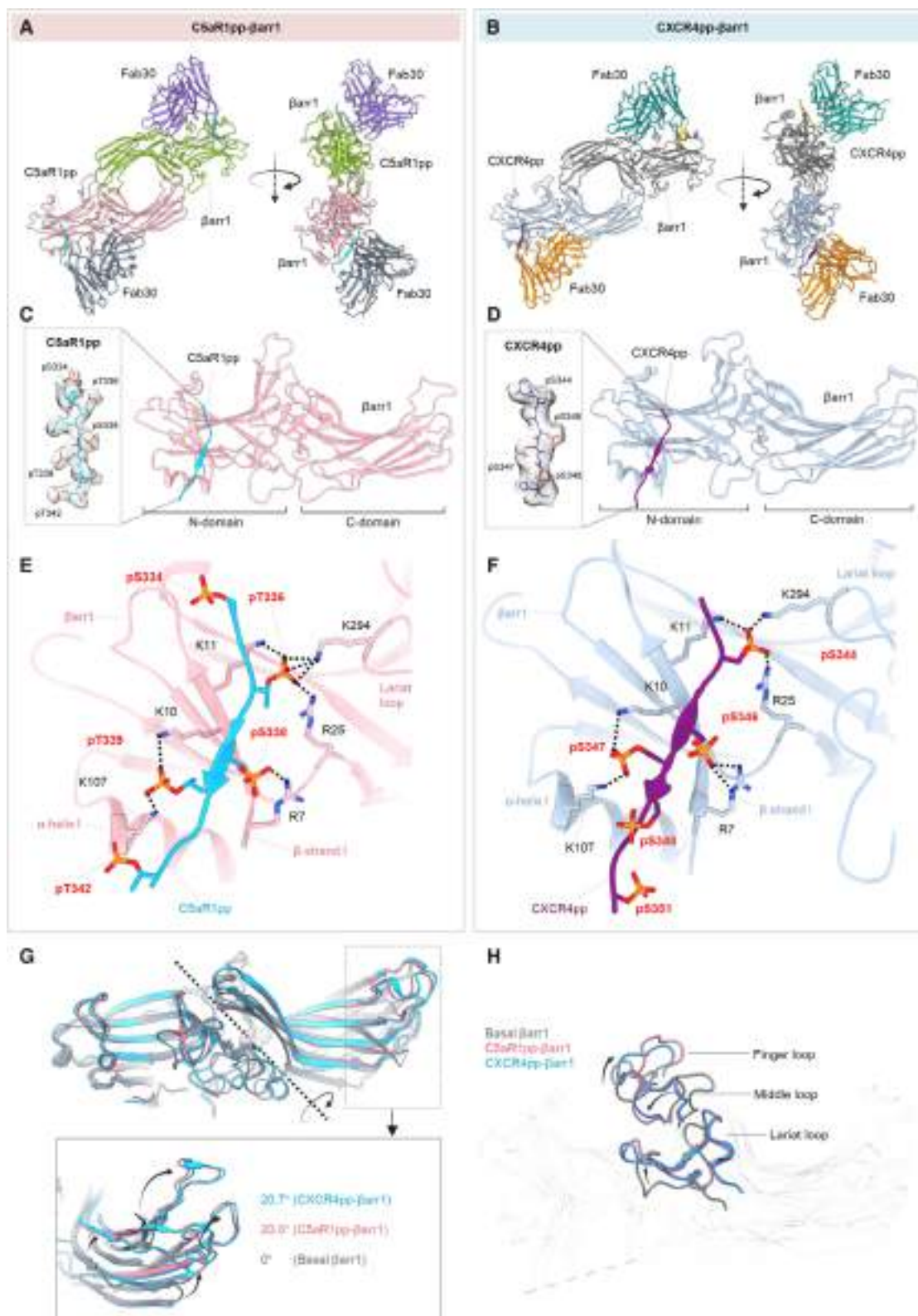
(C and D) Negative-staining EM-based 2D class averages of C5aR1pp- $\beta$ arr1-Fab30 and CXCR4pp- $\beta$ arr1-Fab30 complexes, respectively.

(E and F) Selected 2D class averages and surface representation of C5aR1pp- $\beta$ arr1-Fab30 and CXCR4pp- $\beta$ arr1-Fab30 structures, respectively, determined by cryo-EM. The missing Fab30 constant domain densities were truncated during local refinement. See also [Figures S1–S3](#) and [Table 1](#).

understanding of GPCR signaling and regulatory paradigms governing and fine-tuning signal transduction through this versatile class of receptors.

In this backdrop, here we present cryogenic-electron microscopy (cryo-EM) structures of full-length  $\beta$ arr1 and 2 activated by defined phosphorylation patterns encoded in the form of





phosphopeptides, which are derived from three different GPCRs, namely, the complement C5a receptor subtype 1 (C5aR1), the CXC chemokine receptor subtype 4 (CXCR4), and the vasopressin receptor subtype 2 (V2R). These structural snapshots reveal a P-X-P-P type pattern of phosphorylation in GPCRs that engages a K-K-R-R-K-K sequence in the N-domain of  $\beta$ arrs leading to  $\beta$ arr activation. Interestingly, a large repertoire of GPCRs encodes the P-X-P-P motif either in their carboxyl terminus or in the 3<sup>rd</sup> intracellular loop (ICL3), suggesting a broad implication of this activation mechanism. We further validate the contribution of the P-X-P-P motif with respect to  $\beta$ arr activation in cellular context for several GPCRs using an intrabody-based conformational biosensor and structure-guided mutagenesis studies. Collectively, our data help identify the presence of the P-X-P-P motif in GPCRs and uncover the molecular basis of its ability to activate  $\beta$ arrs.

## RESULTS

Although cryo-EM has been used to determine the structures of GPCR- $\beta$ arr1 complexes,<sup>15–19</sup> all the previous structures of  $\beta$ arrs in basal state,<sup>35–38</sup> bound to phosphopeptides<sup>31–33</sup> or IP6<sup>34</sup> have been determined using X-ray crystallography. Therefore, in order to test the feasibility of structure determination of  $\beta$ arrs in complex with different phosphorylation patterns encoded in GPCR phosphopeptides by cryo-EM, we first reconstituted V2Rpp- $\beta$ arr1 complex together with a set of conformationally selective antigen-binding fragments (Fabs), which recognize activated  $\beta$ arr1.<sup>39,40</sup> Subsequently, we analyzed these complexes using negative-staining single particle EM, which revealed monodisperse particle distribution and 2D class averages where the densities of  $\beta$ arr1 and Fabs were clearly discernible (Figures 1B and S1A). Based on these observations, we synthesized and characterized a set of phosphopeptides corresponding to the carboxyl terminus of the C5aR1 and the CXCR4, and assessed their ability to activate  $\beta$ arr1, measured in terms of Fab30 reactivity (Figures S1B–S1E). We identified the phosphorylation patterns from C5aR1 (C5aR1pp2; referred to as C5aR1pp hereon) and CXCR4 (CXCR4pp2; referred to as CXCR4pp hereon), which elicited maximal Fab30 reactivity as a measure of  $\beta$ arr1 activation (Figures S1B–S1E). Subsequently, we reconstituted C5aR1pp- $\beta$ arr1-Fab30 and CXCR4pp- $\beta$ arr1-Fab30 complexes, validated their monodispersity and architecture using negative-staining EM (Figures 1C, 1D, S1F, and S1G), and subjected

these complexes to cryo-EM. We successfully determined the structures of C5aR1pp- $\beta$ arr1-Fab30 and CXCR4pp- $\beta$ arr1-Fab30 complexes at global resolutions of 3.26 and 4.45 Å, respectively (Figures 1E, 1F, and S2).

### $\beta$ arr1 structures in complex with C5aR1 and CXCR4 phosphorylation patterns

Both structures revealed a dimeric arrangement with the two  $\beta$ arr1 protomers making contacts through the C-edge loops and finger loops (Figures 2A and 2B).  $\beta$ arr1 protomers exhibit nearly identical overall structures with each other (root-mean-square deviation [RMSD] < 0.5 Å) with clear densities of the phosphopeptides visible in the EM map (Figures S3A and S3B). C5aR1pp and CXCR4pp are positioned in a positively charged groove on the N-domain of  $\beta$ arr1 (Figures 2C and 2D), and the phosphate moieties make extensive contacts with Arg/Lys residues at the binding interface (Figures 2E and 2F). Interestingly, three phosphate groups arranged in a P-X-P-P type pattern, where P is a phosphorylated residue and X is any other amino acid, in both the phosphopeptides engage a nearly identical set of Lys/Arg residues in  $\beta$ arr1 (Figures 2E and 2F). Specifically, the pT<sup>336</sup>-R<sup>337</sup>-pS<sup>338</sup>-pT<sup>339</sup> pattern in C5aR1pp and pS<sup>344</sup>-E<sup>345</sup>-pS<sup>346</sup>-pS<sup>347</sup> pattern in CXCR4pp engages K<sup>294</sup>-K<sup>11</sup>-R<sup>25</sup>-R<sup>7</sup>-K<sup>10</sup>-K<sup>107</sup> in  $\beta$ arr1. The other phosphate groups present in the phosphopeptides are either not involved in direct contact or sparsely linked with Arg/Lys or positioned outside the binding groove. The N- and C-domains of  $\beta$ arr1 exhibit an inter-domain rotation of approximately 20° when compared with the basal state of  $\beta$ arr1, in both structures, which is a hallmark of  $\beta$ arr activation upon binding of phosphorylated GPCRs<sup>31</sup> (Figure 2G). Moreover, the three major loops in  $\beta$ arr1 namely, the finger, middle, and lariat loop also exhibit significant reorientation upon binding of C5aR1pp and CXCR4pp compared with the basal state, although their positioning is almost identical between the two structures (Figure 2H). Finally, the three-element interaction and polar-core network in  $\beta$ arr1 are also disrupted upon binding to C5aR1pp and CXCR4pp when compared with the basal state structure through the displacement of the carboxyl terminus of  $\beta$ arr1 from the N-domain and repositioning of the lariat loop through the interaction of K<sup>294</sup> with a phosphate moiety (Figures S4A and S4B). These structural features and interaction interface are analogous to that observed in the V2Rpp- $\beta$ arr1-Fab30 crystal structure determined previously,<sup>31</sup> although the primary sequence and phosphorylation patterns encoded by C5aR1pp and CXCR4pp are distinct from V2Rpp (Figure S4C).

### Figure 2. Overall structures and key structural features of C5aR1pp/CXCR4pp- $\beta$ arr1 complexes

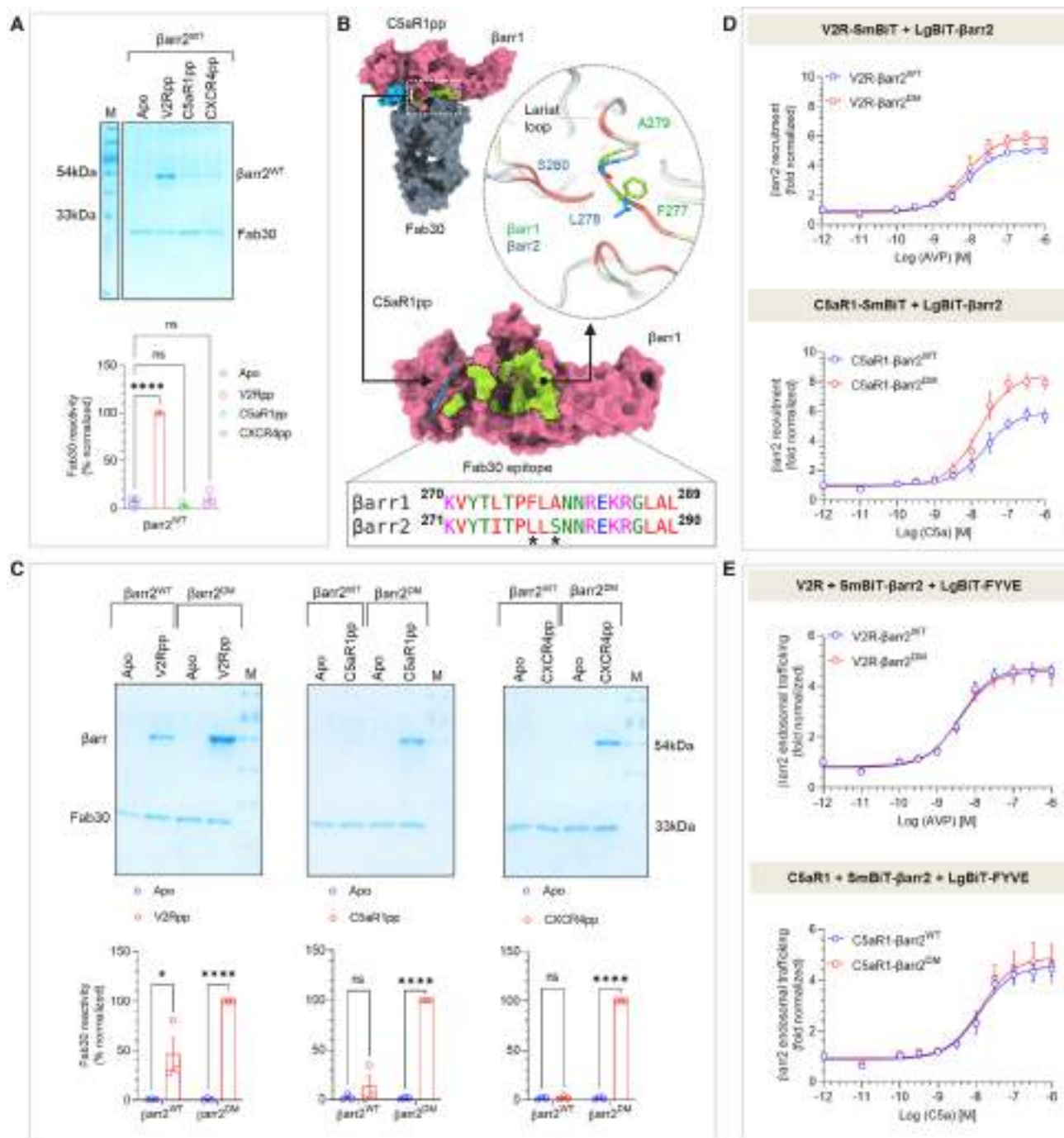
(A and B) Overall structures of C5aR1pp- $\beta$ arr1-Fab30 and CXCR4pp- $\beta$ arr1-Fab30 complexes shown with ribbon representation. The constant domains of Fab30 were masked out during refinement.

(C and D) Structure of individual C5aR1pp- $\beta$ arr1 and CXCR4pp- $\beta$ arr1 complex protomers shown as ribbon representation to indicate the binding of phosphopeptides on the N-domain of  $\beta$ arr1. Cryo-EM densities of corresponding phosphopeptides have been provided in insets.

(E and F) Stabilizing charge-charge interactions of C5aR1pp and CXCR4pp with the N-domain groove residues of  $\beta$ arr1 indicated as dotted lines. pS and pT refer to phospho-Ser and phospho-Thr residues, respectively.

(G) Inter-domain rotation in  $\beta$ arr1 upon binding of C5aR1pp (pink) and CXCR4pp (blue) is compared with the basal conformation of  $\beta$ arr1 determined previously (PDB: 1G4M, gray).

(H) Superimposition of C5aR1pp- and CXCR4pp-bound  $\beta$ arr1 structures with the basal conformation of  $\beta$ arr1 (PDB: 1G4M, gray) indicating the repositioning of finger, middle, and lariat loops upon  $\beta$ arr1 activation. See also Figure S4.



**Figure 3. Generation and characterization of  $\beta$ arr2<sup>DM</sup> for structure determination**

(A) Fab30 reactivity to C5aR1pp and CXCR4pp activated  $\beta$ arr2<sup>WT</sup> was measured by co-immunoprecipitation (coIP) assay. C5aR1pp and CXCR4pp activated  $\beta$ arr2<sup>WT</sup> were not recognized by Fab30 (top). Densitometry-based quantification of the coIP data is presented (bottom) (mean  $\pm$  SEM;  $n = 3$ ; normalized with respect to V2Rpp signal as 100%; one-way ANOVA, Dunnett's multiple comparisons test). The exact p values are as follows: Apo vs. V2Rpp ( $p \leq 0.0001$ ), Apo vs. C5aR1pp ( $p = 0.7719$ ), Apo vs. CXCR4pp ( $p = 0.7899$ ) (\*\*\*\* $p < 0.0001$ ; ns, non-significant).

(B) Comparison of the epitope region of Fab30 in  $\beta$ arr1 with  $\beta$ arr2 reveals that instead of F<sup>277</sup> and A<sup>279</sup> as in  $\beta$ arr1,  $\beta$ arr2 contains L<sup>278</sup> and S<sup>280</sup> in corresponding positions (indicated with an asterisk).

(C) CoIP assay showing the reactivity of Fab30 toward activated  $\beta$ arr2<sup>DM</sup> upon binding of C5aR1pp and CXCR4pp (top). Densitometry-based quantification is presented (bottom) (mean  $\pm$  SEM;  $n = 3$ ; normalized with respect to Fab30 reactivity toward activated  $\beta$ arr2<sup>DM</sup> treated as 100%; two-way ANOVA, Sidák's multiple comparisons test). The exact p values are as follows: for  $\beta$ arr2<sup>WT</sup>: Apo vs. V2Rpp ( $p = 0.018$ ), Apo vs. C5aR1pp ( $p = 0.3369$ ), Apo vs. CXCR4pp ( $p = 0.9338$ ); for  $\beta$ arr2<sup>DM</sup>: Apo vs. V2Rpp ( $p < 0.0001$ ), Apo vs. C5aR1pp ( $p < 0.0001$ ), Apo vs. CXCR4pp ( $p < 0.0001$ ) (\* $p < 0.05$ , \*\*\*\* $p < 0.0001$ ).

(legend continued on next page)



### Structure-guided engineering yields structures of activated $\beta$ arr2

As mentioned earlier, the structural coverage of active  $\beta$ arrs, either in complex with full GPCRs or GPCR-derived phosphopeptides, is limited primarily to  $\beta$ arr1. Activated structures of the other isoform i.e.,  $\beta$ arr2 are represented only by an IP6-bound crystal structure<sup>34</sup> and a complex of truncated  $\beta$ arr2 with a phosphopeptide derived from a  $\beta$ arr-biased seven-transmembrane receptor (7TMR) (CXCR7).<sup>33</sup> Therefore, we set out to reconstitute and determine the structure of  $\beta$ arr2 in complex with the phosphopeptides derived from different receptors, i.e., C5aR1pp and CXCR4pp using cryo-EM. Surprisingly, however, we did not observe a significant Fab30 reactivity to C5aR1pp/CXCR4pp-bound  $\beta$ arr2 while it robustly recognized the V2Rpp- $\beta$ arr2 complex (Figure 3A). Therefore, we analyzed the Fab30 interaction interface on the C5aR1pp-bound  $\beta$ arr1 structure to identify a potential reason for the lack of Fab30 reactivity with  $\beta$ arr2. Interestingly, we observed that Fab30 epitope is conserved between  $\beta$ arr1 and 2 with the exception of two residues, i.e., instead of F<sup>277</sup> and A<sup>279</sup> as in  $\beta$ arr1,  $\beta$ arr2 contains L<sup>278</sup> and S<sup>280</sup> in the corresponding positions, respectively (Figure 3B). Therefore, we generated a  $\beta$ arr2 double mutant, referred to as  $\beta$ arr2<sup>DM</sup>, and tested its reactivity to Fab30 upon activation by distinct phosphopeptides. In line with our hypothesis, we observed a robust interaction of Fab30 with C5aR1pp- $\beta$ arr2<sup>DM</sup> and CXCR4pp- $\beta$ arr2<sup>DM</sup> complex, and we also noticed that the interaction of Fab30 with V2Rpp- $\beta$ arr2<sup>DM</sup> was further enhanced compared with  $\beta$ arr2<sup>WT</sup> (Figure 3C). We also confirmed that  $\beta$ arr2<sup>DM</sup> exhibits a similar pattern of interaction as  $\beta$ arr2<sup>WT</sup> with V2R and C5aR1 in cellular context (Figure 3D) and shows near-identical endosomal trafficking as  $\beta$ arr2<sup>WT</sup> upon the stimulation of V2R and C5aR1 (Figure 3E). The receptor surface expression was assessed by whole-cell-based surface ELISA assay (Figure S5A). Thus,  $\beta$ arr2<sup>DM</sup> provides us with an excellent handle to reconstitute stable complexes with receptor phosphopeptides suitable for cryo-EM. In fact, we successfully managed to reconstitute monodisperse V2Rpp- $\beta$ arr2<sup>DM</sup>-Fab30 and C5aR1pp- $\beta$ arr2<sup>DM</sup>-Fab30 complexes (Figures S5C and S5I) and determine their cryo-EM structures at 3.96 and 4.33 Å resolution, respectively (Figures 4A, 4B, S5D–S5H, and S5J–S5N). In order to simplify the discussion, we refer to  $\beta$ arr2<sup>DM</sup> as  $\beta$ arr2 from here onward unless specified otherwise.

### Structures of $\beta$ arr2 in complex with V2Rpp and C5aR1pp

The V2Rpp- $\beta$ arr2 and C5aR1pp- $\beta$ arr2 structures exhibited a trimeric assembly of  $\beta$ arr2 with the individual protomers arranged through N- to C-domain contacts (Figures 4A–4D and S7; Table S2). The overall structural features of the individual protomers in each structure were nearly identical as reflected by low RMSD (<0.5 Å) with the phosphopeptide densities clearly visible in the EM maps (Figures S3C and S3D). Like  $\beta$ arr1 structures, V2Rpp and C5aR1pp are positioned in a positively charged groove on the N-domain of  $\beta$ arr2 (Figures 4E and 4F),

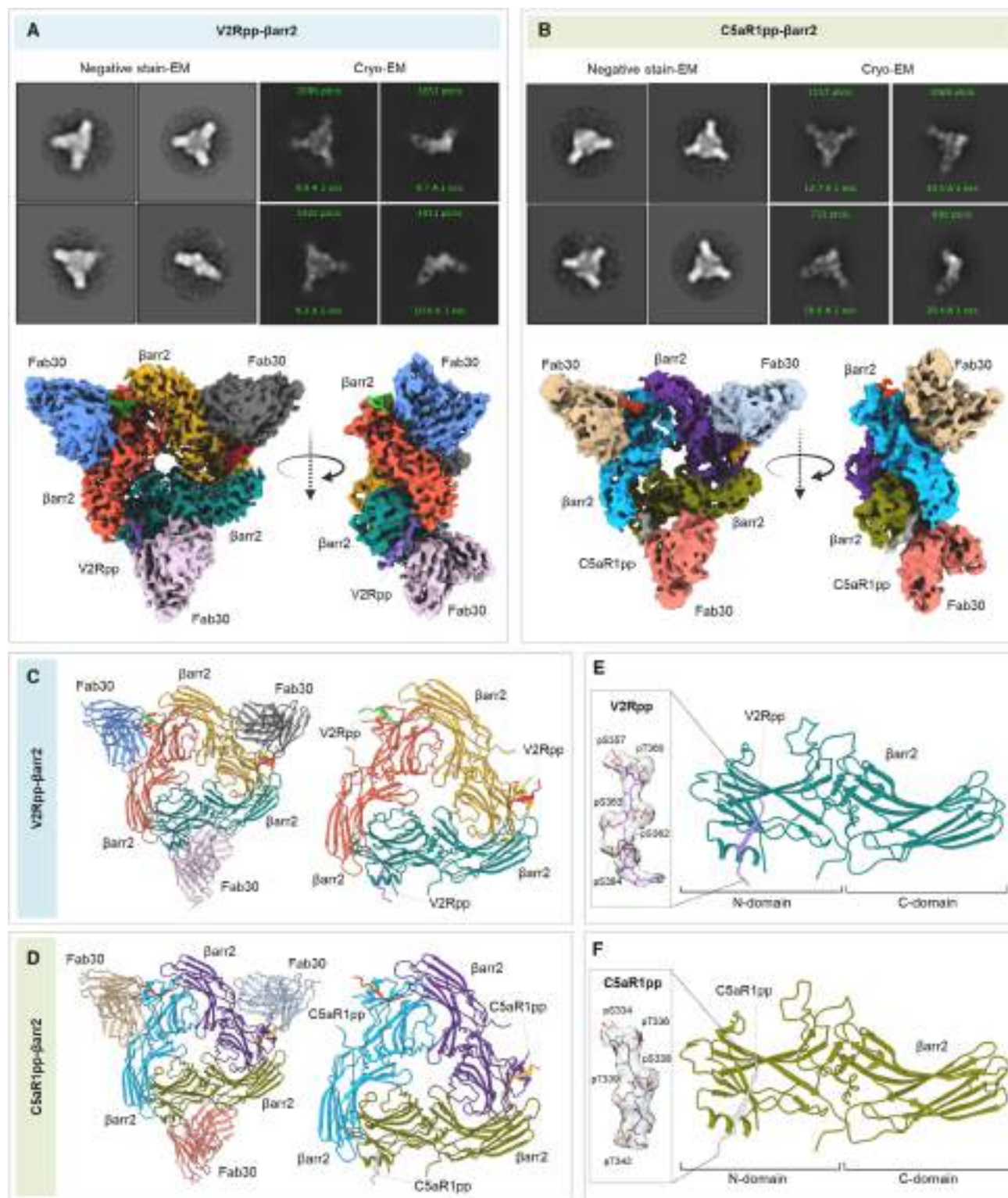
and the phosphate moieties make extensive contacts with Arg/Lys residues at the binding interface (Figures 5A and 5B). Remarkably, we observed that three phosphate groups arranged in a P-X-P-P pattern in the phosphopeptides, engage an analogous set of Lys/Arg residues as in  $\beta$ arr1 (Figures 5A and 5B). Specifically, pT<sup>360</sup>-A<sup>361</sup>-pS<sup>362</sup>-pS<sup>363</sup> in V2Rpp and pT<sup>336</sup>-R<sup>337</sup>-pS<sup>338</sup>-pT<sup>339</sup> in C5aR1pp engage K<sup>295</sup>-K<sup>12</sup>-R<sup>26</sup>-R<sup>8</sup>-K<sup>11</sup>-K<sup>108</sup> in  $\beta$ arr2 (Figures 5A and 5B). Similar to  $\beta$ arr1 structures, the other phosphate groups present in the phosphopeptides are either not involved in direct contact with Arg/Lys or positioned outside the binding groove. The N- and C-domains of  $\beta$ arr2 exhibit an inter-domain rotation of approximately 25° when compared with the basal state of  $\beta$ arr2 in both the structures, which is relatively higher from that observed in phosphopeptide-bound  $\beta$ arr1 (Figure 5C). Moreover, the three major loops in  $\beta$ arr2 namely, the finger, middle, and lariat loop also exhibit significant reorientation upon binding of V2Rpp and C5aR1pp compared with the basal state, although their positioning is almost identical between the two structures (Figure 5D). Finally, the three-element interaction and polar-core network in  $\beta$ arr2 are also disrupted upon binding to V2Rpp and C5aR1pp when compared with the basal state structure through the displacement of the carboxyl terminus of  $\beta$ arr2 from the N-domain and repositioning of the lariat loop through the interaction of K<sup>295</sup> with a phosphate moiety (Figures 5E and 5F).

### Identification of a key phosphorylation motif driving $\beta$ arr activation

As mentioned earlier, the analysis of these structural snapshots in terms of phosphorylation sites revealed a P-X-P-P type pattern with nearly identical interactions with analogous residues in  $\beta$ arr1 and 2 (Figures 6A–6C). Therefore, we analyzed the primary sequence of all non-olfactory and non-orphan GPCRs in their carboxyl terminus and ICL3 to identify the occurrence of the P-X-P-P pattern in these receptors (Table S1). We observed that a large set of GPCRs harbored this motif in their carboxyl terminus sequence and several receptors also included it in their ICL3 (Figures 6F and 6G). In order to validate the functional contribution of this motif in GPCR-induced  $\beta$ arr activation, we employed Ib30-based conformational biosensor in cellular context using three different receptors, which possess P-X-P-P motif either in their C terminus (CXCR3), ICL3 (M2R), or lack it (CXCR7). At the level of  $\beta$ arr1 conformation, the Ib30 sensor reports the degree (>15°) of inter-domain rotation as a proxy of  $\beta$ arr activation upon its interaction with activated and phosphorylated receptors.<sup>40,41–43</sup> In agreement with our hypothesis, we observed robust reactivity of Ib30 sensor with  $\beta$ arr1 for CXCR3 and M2R but not for CXCR7 although CXCR7 is capable of recruiting  $\beta$ arr1 upon agonist stimulation<sup>44–46</sup> (Figure 7A). We note however that the lack of Ib30 reactivity for CXCR7 does not indicate the absence of  $\beta$ arr1 activation but instead, a different active conformation that is not recognized by the Ib30 sensor. Surface expression of the receptors was confirmed in

(D) A side-by-side comparison of agonist-induced  $\beta$ arr2<sup>WT</sup> and  $\beta$ arr2<sup>DM</sup> recruitment to V2R (top) and C5aR1 (bottom) in the NanoBIT assay (Receptor-SmBIT+LgBIT- $\beta$ arr2) (mean  $\pm$  SEM; n = 4; normalized as fold over basal).

(E) A side-by-side comparison of  $\beta$ arr2<sup>WT</sup> and  $\beta$ arr2<sup>DM</sup> endosomal trafficking in response to agonist (arginine vasopressin peptide [AVP] for V2R and C5aR1) in the NanoBIT assay (Receptor+SmBIT- $\beta$ arr2+LgBIT-FYVE) (mean  $\pm$  SEM; n = 5; normalized as fold over basal). See also Figure S5.



**Figure 4. Structures of V2Rpp/C5aR1pp-βarr2 complexes**

(A and B) Overall cryo-EM structures of V2Rpp-βarr2-Fab30 (left) and C5aR1pp-βarr2-Fab30 complexes (right), respectively, in a trimeric assembly with βarr2 and Fab30 molecules colored as individual units. Front and side views of the trimeric complex EM maps have been shown with βarr2 molecules in blue, olive green,

(legend continued on next page)

these assays using the whole-cell ELISA assay (Figure S6B). To further corroborate these findings, we used two different receptors namely, the Bradykinin receptor subtype 2 (B2R) and C5aR1 for structure-based targeted mutagenesis to probe gain of function and loss of function, respectively, in terms of Ib30 reactivity pattern. As presented in Figure 7B, the activation of the wild-type B2R fails to induce an interaction between  $\beta$ arr1 and Ib30 as it lacks a P-X-P-P motif in its C terminus, although B2R is capable of recruiting  $\beta$ arrs.<sup>43,47</sup> Similar to CXCR7, the lack of Ib30 reactivity for B2R likely suggests a distinct activated conformation in  $\beta$ arr1 that is not efficiently recognized by Ib30 sensor. Interestingly however, reconstitution of the P-X-P-P motif in B2R by double mutation ( $\Delta G^{368} + L^{370}T$ ) results in robust Ib30 reactivity upon agonist stimulation (Figure 7B). Along the same lines, a mutant version of C5aR1pp, where the P-X-P-P motif is disrupted by the insertion of an additional arginine residue between pT<sup>336</sup> and pS<sup>338</sup>, completely loses the ability to induce a conformation that is recognizable by Fab30 (Figure 7C). Moreover, the corresponding mutation in C5aR1 also leads to a dramatic loss of Ib30 reactivity in cellular context (Figure 7C). Finally, we also disrupted a P-X-P-P motif present in the ICL3 of M2R by site-directed mutagenesis and measured the reactivity of Ib30 upon agonist stimulation (Figure 7D). We observed that similar to C5aR1, the disruption of the P-X-P-P motif in M2R also results in a marked decrease in Ib30 reactivity (Figure 7D). Taken together these data establish that the P-X-P-P phosphorylation pattern, when present in GPCRs, serves as a critical determinant for interaction and activation of  $\beta$ arrs (Figure 7E). In these experiments, the wild-type and mutant receptors were expressed at comparable levels as assessed in the surface expression assay (Figures S6C–S6E).

## DISCUSSION

Understanding the contribution of specific phosphorylation patterns in GPCRs for  $\beta$ arr recruitment, activation, and functional outcomes has been a key focus area in the field of GPCR biology, and several previous studies have provided interesting insights into this. For example, the bar-code hypothesis suggests differential contribution of distinct GPCR phosphorylation patterns in inducing specific conformations in  $\beta$ arrs that are linked to corresponding functional outcomes.<sup>28,50</sup> Taking lead from the rhodopsin-arrestin fusion protein structure determined by X-ray free electron laser (XFEL), a previous study identified and validated PXPXP and PXXPPX type phosphorylation codes in GPCRs that appear to be critical for  $\beta$ arr recruitment.<sup>51,52</sup> Subsequently, another study used peptide array and biophysical approaches to identify and propose a framework of the “key,” “inhibitory,” and “modulatory” phosphorylation sites in GPCRs that influence  $\beta$ arr binding, activation, and global conformation.<sup>53</sup> Additionally, we have also previously explored the func-

tional contribution of phosphorylation codes in several GPCRs to establish a link between different codes and ERK1/2 mitogen-activated protein (MAP) kinase activation using biochemical and cellular experiments.<sup>43</sup> This study now visualizes the molecular interaction of multiple GPCR phosphorylation patterns with  $\beta$ arrs and identifies a P-X-P-P motif that plays a crucial role in  $\beta$ arr binding and activation when present in the interacting receptor.

The structural snapshots determined here reveal that the P-X-P-P phosphorylation pattern simultaneously engages multiple elements in  $\beta$ arrs that are typically responsible for maintaining the inactive conformation in order to facilitate  $\beta$ arr activation. It is intriguing that the P-X-P-P phosphorylation pattern from different GPCR phospho-peptides utilizes a conserved set of interactions and docking interface on both isoforms of  $\beta$ arrs (Figures 6D and 6E). This conserved binding interface and corresponding interactions ensure the displacement of  $\beta$ arrs' C terminus from the N-domain and repositioning of theariat loop, leading to the release of the two major “breaks” on  $\beta$ arr activation namely, the three-element interaction and the polar-core network<sup>10,11,54</sup> (Figure S6A). It is important to note that some GPCRs tested here such as B2R and CXCR7, which lack the P-X-P-P motif, do not exhibit Ib30 reactivity although they are capable of recruiting  $\beta$ arrs in functionally competent conformation as reported previously.<sup>43–45,47</sup> Taken together, these data indicate that although P-X-P-P motif is not always essential for  $\beta$ arr binding, however, when present in GPCRs, it contributes critically in  $\beta$ arr interaction, and imparts an active conformation on  $\beta$ arrs through a structurally conserved mechanism. On the other hand, in the absence of the P-X-P-P motif, the receptors likely induce a distinct active conformation in  $\beta$ arrs that is also capable of directing functional responses. These scenarios underscore the structural diversity engrained in the GPCR- $\beta$ arr system, which orchestrates their functional diversity and versatility. The P-X-P-P motif, when present in GPCRs, appears to be sufficient to activate  $\beta$ arrs as measured using Ib30 sensor; however, additional phosphorylation sites may also contribute to improve the affinity of GPCR- $\beta$ arr interaction and exert modulatory activities as proposed earlier.<sup>52,53</sup> Moreover, several GPCRs harbor multiple copies of P-X-P-P motif in their carboxyl terminus and ICL3 (Table S2), similar to PXPXP and PXXPPX codes as reported earlier.<sup>52</sup> While it is tempting to speculate that different P-X-P-P motif may be utilized in a context-dependent fashion such as cell type-specific, GRK-specific, and ligand-specific manner, further studies are essential to systematically probe these interesting possibilities.

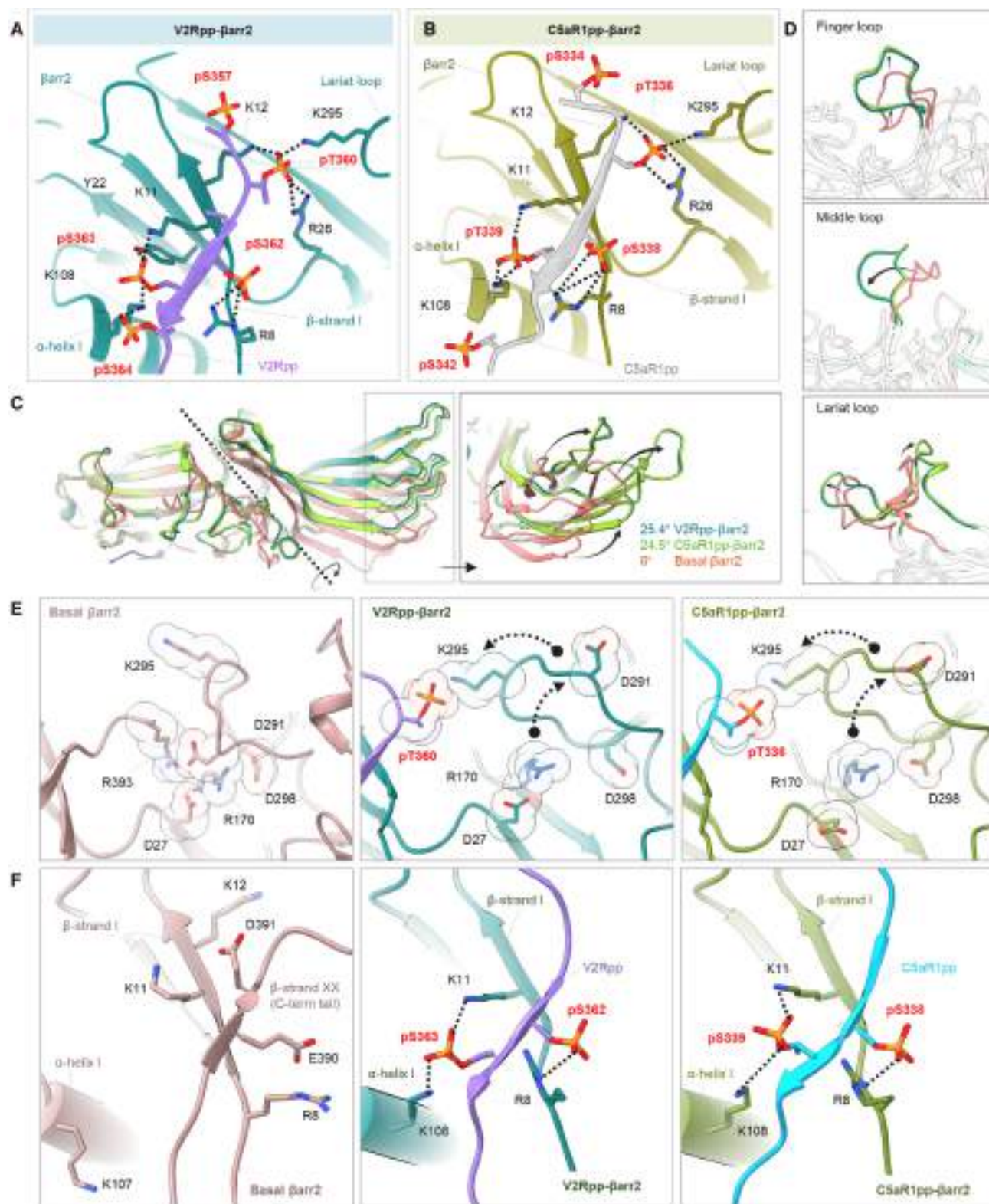
An intriguing observation in these structural snapshots is novel dimer and trimer assemblies of  $\beta$ arrs. Although  $\beta$ arrs have a strong propensity to adopt different oligomeric states,<sup>55–57</sup> the dimer and trimer interfaces observed here differ significantly from previously reported interfaces<sup>34,56,57</sup> (Figure S7). The

and purple; and Fab30 molecules in beige, red, and gray, respectively. The constant domain densities of Fab30 were removed during local refinement and not observed in the final cryo-EM maps.

(C and D) Overall trimeric arrangement of V2Rpp/C5aR1pp- $\beta$ arr2 complexes in the cryo-EM structures shown here as cartoon representation (left). Domain organization of the  $\beta$ arr2 molecules in trimeric assembly without Fab30 shown as cartoon representation (right).

(E and F) Structure of individual  $\beta$ arr2 protomers in V2Rpp/C5aR1pp- $\beta$ arr2 complexes showing the binding of phosphopeptides on the N-domain of  $\beta$ arr2. Coulombic densities of corresponding phosphopeptides have been provided in insets. See also Figures S3 and S5 and Table 1.





**Figure 5. Active conformations of phosphopeptide-bound βarr2**

(A and B) Extensive charge-charge interactions between the phosphate residues in V2Rpp/C5aR1pp with Lys/Arg in the N-domain (represented as black dotted lines) stabilize the V2Rpp and C5aR1pp into the N-domain groove of βarr2.

(legend continued on next page)

overall buried surface area in dimer and trimer assemblies are approximately 1,500 and 4,500 Å<sup>2</sup>, respectively, suggesting a robust and stable oligomeric arrangement. The two protomers in C5aR1pp- and CXCR4pp-bound  $\beta$ arr1 interface with each other through multiple hydrogen bonds, salt bridges, and non-bonded contacts in a manner where the C-edge loop residues of one protomer are positioned into the central crest of the other protomer in the proximity of the finger loop. An analogous set of interactions are also involved in the trimer arrangement of  $\beta$ arr2 in complex with V2Rpp and C5aR1pp. Interestingly, a previous crystal structure of  $\beta$ arr2 in complex with IP6 also shows a trimeric arrangement although the trimer interface is different from that observed here in phosphopeptide-bound conformations.<sup>34</sup> A comprehensive map of dimer and trimer interface with residue-level contacts is listed in Table S2. Considering the functional multiplicity of  $\beta$ arrs in terms of distinct signaling and regulatory outcomes and receptor-specific responses,<sup>58</sup> it is possible that distinct oligomeric interfaces in  $\beta$ arrs may be a modular mechanism to fine-tune the functional contributions by providing distinct possibilities for adaptable protein-protein interaction interfaces for binding partners. Nonetheless, the biological implications of the oligomeric assemblies observed here for activated  $\beta$ arrs remain to be explored further in future studies. The comparison of V2Rpp- and C5aR1pp-bound  $\beta$ arr1 and 2 reveal a significantly higher inter-domain rotation in  $\beta$ arr2 compared with  $\beta$ arr1 as hypothesized earlier based on cellular and biochemical studies,<sup>40</sup> and this may provide a plausible explanation for functional differences between the  $\beta$ arr isoforms as observed for multiple GPCRs.<sup>30</sup> However, structural visualization of GPCR- $\beta$ arr2 complexes would be required to validate this possibility.

In summary, guided by the structural snapshots, we identify and experimentally validate a P-X-P-P motif in GPCRs that imparts an active conformation on  $\beta$ arrs by engaging a conserved network of interacting residues in their N-domain. Our study therefore provides an important missing link in our current conceptual framework of GPCR-mediated  $\beta$ arr activation and paves the way to decipher the structural and functional diversity encoded in GPCR signaling and regulatory paradigms. We also note that a companion manuscript presenting the crystal structures of  $\beta$ arr1 in complex with phosphopeptides derived from the chemokine receptor CCR5 converges to similar conclusions about a critical contribution of the P-X-P-P motif in GPCRs on  $\beta$ arr interaction and activation.<sup>59</sup>

### Limitations of the study

The structural snapshots presented here involve isolated phosphopeptides with defined phosphorylation patterns without the

transmembrane core of the receptors. As the interaction of receptor transmembrane core is also important for generating fully engaged GPCR- $\beta$ arr complexes,<sup>60–63</sup> it is possible that  $\beta$ arrs exhibit additional conformational changes in complex with activated and phosphorylated receptors. However, the ability of the P-X-P-P motif, when present in GPCRs, to engage the K-K-R-R-K-K pattern in  $\beta$ arrs is likely to be maintained and guide  $\beta$ arr activation even in the context of full-length receptors. Additionally, we cannot rule out the possibility that for some receptors, a functional P-X-P-P motif may not be generated in cellular context despite having a suitable primary sequence because all the phosphorylatable residues may not undergo efficient phosphorylation.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human cell lines
  - Insect cells
- METHOD DETAILS
  - General reagents, plasmids, and cell culture
  - Expression and purification of  $\beta$ arrs
  - Expression and purification of Fabs
  - Co-immunoprecipitation assay
  - Reconstitution of phosphopeptide- $\beta$ arr-Fab complexes
  - Negative-staining EM
  - Cryo-EM sample preparation and data acquisition
  - Cryo-EM data processing and model building
  - Model building and refinement
  - NanoBiT assay for  $\beta$ arr2<sup>WT</sup> and  $\beta$ arr2<sup>DM</sup> recruitment
  - NanoBiT assay for  $\beta$ arr trafficking
  - NanoBiT assay for Ib30 reactivity
  - Receptor surface expression
- QUANTIFICATION AND STATISTICAL ANALYSIS

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.molcel.2023.04.025>.

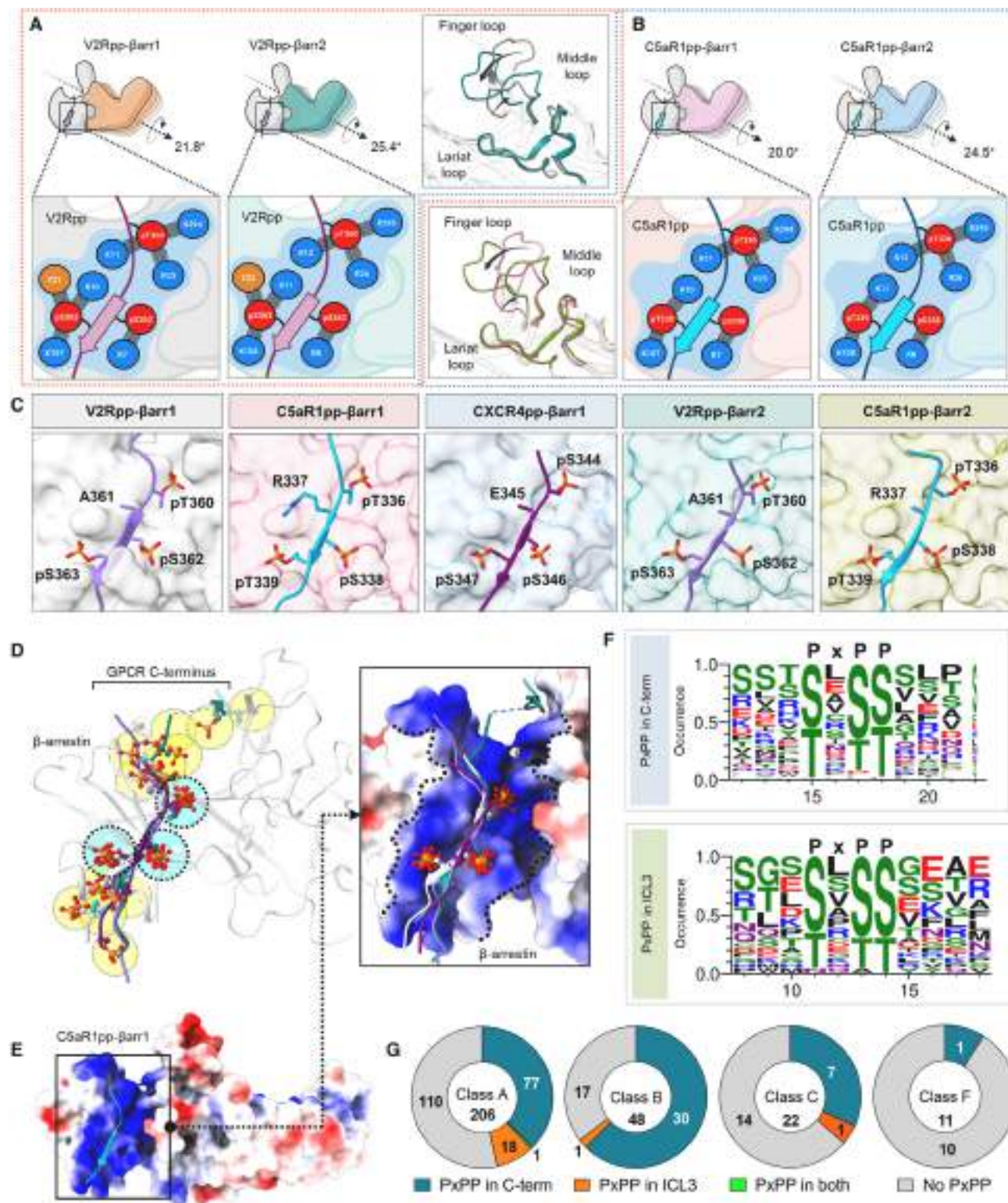
(C) V2Rpp (dark green) and C5aR1pp (light green) activated  $\beta$ arr2 structures were superimposed with the basal state of  $\beta$ arr2 (PDB: 3P2D, orange), and inter-domain rotations were calculated.

(D) Conformational changes observed in the finger (top), middle (middle), and lariat loops (bottom) in the activated  $\beta$ arr2 compared with the basal state crystal structure of  $\beta$ arr2.

(E) Polar-core environment in basal  $\beta$ arr2 (PDB: 3P2D, left) and disruption of polar-core interactions upon binding of V2Rpp (middle) and C5aR1pp to  $\beta$ arr2 (right).

(F) Three-element interaction network consisting of  $\beta$ arr2 C-terminal  $\beta$ -strand XX,  $\alpha$ -helix I, and  $\beta$ -strand I in the basal state of  $\beta$ arr2 (left). Binding of the phosphopeptides V2Rpp and C5aR1pp to  $\beta$ arr2 results in the displacement of the  $\beta$ -strand XX, and engages the phosphopeptide V2Rpp (middle) and C5aR1pp (right) into the N-domain groove of  $\beta$ arr2 through hydrogen bonds and polar interactions. See also Figure S6.





**Figure 6. Identification of a key phosphorylation motif in GPCRs driving βarr activation**

(A) Comparison of the V2Rpp-bound βarr1 and βarr2 structures reveals similar interactions of V2Rpp with both isoforms of βarrs although a slightly higher inter-domain rotation is observed in βarr2 (left). A schematic representation of the interface network between negatively charged phospho-residues (red) and positively

(legend continued on next page)

## ACKNOWLEDGMENTS

Research in A.K.S.'s laboratory is supported by the Senior Fellowship of the DBT Wellcome Trust India Alliance (IA/S/20/1/504916) awarded to A.K.S., Science and Engineering Research Board (SPR/2020/000408 and IPA/2020/000405), Council of Scientific and Industrial Research (37(1730)/19/EMR-II), Indian Council of Medical Research (F.NO.52/15/2020/BIO/BMS), the Young Scientist Award from Lady Tata Memorial Trust, and IIT Kanpur. We thank A. Ranjan, M. Chaturvedi, and H. Dwivedi-Agnihotri for their help with the characterization of the phosphopeptides; M. Ganguly for assisting with GPCR sequence analysis; E. Ghosh for initial characterization of  $\beta$ arr2<sup>DM</sup>; and A. Dalal and N. Zaidi for helping with the functional assays on M2R. Cryo-EM was performed at the BioEM lab of the Biozentrum at the University of Basel, and we thank Carola Alampi and David Kalbermatter for their excellent technical assistance.

## AUTHOR CONTRIBUTIONS

J.M. and M.K.Y. prepared and characterized the  $\beta$ arr complexes. J.M. performed negative-staining EM with R.B. and processed the cryo-EM data with R.B. P.S. carried out all the functional assays related to  $\beta$ arr2<sup>DM</sup> characterization and Ib30 sensor assay. M.K.Y. purified  $\beta$ arrs and carried out the colP experiments with V.S. and Sayantan Saha. Shirsha Saha contributed to the functional characterization of  $\beta$ arr2<sup>DM</sup>. M.C. screened the samples and collected cryo-EM data. A.K.S. supervised and managed the overall project. All authors contributed to data analysis, interpretation, and manuscript writing.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: October 10, 2022

Revised: February 22, 2023

Accepted: April 26, 2023

Published: May 19, 2023

## REFERENCES

- Pierce, K.L., Premont, R.T., and Lefkowitz, R.J. (2002). Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* 3, 639–650. <https://doi.org/10.1038/nrm908>.
- Reiter, E., Ahn, S., Shukla, A.K., and Lefkowitz, R.J. (2012). Molecular mechanism of beta-arrestin-biased agonism at seven-transmembrane receptors. *Annu. Rev. Pharmacol. Toxicol.* 52, 179–197. <https://doi.org/10.1146/annurev.pharmtox.010909.105800>.
- Shenoy, S.K., and Lefkowitz, R.J. (2005). Seven-transmembrane receptor signaling through beta-arrestin. *Sci. STKE* 2005, cm10. <https://doi.org/10.1126/stke.2005/308/cm10>.
- Pierce, K.L., and Lefkowitz, R.J. (2001). Classical and new roles of beta-arrestins in the regulation of G-protein-coupled receptors. *Nat. Rev. Neurosci.* 2, 727–733. <https://doi.org/10.1038/35094577>.
- Maharana, J., Banerjee, R., Yadav, M.K., Sarma, P., and Shukla, A.K. (2022). Emerging structural insights into GPCR-beta-arrestin interaction and functional outcomes. *Curr. Opin. Struct. Biol.* 75, 102406. <https://doi.org/10.1016/j.sbi.2022.102406>.
- Ranjan, R., Dwivedi, H., Baidya, M., Kumar, M., and Shukla, A.K. (2017). Novel structural insights into GPCR-beta-arrestin interaction and signaling. *Trends Cell Biol.* 27, 851–862. <https://doi.org/10.1016/j.tcb.2017.05.008>.
- Ahn, S., Shenoy, S.K., Luttrell, L.M., and Lefkowitz, R.J. (2020). SnapShot: beta-arrestin functions. *Cell* 182, 1362–1362.e1. <https://doi.org/10.1016/j.cell.2020.07.034>.
- Shenoy, S.K., and Lefkowitz, R.J. (2011). beta-arrestin-mediated receptor trafficking and signal transduction. *Trends Pharmacol. Sci.* 32, 521–533. <https://doi.org/10.1016/j.tips.2011.05.002>.
- Kang, D.S., Tian, X., and Benovic, J.L. (2014). Role of beta-arrestins and arrestin domain-containing proteins in G protein-coupled receptor trafficking. *Curr. Opin. Cell Biol.* 27, 63–71. <https://doi.org/10.1016/j.ccb.2013.11.005>.
- Seyedabadi, M., Gharghabi, M., Gurevich, E.V., and Gurevich, V.V. (2021). Receptor-arrestin interactions: the GPCR perspective. *Biomolecules* 11, 218. <https://doi.org/10.3390/biom11020218>.
- Gurevich, V.V., and Gurevich, E.V. (2019). The structural basis of the arrestin binding to GPCRs. *Mol. Cell. Endocrinol.* 484, 34–41. <https://doi.org/10.1016/j.mce.2019.01.019>.
- Gurevich, V.V., and Gurevich, E.V. (2014). Overview of different mechanisms of arrestin-mediated signaling. *Curr. Protoc. Pharmacol.* 67, 2.10.11–2.10.19. <https://doi.org/10.1002/0471141755.ph0210s67>.
- Shiraishi, Y., Kofuku, Y., Ueda, T., Pandey, S., Dwivedi-Agnihotri, H., Shukla, A.K., and Shimada, I. (2021). Biphasic activation of beta-arrestin 1 upon interaction with a GPCR revealed by methyl-TROSY NMR. *Nat. Commun.* 12, 7158. <https://doi.org/10.1038/s41467-021-27482-3>.
- Shukla, A.K., Westfield, G.H., Xiao, K., Reis, R.I., Huang, L.Y., Tripathi-Shukla, P., Qian, J., Li, S., Blanc, A., Oleskie, A.N., et al. (2014). Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature* 512, 218–222. <https://doi.org/10.1038/nature13430>.
- Bous, J., Fouillen, A., Orcel, H., Trapani, S., Cong, X., Fontanel, S., Saint-Paul, J., Lai-Kee-Him, J., Urbach, S., Sibille, N., et al. (2022). Structure of the vasopressin hormone-V2 receptor-beta-arrestin1 ternary complex. *Sci. Adv.* 8, eabo7761. <https://doi.org/10.1126/sciadv.abo7761>.

charged residues (blue) of  $\beta$ arrs is shown (below, zoomed-in box). Although the lariat loops of the two structures align well, significant deviations can be observed for the finger and middle loops (right, inset box).

(B) Comparative analysis of C5aR1pp-bound  $\beta$ arr1 and  $\beta$ arr2 structures uncover similar interactions of C5aR1pp with both  $\beta$ arr isoforms, but again, a higher inter-domain rotation is observed for  $\beta$ arr2. A similar representation of the interface between negatively charged phospho-residues (red) and positively charged residues (blue) of  $\beta$ arrs is shown (below, zoomed-in box).

(C) In all the structures of phosphopeptide-bound  $\beta$ arrs, a conserved motif can be observed with respect to three phospho-residues, referred to as the P-X-P-P motif, where “P” is a phospho-Ser/Thr and “X” can be any other residue.

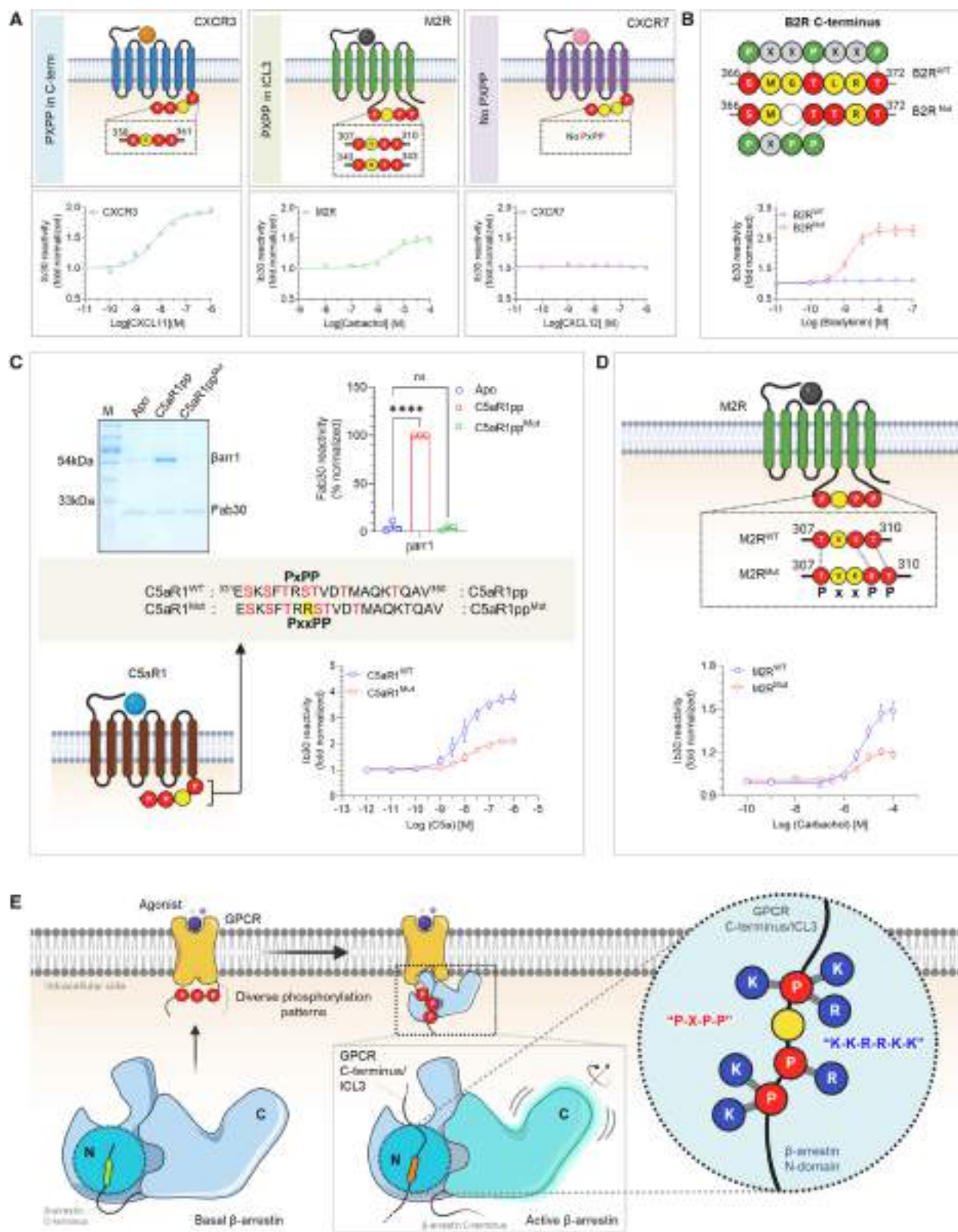
(D) Superposition of V2Rpp- $\beta$ arr1 (PDB 4JQI), V2Rpp- $\beta$ arr2, C5aR1pp- $\beta$ arr1, CXCR4pp- $\beta$ arr1, and C5aR1pp- $\beta$ arr2 shows conservation of phosphates corresponding to P-X-P-P position, whereas other phosphates are distributed throughout the phosphopeptides.

(E) Superposition of phosphopeptides on C5aR1pp- $\beta$ arr1 reveals the conserved phospho-residues on positively charged cleft present on  $\beta$ arrs' N-domain.  $\beta$ arr is shown as Coulombic-charged surface here.

(F) A sequence alignment of the C-terminal tail and ICL3 residues of non-olfactory and non-orphan class A receptors reveal the consensus sequence, “P-X-P-P” required for activation of  $\beta$ arrs. The consensus sequence logo was generated with the WEBLOGO tool<sup>48</sup> and sequence alignment was performed with Kalign.<sup>49</sup> A stretch of 11 amino acid residues has been shown for better representation.

(G) Proportions of GPCRs of class A, B, C, and F having P-X-P-P motif in C terminus or ICL3 have been represented as pie charts. See also Figure S6 and Table S1.





(legend on next page)

16. Lee, Y., Warne, T., Nehmé, R., Pandey, S., Dwivedi-Agnihotri, H., Chaturvedi, M., Edwards, P.C., García-Nafria, J., Leslie, A.G.W., Shukla, A.K., and Tate, C.G. (2020). Molecular basis of beta-arrestin coupling to formoterol-bound beta(1)-adrenoceptor. *Nature* 583, 862–866. <https://doi.org/10.1038/s41586-020-2419-1>.
17. 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. <https://doi.org/10.1038/s41586-020-1953-1>.
18. 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. <https://doi.org/10.1038/s41586-020-1954-0>.
19. Yin, W., Li, Z., Jin, M., Yin, Y.L., de Waal, P.W., Pal, K., Yin, Y., Gao, X., He, Y., Gao, J., et al. (2019). A complex structure of arrestin-2 bound to a G protein-coupled receptor. *Cell Res.* 29, 971–983. <https://doi.org/10.1038/s41422-019-0256-2>.
20. Cao, C., Barros-Álvarez, X., Zhang, S., Kim, K., Dämgen, M.A., Panova, O., Suomivuori, C.M., Fay, J.F., Zhong, X., Krumm, B.E., et al. (2022). Signaling snapshots of a serotonin receptor activated by the prototypical psychedelic LSD. *Neuron* 110, 3154–3167.e7. <https://doi.org/10.1016/j.neuron.2022.08.006>.
21. Latorraca, N.R., Wang, J.K., Bauer, B., Townshend, R.J.L., Hollingsworth, S.A., Olivieri, J.E., Xu, H.E., Sommer, M.E., and Dror, R.O. (2018). Molecular mechanism of GPCR-mediated arrestin activation. *Nature* 557, 452–456. <https://doi.org/10.1038/s41586-018-0077-3>.
22. Eichel, K., Jullié, D., Barsi-Rhine, B., Latorraca, N.R., Masureel, M., Sibarita, J.B., Dror, R.O., and von Zastrow, M. (2018). Catalytic activation of beta-arrestin by GPCRs. *Nature* 557, 381–386. <https://doi.org/10.1038/s41586-018-0079-1>.
23. Reiter, E., and Lefkowitz, R.J. (2006). GRKs and beta-arrestins: roles in receptor silencing, trafficking and signaling. *Trends Endocrinol. Metab.* 17, 159–165. <https://doi.org/10.1016/j.tem.2006.03.008>.
24. Prihandoko, R., Bradley, S.J., Tobin, A.B., and Butcher, A.J. (2015). Determination of GPCR phosphorylation status: establishing a phosphorylation barcode. *Curr. Protoc. Pharmacol.* 69, 2.13.1–2.13.26. <https://doi.org/10.1002/0471141755.ph0213s69>.
25. Tobin, A.B. (2008). G-protein-coupled receptor phosphorylation: where, when and by whom. *Br. J. Pharmacol.* 153, S167–S176. <https://doi.org/10.1038/sj.bjp.0707662>.
26. Tobin, A.B., Butcher, A.J., and Kong, K.C. (2008). Location, location, location...site-specific GPCR phosphorylation offers a mechanism for cell-type-specific signalling. *Trends Pharmacol. Sci.* 29, 413–420. <https://doi.org/10.1016/j.tips.2008.05.006>.
27. Yang, Z., Yang, F., Zhang, D., Liu, Z., Lin, A., Liu, C., Xiao, P., Yu, X., and Sun, J.P. (2017). Phosphorylation of G protein-coupled receptors: from the barcode hypothesis to the flute model. *Mol. Pharmacol.* 92, 201–210. <https://doi.org/10.1124/mol.116.107839>.
28. Nobles, K.N., Xiao, K., Ahn, S., Shukla, A.K., Lam, C.M., Rajagopal, S., Strachan, R.T., Huang, T.Y., Bressler, E.A., Hara, M.R., et al. (2011). Distinct phosphorylation sites on the beta(2)-adrenergic receptor establish a barcode that encodes differential functions of beta-arrestin. *Sci. Signal.* 4, ra51. <https://doi.org/10.1126/scisignal.2001707>.
29. Chen, Q., and Tesmer, J.J.G. (2022). G protein-coupled receptor interactions with arrestins and GPCR kinases: the unresolved issue of signal bias. *J. Biol. Chem.* 298, 102279. <https://doi.org/10.1016/j.jbc.2022.102279>.
30. Srivastava, A., Gupta, B., Gupta, C., and Shukla, A.K. (2015). Emerging functional divergence of beta-arrestin isoforms in GPCR function. *Trends Endocrinol. Metab.* 26, 628–642. <https://doi.org/10.1016/j.tem.2015.09.001>.
31. Shukla, A.K., Manglik, A., Kruse, A.C., Xiao, K., Reis, R.I., Tseng, W.C., Staus, D.P., Hilger, D., Uysal, S., Huang, L.Y., et al. (2013). Structure of active beta-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. *Nature* 497, 137–141. <https://doi.org/10.1038/nature12120>.
32. He, Q.T., Xiao, P., Huang, S.M., Jia, Y.L., Zhu, Z.L., Lin, J.Y., Yang, F., Tao, X.N., Zhao, R.J., Gao, F.Y., et al. (2021). Structural studies of phosphorylation-dependent interactions between the V2R receptor and arrestin-2. *Nat. Commun.* 12, 2396. <https://doi.org/10.1038/s41467-021-22731-x>.
33. Min, K., Yoon, H.J., Park, J.Y., Baidya, M., Dwivedi-Agnihotri, H., Maharana, J., Chaturvedi, M., Chung, K.Y., Shukla, A.K., and Lee, H.H. (2020). Crystal structure of beta-arrestin 2 in complex with CXCR7 phosphopeptide. *Structure* 28, 1014–1023.e4. <https://doi.org/10.1016/j.str.2020.06.002>.
34. Chen, Q., Perry, N.A., Vishnivetskiy, S.A., Berndt, S., Gilbert, N.C., Zhuo, Y., Singh, P.K., Tholen, J., Ohi, M.D., Gurevich, E.V., et al. (2017). Structural basis of arrestin-3 activation and signaling. *Nat. Commun.* 8, 1427. <https://doi.org/10.1038/s41467-017-01218-8>.
35. Han, M., Gurevich, V.V., Vishnivetskiy, S.A., Sigler, P.B., and Schubert, C. (2001). Crystal structure of beta-arrestin at 1.9 Å: possible mechanism of receptor binding and membrane Translocation. *Structure* 9, 869–880. [https://doi.org/10.1016/S0969-2126\(01\)00644-X](https://doi.org/10.1016/S0969-2126(01)00644-X).
36. Zhan, X., Gimenez, L.E., Gurevich, V.V., and Spiller, B.W. (2011). Crystal structure of arrestin-3 reveals the basis of the difference in receptor binding between two non-visual subtypes. *J. Mol. Biol.* 406, 467–478. <https://doi.org/10.1016/j.jmb.2010.12.034>.
37. Milano, S.K., Pace, H.C., Kim, Y.M., Brenner, C., and Benovic, J.L. (2002). Scaffolding functions of arrestin-2 revealed by crystal structure and mutagenesis. *Biochemistry* 41, 3321–3328. <https://doi.org/10.1021/bi015905j>.
38. Kang, D.S., Kern, R.C., Puthenveedu, M.A., von Zastrow, M., Williams, J.C., and Benovic, J.L. (2009). Structure of an arrestin2-clathrin complex reveals a novel clathrin binding domain that modulates receptor trafficking. *J. Biol. Chem.* 284, 29860–29872. <https://doi.org/10.1074/jbc.M109.023366>.
39. Ghosh, E., Srivastava, A., Baidya, M., Kumari, P., Dwivedi, H., Nidhi, K., Ranjan, R., Dogra, S., Koide, A., Yadav, P.N., et al. (2017). A synthetic

### Figure 7. A key phosphorylation motif for $\beta$ arr activation

- (A) NanoBIT-based assay for assessing Ib30 reactivity to CXCR3 (left), M2R (middle), and CXCR7 (right) activated  $\beta$ arr1 (Receptor+SmBiT- $\beta$ arr1+LgBiT-Ib30) (mean  $\pm$  SEM; n = 3; normalized as fold over basal).
- (B) Deletion of G<sup>368</sup> and substitution of L<sup>370</sup> to Thr in B2R engineers the P-X-P-P (referred to as B2R<sup>Mut</sup>) and results in gain of function in terms of Ib30 reactivity as measured using the NanoBIT assay (Receptor+SmBiT- $\beta$ arr1+LgBiT-Ib30) (mean  $\pm$  SEM; n = 3; normalized as fold over basal).
- (C) Addition of an extra Arg between positions 336 and 337 in C5aR1pp to disrupt the P-X-P-P (referred to as C5aR1pp<sup>Mut</sup>) leads to a near-complete loss of Fab30 (top) reactivity as measured in colP assay (mean  $\pm$  SEM; n = 3; densitometry-based data normalized with respect to C5aR1pp signal as 100%; one-way ANOVA, Dunnett's multiple comparisons test). The exact p values are as follows: Apo vs. C5aR1pp2 (p < 0.0001), Apo vs. C5aR1pp4 (p = 0.9160) (\*\*\*\*p < 0.0001, ns, non-significant). Corresponding mutation in C5aR1 to disrupt the P-X-P-P (referred to as C5aR1<sup>Mut</sup>) results in a dramatic decrease in Ib30 reactivity (bottom) as measured using the NanoBIT assay (mean  $\pm$  SEM; n = 5; normalized as fold over basal).
- (D) Insertion of a Lys residue between positions 308 and 309 to disrupt the P-X-P-P pattern in ICL3 of M2R (referred to as M2R<sup>Mut</sup>) shows reduced Ib30 reactivity compared with the wild type (M2R<sup>WT</sup>) as measured using the NanoBIT assay (mean  $\pm$  SEM; n = 4; normalized as fold over basal).
- (E) Schematic representation summarizing the identification of a phosphorylation motif in GPCRs that drives  $\beta$ arr activation. See also Figure S6.

**Table 1. Cryo-EM data collection, processing, and refinement statistics, related to Figures 1, 4, S2, and S5**

	C5aR1pp- βarr1-Fab30	C5aR1pp-βarr1- Fab30 (local refined)	V2Rpp- βarr2-Fab30	V2Rpp-βarr2-Fab30 (local refined)	C5aR1pp- βarr2-Fab30	C5aR1pp-βarr2- Fab30 (local refined)	CXCR4pp- βarr1-Fab30	CXCR4pp-βarr1- Fab30 (local refined)
Code	PDB: 8GO8 EMD 34173	PDB: 8I0N EMD 35104	PDB: 8GOC EMD 34175	PDB: 8I10 EMD 35115	PDB: 8GOO EMD 34178	PDB: 8I0Z EMD 35114	PDB: 8GP3 EMD 34188	PDB: 8I0Q EMD 35106
Microscope	Titan Krios	Titan Krios	Titan Krios	Titan Krios	Glacios	Glacios	Glacios	Glacios
Camera	GIF/K2	GIF/K2	GIF/K2	GIF/K2	Gatan K3	Gatan K3	Gatan K3	Gatan K3
Magnification	165,000	165,000	165,000	165,000	46,000	46,000	46,000	46,000
Voltage (kV)	300	300	300	300	200	200	200	200
Defocus range (μm)	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5	0.5–2.5
Exposure time (s)	5	5	4	4	3	3	2.9	2.9
Total dose (e <sup>-</sup> /Å <sup>2</sup> )	49	49	48.7	48.7	51	51	49.38	49.38
Number of frames	40	40	40	40	40	40	40	40
Pixel size (Å)	0.82	0.82	0.82	0.82	0.878	0.878	0.878	0.878
Micrographs (no.)	6,212	6,212	9,720	9,720	8,614	8,614	5,637	5,637
Initial particles (no.)	4,304,237	4,304,237	2,444,407	2,444,407	4,012,616	4,012,616	3,236,193	3,236,193
Symmetry imposed	C2	C2	C3	C3	C3	C3	C2	C2
Final particles (no.)	80,437	80,437	92,018	92,018	38,206	38,206	86,525	86,525
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution (Å)	3.41	3.26	4.18	3.96	4.41	4.33	4.81	4.45
<b>Refinement</b>								
Initial code (PDB)	4JQI	8GO8	5TV1	8GOC	8GOC	8GOO	8GO8	8GP3
Model resolution (Å)	3.5	3.4	4.7	4.1	4.7	4.7	5.3	5.7
FSC threshold	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
<b>Model composition</b>								
Non-hydrogen atoms	11,519	9,141	16,942	13,205	16,966	13,163	11,494	8,842
Protein residues	1,522	1,190	2,223	1,729	2,223	1,725	1,522	1,192
Ligand atoms	0	0	0	0	0	0	0	0
<b>RMSD</b>								
Bond length (Å)	0.004	0.005	0.003	0.007	0.004	0.006	0.005	0.004
Bond angle (°)	0.942	1.001	0.716	1.256	0.665	1.128	0.982	0.967
<b>Validation</b>								
Favored (%)	97.23	95.43	95.05	93.63	94.90	93.24	96.62	94.92
Allowed (%)	2.77	4.57	4.76	6.37	4.96	6.76	3.38	5.08
Disallowed (%)	0	0	0.18	0	0.14	0	0	0
MolProbity score	1.41	1.46	1.97	1.88	1.91	1.77	1.77	1.85
Clash score	5.23	3.49	13.07	8.60	10.83	6.12	10.81	9.44



- intrabody-based selective and generic inhibitor of GPCR endocytosis. *Nat. Nanotechnol.* 12, 1190–1198. <https://doi.org/10.1038/nnano.2017.188>.
40. Ghosh, E., Dwivedi, H., Baidya, M., Srivastava, A., Kumari, P., Stepniowski, T., Kim, H.R., Lee, M.H., van Gastel, J., Chaturvedi, M., et al. (2019). Conformational sensors and domain swapping reveal structural and functional differences between beta-arrestin isoforms. *Cell Rep.* 28, 3287–3299.e6. <https://doi.org/10.1016/j.celrep.2019.08.053>.
41. Baidya, M., Chaturvedi, M., Dwivedi-Agnihotri, H., Ranjan, A., Devost, D., Namkung, Y., Stepniowski, T.M., Pandey, S., Baruah, M., Panigrahi, B., et al. (2022). Allosteric modulation of GPCR-induced beta-arrestin trafficking and signaling by a synthetic intrabody. *Nat. Commun.* 13, 4634. <https://doi.org/10.1038/s41467-022-32386-x>.
42. Dwivedi-Agnihotri, H., Chaturvedi, M., Baidya, M., Stepniowski, T.M., Pandey, S., Maharana, J., Srivastava, A., Caengprasath, N., Hanyaloglu, A.C., Selent, J., and Shukla, A.K. (2020). Distinct phosphorylation sites in a prototypical GPCR differently orchestrate beta-arrestin interaction, trafficking, and signaling. *Sci. Adv.* 6, eabb8368. <https://doi.org/10.1126/sciadv.abb8368>.
43. Baidya, M., Kumari, P., Dwivedi-Agnihotri, H., Pandey, S., Chaturvedi, M., Stepniowski, T.M., Kawakami, K., Cao, Y., Laporte, S.A., Selent, J., et al. (2020). Key phosphorylation sites in GPCRs orchestrate the contribution of beta-arrestin in ERK1/2 activation. *EMBO Rep.* 21, e49886. <https://doi.org/10.15252/embr.201949886>.
44. Rajagopal, S., Kim, J., Ahn, S., Craig, S., Lam, C.M., Gerard, N.P., Gerard, C., and Lefkowitz, R.J. (2010). Beta-arrestin- but not G protein-mediated signaling by the "decoy" receptor CXCR7. *Proc. Natl. Acad. Sci. USA* 107, 628–632. <https://doi.org/10.1073/pnas.0912852107>.
45. Nguyen, H.T., Reyes-Alcaraz, A., Yong, H.J., Nguyen, L.P., Park, H.K., Inoue, A., Lee, C.S., Seong, J.Y., and Hwang, J.I. (2020). CXCR7: a beta-arrestin-biased receptor that potentiates cell migration and recruits beta-arrestin2 exclusively through Gbetagamma subunits and GRK2. *Cell Biosci.* 10, 134. <https://doi.org/10.1186/s13578-020-00497-x>.
46. Reyes-Alcaraz, A., Lee, Y.N., Yun, S., Hwang, J.I., and Seong, J.Y. (2018). Conformational signatures in beta-arrestin2 reveal natural biased agonism at a G-protein-coupled receptor. *Commun. Biol.* 1, 128. <https://doi.org/10.1038/s42003-018-0134-3>.
47. Zimmerman, B., Simaan, M., Akoume, M.Y., i, N., Chevallier, S., Séguéla, P., and Laporte, S.A. (2011). Role of ssarrestins in bradykinin B2 receptor-mediated signalling. *Cell. Signal.* 23, 648–659. <https://doi.org/10.1016/j.cellsig.2010.11.016>.
48. Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: A sequence logo generator. *Genome Res.* 14, 1188–1190. <https://doi.org/10.1101/gr.849004>.
49. Lassmann, T. (2019). Kalign 3: multiple sequence alignment of large data sets. *Bioinformatics* 36, 1928–1929. <https://doi.org/10.1093/bioinformatics/btz795>.
50. Shukla, A.K., Violin, J.D., Whalen, E.J., Gesty-Palmer, D., Shenoy, S.K., and Lefkowitz, R.J. (2008). Distinct conformational changes in beta-arrestin report biased agonism at seven-transmembrane receptors. *Proc. Natl. Acad. Sci. USA* 105, 9988–9993. <https://doi.org/10.1073/pnas.0804246105>.
51. Kang, Y., Zhou, X.E., Gao, X., He, Y., Liu, W., Ishchenko, A., Barty, A., White, T.A., Yefanov, O., Han, G.W., et al. (2015). Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* 523, 561–567. <https://doi.org/10.1038/nature14656>.
52. Zhou, X.E., He, Y., de Waal, P.W., Gao, X., Kang, Y., Van Eps, N., Yin, Y., Pal, K., Goswami, D., White, T.A., et al. (2017). Identification of phosphorylation codes for arrestin recruitment by G protein-coupled receptors. *Cell* 170, 457–469.e13. <https://doi.org/10.1016/j.cell.2017.07.002>.
53. Mayer, D., Damberger, F.F., Samarasinghareddy, M., Feldmueller, M., Vuckovic, Z., Flock, T., Bauer, B., Mutt, E., Zosel, F., Allain, F.H.T., et al. (2019). Distinct G protein-coupled receptor phosphorylation motifs modulate arrestin affinity and activation and global conformation. *Nat. Commun.* 10, 1261. <https://doi.org/10.1038/s41467-019-09204-y>.
54. Gurevich, V.V., and Gurevich, E.V. (2004). The molecular acrobatics of arrestin activation. *Trends Pharmacol. Sci.* 25, 105–111. <https://doi.org/10.1016/j.tips.2003.12.008>.
55. Gurevich, V.V., and Gurevich, E.V. (2022). Solo vs. Chorus: monomers and Oligomers of arrestin Proteins. *Int. J. Mol. Sci.* 23, 7253. <https://doi.org/10.3390/ijms23137253>.
56. Chen, Q., Zhuo, Y., Kim, M., Hanson, S.M., Francis, D.J., Vishnivetskiy, S.A., Altenbach, C., Klug, C.S., Hubbell, W.L., and Gurevich, V.V. (2014). Self-association of arrestin family members. *Handb. Exp. Pharmacol.* 219, 205–223. [https://doi.org/10.1007/978-3-642-41199-1\\_11](https://doi.org/10.1007/978-3-642-41199-1_11).
57. Chen, Q., Zhuo, Y., Sharma, P., Perez, I., Francis, D.J., Chakravarthy, S., Vishnivetskiy, S.A., Berndt, S., Hanson, S.M., Zhan, X., et al. (2021). An eight amino acid Segment Controls Oligomerization and Preferred Conformation of the two Non-visual arrestins. *J. Mol. Biol.* 433, 166790. <https://doi.org/10.1016/j.jmb.2020.166790>.
58. Gurevich, V.V., and Gurevich, E.V. (2019). Plethora of functions packed into 45 kDa arrestins: biological implications and possible therapeutic strategies. *Cell. Mol. Life Sci.* 76, 4413–4421. <https://doi.org/10.1007/s00018-019-03272-5>.
59. Isaikina, P., P.I., Jakob, R.P., Sarma, P., Ranjan, A., Baruah, M., Panwalkar, V., Maier, T., Shukla, A.K., and Grzesiek, S. (2022). A key GPCR phosphorylation motif discovered in arrestin2•CCR5 phosphopeptide complexes. *Mol. Cell.* 83. <https://doi.org/10.1016/j.molcel.2023.05.002>.
60. Kumari, P., Srivastava, A., Banerjee, R., Ghosh, E., Gupta, P., Ranjan, R., Chen, X., Gupta, B., Gupta, C., Jaiman, D., and Shukla, A.K. (2016). Functional competence of a partially engaged GPCR-beta-arrestin complex. *Nat. Commun.* 7, 13416. <https://doi.org/10.1038/ncomms13416>.
61. Kumari, P., Srivastava, A., Ghosh, E., Ranjan, R., Dogra, S., Yadav, P.N., and Shukla, A.K. (2017). Core engagement with beta-arrestin is dispensable for agonist-induced vasopressin receptor endocytosis and ERK activation. *Mol. Biol. Cell* 28, 1003–1010. <https://doi.org/10.1091/mbc.E16-12-0818>.
62. Cahill, T.J., 3rd, Thomsen, A.R., Tarrasch, J.T., Plouffe, B., Nguyen, A.H., Yang, F., Huang, L.Y., Kahsai, A.W., Bassoni, D.L., Gavino, B.J., et al. (2017). Distinct conformations of GPCR-beta-arrestin complexes mediate desensitization, signaling, and endocytosis. *Proc. Natl. Acad. Sci. USA* 114, 2562–2567. <https://doi.org/10.1073/pnas.1701529114>.
63. Asher, W.B., Terry, D.S., Gregorio, G.G.A., Kahsai, A.W., Borgia, A., Xie, B., Modak, A., Zhu, Y., Jang, W., Govindaraju, A., et al. (2022). GPCR-mediated beta-arrestin activation deconvoluted with single-molecule precision. *Cell* 185, 1661–1675.e16. <https://doi.org/10.1016/j.cell.2022.03.042>.
64. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>.
65. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
66. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. <https://doi.org/10.1107/S0907444910007493>.
67. 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. <https://doi.org/10.1038/nmeth.4169>.
68. Laskowski, R.A., Jabłońska, J., Pravda, L., Vařeková, R.S., and Thornton, J.M. (2018). PDBsum: structural summaries of PDB entries. *Protein Sci.* 27, 129–134. <https://doi.org/10.1002/pro.3289>.

69. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, 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. <https://doi.org/10.1107/S2059798319011471>.
70. Zivanov, J., Nakane, T., and Scheres, S.H.W. (2020). Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* 7, 253–267. <https://doi.org/10.1107/S2052252520000081>.
71. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* 152, 36–51. <https://doi.org/10.1016/j.jsb.2005.07.007>.
72. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
73. Baidya, M., Kumari, P., Dwivedi-Agnihotri, H., Pandey, S., Sokrat, B., Sposini, S., Chaturvedi, M., Srivastava, A., Roy, D., Hanyaloglu, A.C., et al. (2020). Genetically encoded intrabody sensors report the interaction and trafficking of beta-arrestin 1 upon activation of G-protein-coupled receptors. *J. Biol. Chem.* 295, 10153–10167. <https://doi.org/10.1074/jbc.RA120.013470>.
74. Peisley, A., and Skiniotis, G. (2015). 2D projection analysis of GPCR complexes by negative stain electron microscopy. *Methods Mol. Biol.* 1335, 29–38. [https://doi.org/10.1007/978-1-4939-2914-6\\_3](https://doi.org/10.1007/978-1-4939-2914-6_3).
75. Schrödinger, L., and DeLano, W. (2020). PyMOL. <http://www.pymol.org/pymol>.
76. Kawakami, K., Yanagawa, M., Hiratsuka, S., Yoshida, M., Ono, Y., Hiroshima, M., Ueda, M., Aoki, J., Sako, Y., and Inoue, A. (2022). Heterotrimeric Gq proteins act as a switch for GRK5/6 selectivity underlying beta-arrestin transducer bias. *Nat. Commun.* 13, 487. <https://doi.org/10.1038/s41467-022-28056-7>.
77. Dwivedi-Agnihotri, H., Sarma, P., Deeksha, S., Kawakami, K., Inoue, A., and Shukla, A.K. (2022). An intrabody sensor to monitor conformational activation of beta-arrestins. *Methods Cell Biol.* 169, 267–278. <https://doi.org/10.1016/bs.mcb.2021.12.023>.
78. Pandey, S., Roy, D., and Shukla, A.K. (2019). Measuring surface expression and endocytosis of GPCRs using whole-cell ELISA. *Methods Cell Biol.* 149, 131–140. <https://doi.org/10.1016/bs.mcb.2018.09.014>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Monoclonal ANTI-FLAG M2-HRP antibody	Sigma-Aldrich	Cat# A8592; RRID: AB_439702
<b>Chemicals, peptides, and recombinant proteins</b>		
TRIS	SRL	Cat# 71033
HEPES	SRL	Cat# 63732
NaCl	SRL	Cat# 41721
EDTA	SRL	Cat# 12070
Phenylmethanesulfonyl Fluoride (PMSF)	SRL	Cat# 84375 (84375)
Benzamidine Hydrochloride	SRL	Cat# 93014 (0248255)
Lysozyme	SRL	Cat# 45822
Glycerol	SRL	Cat# 77453
Dithiothreitol	HiMedia	Cat# MB070
Lauryl Maltose Neopentyl Glycol (MNG)	Anatrace	Cat# NG310, CAS no.1257852-96-2
Paraformaldehyde (PFA)	Sigma Aldrich	Cat# P6148, CAS no. 30525-89-4
Poly-D-lysine	Sigma Aldrich	Cat# P0899
TMB (Tetramethylbenzidine)	Thermo Fisher Scientific	Cat# 34028
Janus Green B	Sigma Aldrich	Cat# 201677
PEI (Polyethylenimine)	Polysciences	Cat# 23966
Bovine Serum Albumin, BSA	SRL	Cat# 83803 (0140105)
HBSS - Hank's Balanced Salt Solution	Thermo Fisher Scientific	Cat# 14065
GIBCO Fetal Bovine Serum	Thermo Fisher Scientific	Cat# 10270-106
DMEM	Cellclone	Cat# CC3004
Phosphate-buffered saline (PBS)	Sigma Aldrich	Cat# D1283
GIBCO Penicillin-Streptomycin	Thermo Fisher Scientific	Cat# 15140122
Coelenterazine	Goldbio	Cat# CZ05
Glyco-diosgenin (GDN)	Anatrace	GDN101
Cholesteryl Hemisuccinate	Sigma	C6512
Coomassie brilliant Blue	SRL	Cat# 64222
Uranyl formate	Polysciences	Cat# 24762-1
Recombinant rat $\beta$ -arrestin1	Purified	N/A
Recombinant bovine $\beta$ -arrestin2	Purified	N/A
C5aR1pp1	Chemically synthesized	N/A
C5aR1pp2	Chemically synthesized	N/A
C5aR1pp3	Chemically synthesized	N/A
V2Rpp	Chemically synthesized	N/A
CXCR4pp1	Chemically synthesized	N/A
CXCR4pp2	Chemically synthesized	N/A
CXCR4pp3	Chemically synthesized	N/A
CXCR4pp4	Chemically synthesized	N/A
Recombinant human C5a	Purified	N/A
Bradykinin	Genscript	N/A
Arginine Vasopressin Peptide (AVP)	Genscript	N/A
Uranyl formate	Polysciences	Cat# 24762-1

(Continued on next page)

### Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Formvar/carbon coated 300 mesh copper grids	PELCO (Ted Pella)	Cat# 01753-F
<b>Critical commercial assays</b>		
Site directed mutagenesis kit	NEB	Cat# E0554
NanoBiT assay	Promega	N/A
<b>Deposited data</b>		
C5aR1pp- $\beta$ arr1-Fab30	This study	PDB: 8GO8, EMD-34173
V2Rpp- $\beta$ arr2-Fab30	This study	PDB: 8GOC, EMD-34175
C5aR1pp- $\beta$ arr2-Fab30	This study	PDB: 8GOO, EMD-34178
CXCR4pp- $\beta$ arr1-Fab30	This study	PDB: 8GP3, EMD-34188
C5aR1pp- $\beta$ arr1-Fab30-Local-refine	This study	PDB: 8I0N, EMD-35104
V2Rpp- $\beta$ arr2-Fab30-Local-refine	This study	PDB: 8I10, EMD-35115
C5aR1pp- $\beta$ arr2-Fab30-Local-refine	This study	PDB: 8I0Z, EMD-35114
CXCR4pp- $\beta$ arr1-Fab30-Local-refine	This study	PDB: 8I0Q, EMD-35106
<b>Experimental models: Cell lines</b>		
Human: HEK293	ATCC	Cat# CRL-3216
<b>Oligonucleotides</b>		
C5aR1_R-insertion SDM primer Forward: CGCTCCACAGTGGACACTATGG	This study	N/A
C5aR1_R-insertion SDM primer Reverse: GCGCGTGAATGACTTGCT	This study	N/A
B2R <sup>mut</sup> Forward: AACACGGACCTCCATCTCCGTG	This study	N/A
B2R <sup>mut</sup> Reverse: GTCATGGAGTTCTCCATCTGAATGGG	This study	N/A
M2R <sup>mut</sup> Forward: AAGTCTACTTCACTGGGCCAC	This study	N/A
M2R <sup>mut</sup> Reverse: GACGGTGTTTTCGTCCTG	This study	N/A
$\beta$ -arrestin2 Double mutant ( $\beta$ arr2DM) Fw: ggcgAACAACCGTGAAAAACGTG	This study	N/A
$\beta$ -arrestin2 Double mutant ( $\beta$ arr2DM) Rv: aggaaCGGGGTGATGGTGTAAAC	This study	N/A
<b>Recombinant DNA</b>		
PcDNA_3.1 (empty vector)	Dr. Arun K Shukla	N/A
pcDNA3.1_V2R-WT	Dr. Arun K Shukla	N/A
pcDNA3.1_C5aR1-WT	Dr. Arun K Shukla	N/A
pcDNA3.1_CXCR3-WT	Dr. Arun K Shukla	N/A
pcDNA3.1_CXCR7-WT	Dr. Arun K Shukla	N/A
pcDNA3.1_B2R-WT	Dr. Arun K Shukla	N/A
pcDNA3.1_M2R-WT	Dr. Arun K Shukla	N/A
pcDNA3.1_C5aR1-mut <sup>TRRST</sup>	This study	N/A
pcDNA3.1_B2R-mut <sup>ΔG368/L370T</sup>	Dr. Arun K Shukla	N/A
pcDNA3.1_M2R-mut <sup>TVKST</sup>	This study	N/A
pCAGGS_V2R-WT-SmBiT	Dr. Arun K Shukla	N/A
pCAGGS_C5aR1-SmBiT	Dr. Arun K Shukla	N/A
pCAGGS_SmBiT- $\beta$ arr2 <sup>WT</sup>	Dr. Asuka Inoue	N/A
pCAGGS_SmBiT- $\beta$ arr2 <sup>DM</sup>	This study	N/A
pCAGGS_LgBiT- $\beta$ arr2 <sup>WT</sup>	Dr. Asuka Inoue	N/A
pCAGGS_LgBiT- $\beta$ arr2 <sup>DM</sup>	This study	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pCAGGS_LgBiT-FYVE	Dr. Asuka Inoue	N/A
pCAGGS_LgBiT-Ib30	Dr. Asuka Inoue	N/A
<b>Software and algorithms</b>		
UCSF Chimera	Pettersen et al. <sup>64</sup>	<a href="https://www.cgl.ucsf.edu/chimera/">https://www.cgl.ucsf.edu/chimera/</a>
UCSF Chimera X	Pettersen et al. <sup>65</sup>	<a href="https://www.rbvi.ucsf.edu/chimerax/">https://www.rbvi.ucsf.edu/chimerax/</a>
COOT	Emsley et al. <sup>66</sup>	<a href="https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/">https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/</a>
cryoSPARC	Punjani et al. <sup>67</sup>	<a href="https://cryosparc.com/">https://cryosparc.com/</a>
PDBsum	Laskowski et al. <sup>68</sup>	<a href="http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/">http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/</a>
Phenix	Liebschner et al. <sup>69</sup>	<a href="https://www.phenix-online.org/">https://www.phenix-online.org/</a>
PyMol	Schrodinger	<a href="https://pymol.org/2/">https://pymol.org/2/</a>
Prism 8	GraphPad Software	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
Relion3.1.2	Zivanov et al. <sup>70</sup>	<a href="https://www3.mrc-lmb.cam.ac.uk/relion/index.php?title=Main_Page">https://www3.mrc-lmb.cam.ac.uk/relion/index.php?title=Main_Page</a>
SerialEM	Mastronarde <sup>71</sup>	<a href="https://bio3d.colorado.edu/SerialEM/">https://bio3d.colorado.edu/SerialEM/</a>
Graphpad Prism 9	GraphPad Software, San Diego, California USA	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
ImageJ	Schneider et al. <sup>72</sup>	<a href="https://imagej.nih.gov/ij/download.html">https://imagej.nih.gov/ij/download.html</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for reagents should be addressed to the lead contact, Dr. Arun K. Shukla ([arshukla@iitk.ac.in](mailto:arshukla@iitk.ac.in)).

### Materials availability

Reagents described in this manuscript are available upon reasonable request from the [lead contact](#).

### Data and code availability

- All three-dimensional cryo-EM density maps, coordinates for the atomic models and local-refined maps generated in this study have been deposited and are publicly available as of the date of publication. Accession numbers (EMDB and PDB IDs) are listed in the [key resources table](#). Original gel images have been deposited to Mendeley data, and they are publicly available after publication. The DOI is listed in the [key resources table](#).
- This paper does not report any original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Human cell lines

HEK-293 cells were purchased from ATCC for all the cellular experiments performed in the study. The cell line was examined frequently under the microscope for proper morphology, but they were not authenticated. They were cultured in DMEM with fetal bovine serum (FBS) at 37°C in 5% CO<sub>2</sub>. In this study, any stable, knockout, or knockdown cell lines were not generated, and the details of previously generated cell lines are referenced in the manuscript.

### Insect cells

Sf9 cells were obtained from Expression systems, and they were routinely monitored under the microscope for proper morphology. These cells were maintained in a shaker incubator at 27°C with 135rpm shaking, and sub-cultured in protein-free insect cell medium purchased from Expression Systems.



## METHOD DETAILS

## General reagents, plasmids, and cell culture

Most of the general reagents were purchased from Sigma Aldrich unless mentioned otherwise. Dulbecco's Modified Eagle's Medium (DMEM), Dulbecco's Phosphate buffer saline (PBS), Fetal-Bovine Serum (FBS), Trypsin-EDTA, Hank's balanced salt solution (HBSS), and penicillin-streptomycin solution were purchased from Thermo Fisher Scientific. HEK-293 cells were obtained from ATCC and maintained in DMEM (Gibco, Cat no. 12800-017) supplemented with 10% FBS (Gibco, Cat no. 10270-106) and 100U ml<sup>-1</sup> penicillin (Gibco, Cat no. 15140122) and 100μg ml<sup>-1</sup> streptomycin (Gibco, Cat no. 15140-122) at 37°C under 5% CO<sub>2</sub>. The cDNA coding region for the mentioned receptors namely, V2R, C5aR1, B2R, M2R, CXCR3, and CXCR7 were cloned in pcDNA3.1 consist of HA signal sequence followed by FLAG tag at the N-terminus of the receptor. The mutants generated for the study are as follows: deletion of G368 and a substitution L370T in B2R; insertion of an Arg residue between R337 and S338 in C5aR1; insertion of a Lys residue between V308 and S309 in M2R; using Q5 site-directed mutagenesis kit (NEB, Cat. no. E0554S). For the NanoBiT assay, receptors harboring SmBiT at the C-terminus were generated by subcloning in the lab, and other constructs have been described previously.<sup>41,43,73</sup> All the constructs were verified by DNA sequencing (Macrogen).

## Expression and purification of βarrs

Full length rat βarr1, βarr2<sup>WT</sup> and bovine βarr2<sup>DM</sup> were cloned into pGEX-4T3 vector with thrombin cleavage site between GST tag and βarr. Similar protocol was followed for purifying all three forms of βarr. βarrs were expressed in *E. coli* BL21 cells and grown in Terrific broth media supplemented with 100μg ml<sup>-1</sup> ampicillin. A primary culture of 50ml volume was inoculated with an isolated colony from freshly transformed LB-amp plate. Primary culture was grown till a cell optical density at 600nm (OD<sub>600</sub>) of 0.8-1 and further inoculated into a secondary culture of TB-Amp of 1.5L volume till OD<sub>600</sub> 0.8-1. The expression of βarrs was then induced with 25μM IPTG concentration and cells were allowed to grow till 16h at 18°C. Cultures were harvested and stored at -80°C until further use. Harvested pellets were of 12-15g in mass.

For purification, cells were lysed by sonication in lysis buffer; 25mM Tris, pH 8.5, 150mM NaCl, 1mM PMSF (phenylmethylsulfonyl fluoride), 2mM Benzamidine, 1mM EDTA (Ethylenediaminetetraacetic acid), 5% Glycerol, 2mM Dithiothreitol (DTT) and 1mg ml<sup>-1</sup> Lysozyme. The lysate was centrifuged at 18,000-20,000rpm at 4°C and supernatant was allowed to bind to Glutathione resin (GS resin) (Glutathione Sepharose<sup>TM</sup> 4 Fast Flow, GE Healthcare Cat. no. 17-5132-02) in a batch binding mode for overnight at 4°C. GS-resin bound GST-βarr was transferred into Econo columns (Biorad, Cat. no. 7372512) and washed rigorously with wash buffer (25mM Tris, pH 8.5, 150mM NaCl, 2mM DTT and 0.02% n-dodecyl-β-D-maltopyranoside (DDM)). Afterward, on-column cleavage was set up by adding thrombin to 1:1 resin:buffer slurry at room temperature for 2h. βarrs were then eluted with gravity flow and further with buffer 25mM Tris, pH 8.5, 350mM NaCl and 0.02% DDM and 2mM DTT. Eluted proteins were concentrated and further purified on a HiLoad 16/600 Superdex column in buffer 25mM Tris, pH 8.5, 350mM NaCl, 2mM DTT and 0.02% DDM. Fractions corresponding to pure βarr were flash frozen with 10% glycerol and stored at -80°C until further use.

## Expression and purification of Fabs

A similar protocol for expression and purification was followed for all the Fabs and they were purified as previously mentioned.<sup>39</sup> Briefly, Fabs were expressed in the periplasmic fraction of *E. coli* M55244 cells (ATCC) and purified using Protein L resin (GE Healthcare Cat. no. 17547802) with gravity flow affinity chromatography. Cells transformed with Fab plasmid were grown in 50ml 2xYT media and allowed to grow overnight at 30°C. 1L 2xYT media was inoculated with 5% volume of initial inoculum and grown for an additional 8h at 30°C. Cells were collected and resuspended in an equal volume of CRAP medium supplemented with 100μg ml<sup>-1</sup> ampicillin, and grown further for 16h at 30°C. For purification, cells were lysed in lysis buffer (50mM HEPES-Na<sup>+</sup>, pH 8.0, 0.5M NaCl, 0.5% (v/v) Triton X-100, 0.5mM MgCl<sub>2</sub>) by sonication. Cell lysate was heated in a 65°C water bath for 30min and cooled immediately on ice for 5min. Lysate was centrifuged at 20,000rpm and passed through pre-equilibrated Protein L resin packed gravity flow affinity columns. Binding was performed at room temperature and beads were washed extensively with wash buffer (50mM HEPES-Na<sup>+</sup>, pH 8.0, 0.5M NaCl). Fabs were eluted with 100mM acetic acid into tubes containing 10% volume of 1M HEPES, pH 8.0 for neutralization. Eluted samples were desalted into buffer (20mM HEPES-Na<sup>+</sup>, pH 8.0, 0.1M NaCl) using a pre-packed PD-10 column (GE Healthcare Cat. no. 17085101). Purified Fabs were flash-frozen and stored at -80°C supplemented with 10% (v/v) glycerol until further use.

## Co-immunoprecipitation assay

For co-immunoprecipitation assay, 2.5μg of β-arrestins were incubated with different phosphopeptides at 10-fold molar excess in binding buffers (20mM HEPES, pH 7.4, 150mM NaCl) for 1h at room temperature for activation. Post peptide-induced activation, 5μg Fab30 was added, and reaction was incubated for an additional 1h at room temperature. After 1h, 25μl of pre-equilibrated protein L beads (Capto<sup>TM</sup> L resin, GE Healthcare Cat. no. 17547802) was added and reaction was incubated for 90min at room temperature, followed by five washes with binding buffer containing 0.01% LMNG. Bound protein was eluted with 2X SDS loading buffer and 15μl sample was analyzed on 12% SDS-PAGE. For statistical analyses, protein bands were quantified using ImageJ software suite<sup>72</sup> and the values were plotted using GraphPad Prism software v 9.5. The data were normalized with respect to their respective experimental control and appropriate statistical analyses were performed as indicated in the corresponding figure legend.

### Reconstitution of phosphopeptide- $\beta$ arr-Fab complexes

A previously published protocol was followed for complex purification with minor modifications.<sup>31</sup> Briefly,  $\beta$ arrs were activated with corresponding phosphopeptides at a 1:3 molar ratios of  $\beta$ arr:phosphopeptide for 30–40 min at room temperature. Respective Fabs were added to the phosphopeptide- $\beta$ arr mixture at 1:1.5 molar ratio of  $\beta$ arr:Fab and incubated for 1 h at room temperature. To remove excess Fabs, the phosphopeptide- $\beta$ arr-Fab complexes were concentrated with 30,000 MWCO concentrators (Vivaspin, Cytiva Cat. no. 28932361) and injected into Superose 6 Increase 10/300 GL (Cytiva Cat. no. 29091596) gel-filtration column. Fractions were further analyzed on SDS-PAGE and selected fractions were pooled and concentrated for structural studies.

### Negative-staining EM

Complex formation, homogeneity, and particle quality of the samples were judged through negative staining of the samples prior to data collection under cryogenic conditions for high resolution reconstructions. Negative staining of the samples was performed with uranyl formate in accordance with the previously published protocols.<sup>74</sup> For imaging, 3.5  $\mu$ l of the samples were dispensed on glow discharged carbon/formvar coated 300 mesh Cu grids (PELCO, Ted Pella) and allowed to adsorb for 1 min, followed by blotting off the sample using a filter paper. The grid was then touched on a first drop of freshly prepared 0.75% (w/v) uranyl formate stain and immediately blotted off, followed by staining for 30 sec on a second drop of stain. Imaging of the negatively stained samples were performed on a FEI Tecnai G2 12 Twin TEM (LaB6) operating at 120 kV and equipped with a Gatan CCD camera (4k x 4k) at 30,000x magnification. Data processing of the collected micrographs for the individual samples were performed with Relion 3.1.2.<sup>70</sup> Approximately 10,000 particles were autopicked using the gaussian blob picker within Relion and the extracted particles were subjected to reference free 2D classification.

### Cryo-EM sample preparation and data acquisition

Quantifoil holey carbon grids (Cu or Au, R2/1 or R2/2) were glow discharged for 45 sec with a Glocube glow discharge system (Quorum technologies Ltd, UK). 3  $\mu$ l of the complex was dispensed on the glow discharged grid, blotted for 3 sec with a Whatman paper filter no. 1 at 10°C and maintained at 90% humidity and then plunge frozen into liquid ethane (−180 °C) using a Leica GP plunger (Leica Microsystems, Austria).

For C5aR1pp- $\beta$ arr1-Fab30 complex, cryo-EM data collection was performed on R2/2 Cu 300 mesh grid using a Titan Krios electron microscope (ThermoFisher Scientific, USA) operating at 300 kV equipped with the Gatan Energy Filter. Movies were recorded in counting mode with a Gatan K2 Summit DED (Gatan, USA) using the automated SerialEM software<sup>71</sup> at a nominal magnification of 165,000x and a pixel size of 0.82 Å at sample level. 6,212 movie stacks consisting of 40 frames were recorded over a defocus range of 0.5 to 2.5  $\mu$ m with a total dose of 49 e<sup>−</sup>/Å<sup>2</sup> and total exposure time of 5 sec.

For CXCR4pp- $\beta$ arr1-Fab30 complex, cryo-EM data collection was performed on R2/2 Cu 300 mesh grid using a Glacios electron microscope (ThermoFisher Scientific, USA) operating at 200 kV. Movies were recorded in counting mode with a Gatan K3 DED (Gatan, USA) using the automated SerialEM software at a nominal magnification of 46,000x and a pixel size of 0.878 Å at sample level. 5,637 movie stacks consisting of 40 frames were recorded over a defocus range of 0.5 to 2.5  $\mu$ m with a total dose of 49.3 e<sup>−</sup>/Å<sup>2</sup> and total exposure time of 2.9 sec.

For V2Rpp- $\beta$ arr2-Fab30 complex, cryo-EM data collection was performed on R2/2 Au 200 mesh grid using a Titan Krios electron microscope (ThermoFisher Scientific, USA) operating at 300 kV equipped with the Gatan Energy Filter. Movies were recorded in counting mode with a Gatan K2 Summit DED (Gatan, USA) using the automated SerialEM software at a nominal magnification of 165,000x and a pixel size of 0.82 Å at sample level. 9,720 movie stacks consisting of 40 frames were recorded over a defocus range of 0.5 to 2.5  $\mu$ m with a total dose of 48.7 e<sup>−</sup>/Å<sup>2</sup> and total exposure time of 4 sec.

For C5aR1pp- $\beta$ arr2-Fab30 complex, cryo-EM data collection was performed on R2/2 Cu 300 mesh grid using a Glacios electron microscope (ThermoFisher Scientific, USA) operating at 200 kV. Movies were recorded in counting mode with a Gatan K3 DED (Gatan, USA) using the automated SerialEM software at a nominal magnification of 46,000x and a pixel size of 0.878 Å at sample level. 8,614 movie stacks consisting of 40 frames were recorded over a defocus range of 0.5 to 2.5  $\mu$ m with a total dose of 51 e<sup>−</sup>/Å<sup>2</sup> and total exposure time of 3 sec.

### Cryo-EM data processing and model building

All image processing steps were performed in cryoSPARC version 3.3.2<sup>67</sup> unless otherwise stated. In brief, for the C5aR1pp- $\beta$ arr1-Fab30 complex, 6,212 movie stacks were subjected to patch motion correction (multi), followed by CTF refinement with patch CTF multi. 5,790 motion corrected micrographs with CTF fit resolution better than 4.5 Å were selected for further processing. 4,304,237 particle projections were automatically picked with blob picker, extracted with a box size of 480 pixels and fourier cropped to 64 pixels. The particle stack so obtained was subjected to multiple rounds of 2D classification. The class averages with clear secondary structural features were selected and re-extracted with a box size of 480 pixels and fourier cropped to 256 pixels resulting in a pixel size of 1.5375 Å. 295,922 re-extracted particles were then subjected to Ab-initio reconstruction and 3D classification/Heterogeneous refinement with C1 symmetry yielding 4 models. 80,437 particles corresponding to a dimer and containing 47.9% of the total particles were subjected to non-uniform refinement with C2 symmetry to yield a map with an estimated resolution of 3.41 Å (voxel size of 1.5375 Å). Further, local refinement was performed by masking out the variable domains of Fab30 resulting into an estimated resolution of 3.26 Å. Local resolution of all reconstructions was estimated using the Blocres within cryoSPARC version 3.3.2.

For the CXCR4pp- $\beta$ arr1-Fab30 complex data set, 5,637 movies were motion corrected using a patch of 5x5 patch within patch motion correction (multi). Following CTF estimation, 5,236 motion corrected micrographs with CTF fit resolution better than 6Å were curated for further processing. 3,236,193 particles were automatically picked using the blob-picker sub-program and subsequently extracted with a box size of 480pixels and fourier cropped to 64pixels. The extracted particles were subjected to several rounds of 2D classification to remove junk particles. 104,707 particles corresponding to the clean class averages were selected, re-extracted with a box size of 480pixels and fourier cropped to 256pixels (pixel size of 1.65) and used to produce two ab-initio models. The particles corresponding to the two ab-initio models were subjected to heterogenous refinement/3D classification which produced a 3D class with clear dimeric conformation and a particle count of 86,525. This particle set was re-extracted with full box size of 480pixels (pixel size of 0.878) and subjected to non-uniform refinement with C2 symmetry which converge to a map with 4.81Å resolution as estimated using the gold standard Fourier Shell Correlation (GFSC) using the 0.143 criterion. Local refinement was performed by masking out the variable domains of Fab30 which resulted into an estimated resolution of 4.45Å.

For the V2Rpp- $\beta$ arr2-Fab30 complex, 9,720 movies were motion corrected with 5x5 patches followed by CTF estimation with patch CTF (multi). Following CTF refinement, 8,295 movies with CTF fit resolution better than 4.5Å were used for further processing. Particle picking from the curated micrographs was performed automatically with the blob picker sub-program to obtain an initial stack of 2,444,407 particles. The particles were then extracted with a box size of 512pixels and fourier cropped to a box size of 64pixels. The extracted particles were subjected to several rounds of reference free 2D classification. 2D class averages with evident secondary features containing 161,436 particles were extracted with a box size of 512pixels and fourier cropped to a box size of 256 (pixel size of 1.64). This sub-set of particles was used for ab-initio reconstruction and subsequent rounds of 3D/Heterogeneous classification with C1 symmetry to obtain 2 models. 92,018 particles corresponding to a trimer were re-extracted with full box size of 512pixels which refined to an overall resolution of 4.18Å (voxel size of 0.82Å) with NU refinement (C3 symmetry) according to the gold standard Fourier shell correlation (FSC) criterion of 0.143. Subsequently, local refinement was performed on  $\beta$ arr and the variable domains of Fab30 and resulted into an estimated resolution of 3.96Å.

For the C5aR1pp- $\beta$ arr2-Fab30 complex data set, 8,614 movies were motion corrected using patch motion correction (multi) and subsequent CTF estimation was performed through patch CTF (multi). 8,157 micrographs with CTF fit resolution better than 6Å were curated for particle picking using the blob picker sub-program. 4,012,616 particles were automatically picked and extracted with a box size of 512pixels and fourier cropped to 64pixels. Reasonable class averages after several rounds of reference free 2D classification yielded a particle set containing 54,193 particle projections, which was re-extracted with a box size of 512pixels and fourier cropped to 360pixels (pixel size of 1.2487Å) for subsequent used for ab-initio reconstruction generating two ab-initio models. Following heterogenous refinement/3D classification, the 3D class with evident features of a trimer and containing 38,206 particles was subjected to non-uniform refinement with C3 symmetry to yield a reconstruction at 4.41Å (final voxel size of 1.2487Å) as determined by gold standard Fourier Shell Correlation (FSC) using the 0.143 criterion. Local refinement was performed masking out the variable domains of Fab30 which improved the resolution to 4.33Å.

### Model building and refinement

Coordinates from a previously solved V2Rpp bound  $\beta$ arr1 structure (PDB 4JQI) was used to dock the model into the EM density map of C5aR1pp- $\beta$ arr1-Fab30 using Chimera.<sup>64</sup> The EM map was then used for manual rebuilding of the  $\beta$ arr1 residues and placing the phosphopeptide in COOT.<sup>66</sup> The rebuilt model was subjected to real space refinement in Phenix<sup>69</sup> to obtain a model with 97.23% of the residues in most favored region and 2.77% in the allowed region of the Ramachandran plot.

The protomeric structure from the IP6- $\beta$ arr2 (PDB 5TV1) complex solved in a previous study was used as an initial model to dock into the density map of V2Rpp- $\beta$ arr2-Fab30 complex and regenerate the trimeric complex with C3 symmetry. The rigid body fitted trimeric model and the phosphopeptides were then rebuilt manually into the EM density map. The rebuilt trimeric coordinates with the phosphopeptides were subsequently subjected to real space refinement in Phenix to reach a final model with 95.05% in the favored region and 4.76% in the allowed region of the Ramachandran plot.

For model building into the 4.33Å C5aR1pp- $\beta$ arr2-Fab30 coulombic map, the co-ordinates corresponding to V2Rpp peptide were deleted from the trimeric co-ordinates of V2Rpp- $\beta$ arr2-Fab30 complex (PDB 8GOC), and the resulting model was docked into the EM map in Chimera. The “all atom refine” sub-module within the “refine” module in COOT was used for initial fitting of the model into the EM map, followed by manual rebuilding of the phosphopeptides. Multiple rounds of Phenix real space refinement combined with iterative model building yielded a model with 94.9% of the residues residing in the most favored region of the Ramachandran plot.

The dimeric co-ordinates from the cryo-EM structure of C5aR1pp- $\beta$ arr1-Fab30 (PDB 8GO8) without the phosphopeptide was used as an initial model to dock into the CXCR4pp- $\beta$ arr1-Fab30 EM map using Chimera. The docked model along with the coulombic map were imported into COOT and the model was subjected to “all atom refine” for fitting the atoms into the density. The phosphopeptide was manually built into the density to yield a complete model, which was subsequently used to refine the model against the EM map with Phenix real space refinement. The final refined model had 96.62% residues in the most favored regions and 3.38% in the allowed regions of the Ramachandran plot. For model building into the locally refined maps, coordinates of the individual full-length structures were used to dock into the corresponding maps followed by iterative rounds of manual adjustments in COOT and real space refinement in Phenix.

All the refined models were validated using “Comprehensive Validation (cryo-EM)” sub-module in Phenix. 3D reconstruction and model refinement statistics for both full and local-refined structures are provided as Table 1. Figures in the manuscript have been

prepared with Chimera<sup>64</sup> and ChimeraX<sup>65</sup> software. Domain rotation analysis was performed with PyMOL.<sup>75</sup> The interaction interface of  $\beta$ arr oligomers in the cryo-EM structures were identified using PDBSum.<sup>68</sup>

### NanoBiT assay for $\beta$ arr2<sup>WT</sup> and $\beta$ arr2<sup>DM</sup> recruitment

$\beta$ arr2<sup>WT</sup> and  $\beta$ arr2<sup>DM</sup> recruitment downstream of V2R and C5aR1 in response to AVP and C5a, respectively, was measured using NanoBiT (Enzyme linked complementation-based assay) assay following the protocol described earlier.<sup>76</sup> Receptor constructs were tagged with SmBiT at the carboxyl-terminus, and  $\beta$ arr2 constructs were N-terminally tagged with LgBiT. Briefly, HEK-293 cells were transfected with indicated receptor constructs (3.5  $\mu$ g) and  $\beta$ arr2 ( $\beta$ arr2<sup>WT/DM</sup>) constructs (3.5  $\mu$ g) using polyethylenimine (PEI) linear (Polysciences, Cat. no. 19850) at a ratio of 1:3 (DNA:PEI linear). After 16–18 h of transfection, cells were trypsinized, harvested, and resuspended in assay buffer (1X HBSS, 5 mM HEPES, pH 7.4, 0.01% BSA) containing 10  $\mu$ M coelenterazine (GoldBio, Cat. no. CZ05). Resuspended cells were seeded in a white flat bottom 96-well plate (100  $\mu$ l well<sup>-1</sup>). After 2 h of incubation (90 min at 37°C and 30 min at room temperature), basal luminescence was recorded using a multimode plate reader (FLUOstar Omega, BMG Labtech). Later, cells were stimulated with varying doses of indicated ligands followed by measurement of luminescence signal for 20 cycles. For data analysis, ligand induced change in signals were taken and normalized with the lowest ligand dose luminescence value, and fold normalized data was plotted using nonlinear regression three-parameter sigmoidal concentration-response curve in GraphPad Prism v 9.5 software.

### NanoBiT assay for $\beta$ arr trafficking

Agonist-induced  $\beta$ arr2<sup>WT</sup> and  $\beta$ arr2<sup>DM</sup> endosomal trafficking downstream of the receptors mentioned above was studied using NanoBiT assay as described in the recruitment experiment. The only exception from the recruitment assay was that the receptor constructs were not tagged with SmBiT, but rather  $\beta$ arr2 ( $\beta$ arr2<sup>WT/DM</sup>) and FYVE constructs N-terminally fused with SmBiT and LgBiT respectively were used for enzyme complementation. For each experiment, 3  $\mu$ g of indicated receptors, 2  $\mu$ g of SmBiT- $\beta$ arr2<sup>WT/DM</sup>, and 5  $\mu$ g of LgBiT-FYVE were used. Fold normalized change in signals were plotted using nonlinear regression three-parameter sigmoidal concentration-response curve in GraphPad Prism v 9.5 software.

### NanoBiT assay for Ib30 reactivity

To assess Ib30 reactivity in response to an agonist for the mentioned receptors, NanoBiT assay was used following the same protocol as discussed in the  $\beta$ arr2<sup>WT</sup> and  $\beta$ arr2<sup>DM</sup> recruitment assay.<sup>77</sup> For enzyme complementation, N-terminally SmBiT fused  $\beta$ arr1 and N-terminally LgBiT fused Ib30 were used. For transfection, 3  $\mu$ g receptor except for M2R, CXCR7 (5  $\mu$ g), and CXCR3 (7  $\mu$ g), 2  $\mu$ g SmBiT- $\beta$ arr1, and 5  $\mu$ g LgBiT-Ib30 were used. Transfected cells were stimulated with varying doses of respective ligands (mentioned in corresponding figures). Fold normalized change in luminescence were plotted using nonlinear regression three/four-parameter sigmoidal concentration-response curve in GraphPad Prism v 9.5 software.

### Receptor surface expression

Receptor surface expression in various assays was measured using a previously described whole cell-based surface ELISA assay.<sup>78</sup> To study the surface expression of the receptor, cells transfected with a particular receptor were seeded into a 0.01% poly-D-Lysine pre-coated 24-well plate at a density of 2  $\times$  10<sup>5</sup> cells well<sup>-1</sup>. Post 24 h of seeding cells were washed once with ice-cold 1XTBS, fixed with 4% PFA (w/v in 1XTBS) on ice for 20 min, washed again three times with 1XTBS, and blocked with 1% BSA (prepared in 1XTBS) at room temperature for 1.5 h. Afterward, cells were incubated with anti-FLAG M2-HRP antibody at 1:5000 dilution (Sigma, Cat. no. A8592) for 1.5 h, which was followed by three washes in 1% BSA. Subsequently, incubated with TMB-ELISA substrate (Thermo Fisher Scientific, Cat. no. 34028) until a light blue color appeared. To quench the reaction, 100  $\mu$ l of the colored solution was transferred to another 96-well plate containing 100  $\mu$ l of 1M H<sub>2</sub>SO<sub>4</sub>, and the absorbance was measured at 450 nm. Afterward, the TMB substrate was removed, washed twice with 1XTBS, and incubated with 0.2% (w/v) Janus Green (Sigma; Cat. no. 201677) for 15 min at room temperature. Later, cells were washed with water to remove the excess stain, followed by the addition of 800  $\mu$ l of 0.5N HCl in each well. Thereupon, the colored solution was transferred to a 96-well plate for measuring the absorbance at 595 nm. The signal intensity was normalized by calculating the ratio of A<sub>450</sub>/A<sub>595</sub> values followed by quantifying fold increase with respect to the A<sub>450</sub>/A<sub>595</sub> value of negative control (mock transfection) and plotted using the GraphPad Prism v 9.5.

### QUANTIFICATION AND STATISTICAL ANALYSIS

GraphPad Prism v9.5 was used to plot and analyze all the functional data presented in this manuscript, and all the relevant details such as number of replicates, data normalization, mean  $\pm$  sem, and statistical analyses are mentioned in the corresponding figure legends.

# Structure-guided engineering of biased-agonism in the human niacin receptor via single amino acid substitution

Manish K. Yadav<sup>1#</sup>, Parishmita Sarma<sup>1#</sup>, Manisankar Ganguly<sup>1</sup>, Sudha Mishra<sup>1</sup>, Jagannath Maharana<sup>1</sup>, Nashrah Zaidi<sup>1</sup>, Annu Dalal<sup>1</sup>, Vinay Singh<sup>1</sup>, Sayantan Saha<sup>1</sup>, Gargi Mahajan<sup>1</sup>, Saloni Sharma<sup>1</sup>, Mohamed Chami<sup>1</sup>, Ramanuj Banerjee<sup>1\*</sup> and Arun K. Shukla<sup>1\*</sup>

<sup>#</sup>Joint 1<sup>st</sup> author

<sup>1</sup>Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur 208016, India; <sup>2</sup>BioEM Lab, Biozentrum, Universität Basel, Basel, Switzerland.

\*Corresponding authors (ramanujb@iitk.ac.in or arshukla@iitk.ac.in)

**Keywords:** GPCRs, G-proteins, dyslipidemia, biased-agonism, drug discovery



## Abstract

The Hydroxycarboxylic acid receptor 2 (HCA2), also known as the niacin receptor or GPR109A, is a prototypical G protein-coupled receptor that plays a central role in the inhibition of lipolytic and atherogenic activities in our body. Interestingly, GPR109A activation also results in vasodilation that is linked to the side-effect of flushing associated with dyslipidemia drugs such as niacin. This receptor continues to be a key target for developing novel pharmacophores and lead compounds as potential therapeutics in dyslipidemia with minimized flushing response, however, the lack of structural insights into agonist-binding and receptor activation has limited the efforts. Here, we present five different cryo-EM structures of the GPR109A-G-protein complexes with the receptor bound to dyslipidemia drugs, niacin or acipimox, non-flushing agonists, MK6892 or GSK256073, and recently approved psoriasis drug, monomethyl fumarate (MMF). These structures allow us to visualize the binding mechanism of agonists with a conserved molecular interaction network, and elucidate the previously lacking molecular basis of receptor activation and transducer-coupling. Importantly, cellular pharmacology experiments, guided by the structural framework determined here, elucidate pathway-selective biased signaling elicited by the non-flushing agonists. Finally, taking lead from the structural insights, we successfully engineered receptor mutants via single amino acid substitutions that either fail to elicit agonist-induced transducer-coupling or exhibits G-protein signaling bias. Taken together, our study provides previously lacking structural framework to understand the agonist-binding and activation of GPR109A, and opens up the possibilities of structure-guided novel drug discovery targeting this therapeutically important receptor.

## Introduction

The Hydroxycarboxylic acid receptor 2 (HCA2), also known as the niacin receptor or GPR109A, belongs to the superfamily of G protein-coupled receptors (GPCRs), and it is expressed primarily in the adipose tissues<sup>1-3</sup>, keratinocytes<sup>4</sup>, immune cells such as neutrophils and Langerhans cells<sup>5</sup> in our body. Upon activation by agonists, GPR109A couples to Gai sub-family of heterotrimeric G-proteins leading to lowering of cAMP response<sup>6-9</sup>. In addition, activated GPR109A also recruits  $\beta$ -arrestins<sup>10</sup>, which are multifunctional proteins involved in GPCR desensitization, trafficking and downstream signaling<sup>11</sup>. Interestingly, GPR109A was identified as the molecular target for the action of nicotinic acid (aka, niacin), an effective drug prescribed for lowering the triglycerides, almost two decades ago<sup>12</sup>. Moreover, GPR109A activation also mediates the lowering of LDL (aka, bad cholesterol), enhancing the levels of HDL (aka, good cholesterol)<sup>13</sup>. Furthermore, monomethyl fumarate (MMF), the active metabolite of a psoriasis drug, Fumaderm, and also a therapeutic agent for the treatment of relapsing forms of multiple sclerosis, has also been identified as an agonist of GPR109A<sup>14-16</sup>. However, activation of GPR109A is also responsible for driving the troublesome side effect of flushing response associated with niacin, acipimox and MMF<sup>10,13,16-19</sup>. This represents a potential limitation with their therapeutic usage, and therefore, additional small molecule agonists targeting GPR109A remains a key focus area<sup>7,19-21</sup>.

Several non-flushing agonists, such as MK6892 and GSK256073 with high affinity for GPR109A have been developed and characterized using *in-vitro* and animal studies although none of these compounds is yet approved for clinical usage<sup>22,23</sup>. In addition, a comprehensive study has also demonstrated that the side effect of niacin-induced flushing response in mouse is driven primarily by  $\beta$ -arrestin-mediated downstream signaling, and therefore, G-protein-biased agonists of GPR109A may represent improved therapeutics compared to niacin<sup>10</sup> (**Figure 1A**). In the same study, a previously developed agonist MK0354 was reported to maintain the anti-lipolytic effect with significant reduction in flushing response, and it was further characterized as a G-protein-biased agonist<sup>10</sup>. Still however, direct structural visualization and molecular mechanism of agonist-binding and activation of GPR109A remain primarily elusive and represent an important knowledge gap to efficiently target this receptor for therapeutic benefits.

Here, we present five different cryo-EM structures of GPR109A-G-protein complexes where the receptor is activated either by niacin, acipimox, MK6892, GSK256073 and MMF. Comparison of these structural snapshots provides the molecular basis of ligand recognition, activation, and transducer-coupling by GPR109A. Importantly, the structural insights allow us to rationally design receptor mutants harbouring single amino acid substitution that either render the receptor completely inactive with respect to transducer-coupling, or, impart significant transducer-coupling bias. Our study not only illuminates the structural pharmacology of GPR109A ligands and paves the way for structure-guided discovery of novel therapeutics but also offers a framework to leverage the structural information to rationally encode signaling-bias in GPCRs.

## Results

### GPR109A agonists used for structural analysis

In order to visualize the molecular framework of ligand recognition and receptor activation, we selected five different ligands namely, niacin, acipimox, MK6892, GSK256073 and MMF (**Figure 1B**). Of these, niacin and acipimox are clinically prescribed drugs to treat dyslipidemia, while MK6892 and GSK256073 have been developed as non-flushing agonists of GPR109A. MK6892 is a biaryl cyclohexene carboxylic acid derivative that was reported to exhibit high affinity for GPR109A without significant off-target profile, and also displayed reduced vasodilation in animal studies while maintaining free fatty acid reduction similar to niacin<sup>22</sup>. GSK256073 was reported to display robust specificity for GPR109A over the other hydroxycarboxylic acid receptor subtypes, maintain the ability to lower the levels of non-esterified fatty acids in pre-clinical animal studies with reduced flushing response, and even promising outcomes in healthy male subjects<sup>23</sup>. MMF on the other hand, is the active metabolite of psoriasis drug, Fumaderm, and is also used as a therapeutic agent in multiple sclerosis<sup>14,15,24</sup>. Our selection of these ligands was based on their diverse chemical structures, therapeutic profile, and associated side effects with the goal to understand their interaction with GPR109A and potentially link the structural insights with their therapeutic profile.

We measured the pharmacological profile of acipimox, MK6892, GSK250673 and MMF with niacin as a reference agonist of GPR109A in G-protein response and  $\beta$ -arrestin recruitment assays (**Figure 1C-E**). We observed that acipimox behaved as a full agonist but with lower potency

in G-protein dissociation, cAMP response, and  $\beta$ -arrestin recruitment assay (**Figure 1C**). Moreover, MK6892 exhibited higher efficacy in G-protein response and similar efficacy but lower potency in  $\beta$ -arrestin recruitment compared to niacin (**Figure 1D**). In the case of GSK256073, we observed a higher response in G-protein dissociation but similar efficacy in cAMP response, and it also displayed higher efficacy in  $\beta$ -arrestin recruitment as compared to niacin (**Figure 1D**). On the other hand, MMF behaved as a full agonist in both G-protein and  $\beta$ -arrestin recruitment assay with similar efficacy as niacin but slightly weaker potency in  $\beta$ -arrestin assays (**Figure 1E**). Analysis of these pharmacology data and calculation of the bias factor suggest that MK6892 acts as a G-protein biased agonist at GPR109A (**Figure 1F**). In these cellular assays, the surface expression of GPR109A was measured using a previously described whole cell-based ELISA method with mock-transfected cells as negative control (**Supplementary Figure 1**).

### **Overall structure of agonist-GPR109A-G-protein complexes**

We reconstituted the agonist-GPR109A-G-protein complexes using purified components following state-of-the-art methodology successfully applied to other GPCR-G-protein complexes (REF). We determined the cryo-EM structures of these complexes at estimated resolutions of 3.37 Å, 3.45 Å, 3.45 Å, 3.26 Å and 3.56 Å respectively, for the niacin, acipimox, MK6892, GSK256073 and MMF-bound receptor (**Figure 2A-E, Supplementary Figure 2-10**). The unambiguous densities of the coulombic maps enabled us to assign nearly the entire transmembrane domain of the receptor although the first seven residues at the N-terminus of the receptor and the last forty-five residues at the carboxyl-terminus were not resolved in the structures potentially due to their inherent flexibility (**Supplementary Figure 11**). Still however, in each of these complexes, the ligand densities were clearly discernible, allowing us to visualize ligand-receptor interactions, and the high map quality at the receptor-G-protein interface facilitated the identification of residue level interactions driving G-protein coupling to the receptor (**Supplementary Figure 12**). The precise sequence of the components resolved in these structures is listed in **Supplemental Figure 11**. The overall structures of GPR109A in all five complexes are highly similar with an RMSD of 0.6-1.0 Å<sup>2</sup> along the C $\alpha$  of the receptor interface (**Figure 3A**) and the key differences are observed in the ligand-receptor interaction as outlined in the sections below.



## **An extracellular lid in GPR109A**

Structural analysis of the receptor component in these structures uncovered several interesting features. For example, the N-terminus of the receptor in all five structures adopts a twisted  $\beta$ -hairpin structure that positions itself above the extracellular opening of the receptor (**Figure 3B**). Moreover, the Ile169-Leu176 segment in the ECL2 adopts a twisted antiparallel  $\beta$ -hairpin conformation while Ser178<sup>4,51</sup>-Ser181<sup>5,31</sup> dips down into the core of the TM bundle to form part of the orthosteric binding pocket (**Figure 3C, D**). Interestingly, the ECL2 hairpin interacts with the  $\beta$ -hairpin formed by the N-terminal residues Leu11-Cys18 to form a “lid-like” architecture that covers the extracellular opening of the receptor (**Figure 3E**). This observation can be attributed to the presence of three disulfide bridges (N-terminus Cys18 with Cys183 of ECL2, N-terminus Cys19 with Cys266 of TM7, and Cys177 of ECL2 with Cys100 of TM3) which helps to further stabilize the N-terminus-ECL2 lid (**Figure 3E**). These disulfide bonds might impose additional constraints towards the flexibility of the lid, and facilitate docking of the ligand within the orthosteric pocket of the receptor. It is interesting to note that a similar “lid-like” conformation adopted by the N-terminus and ECL2 has been previously reported for several GPCRs such as b2AR and CXCR4, removal of the disulfide constraints resulting in either complete loss or decreased agonist affinity in the receptors (REF).

## **Agonist-receptor interaction in the orthosteric binding pocket**

The ligand binding site in GPR109A is positioned approximately 20 Å deep in the receptor core (measured from N-terminal Phe21 to Phe244<sup>6,48</sup>) (**Figure 3F**), and all five agonists share a common interaction interface, at least in part, on the receptor where chemically similar moieties of the ligands are positioned (**Figure 3F**). An array of aromatic residues namely, Phe276<sup>7,35</sup>, Phe277<sup>7,36</sup>, Trp91<sup>ECL1</sup>, Phe180<sup>ECL2</sup> and Phe193<sup>5,43</sup>, and hydrophobic residues namely, Leu83<sup>2,60</sup>, Leu104<sup>3,29</sup> and Leu107<sup>3,32</sup> are found lining the orthosteric pocket of the receptor, and together, they form the microenvironment for the binding of the ligands. The interactions between niacin, acipimox, MK6892, GSK256073 and MMF with the receptor are mainly ionic, hydrophobic, and aromatic, including residues predominantly from TM2, TM3, TM7 and ECL2, and a complete list of interactions are listed in **Supplementary Figure 12**.

The comparison of the ligand binding pocket in all five structures reveals that Arg111<sup>3,36</sup> forms the most important residue for binding to the negatively charged carboxyl group of ligands,

niacin, acipimox, GSK256073, MK6892 and MMF through hydrogen bond (**Figure 3G**). Previous studies have suggested that this carboxyl group is critical for receptor activation, and substitution with an amide group abolishes GPR109A activity, and our structural snapshots provide a mechanistic basis for these functional observations (REF). Three more pairs of hydrogen bonds can be observed, one between the carboxyl moiety of niacin (or acipimox, GSK256073 and MMF) with the side chain Tyr284<sup>7,43</sup>, two between the chloride moiety of GSK256073 with the side chain Ser179<sup>45,52</sup> and backbone N-atom of Phe180<sup>ECL2</sup> (**Figure 3G**) and one between oxo-group at position 4 of MMF with Ser179<sup>45,52</sup>. Furthermore, activation of GPR109A appears to require a hydrophobic environment within the orthosteric pocket, and several hydrophobic contacts can be found to stabilize niacin and acipimox within the ligand binding pocket mediated by hydrophobic residues namely, Leu83<sup>2,60</sup>, Leu104<sup>3,29</sup>, Leu107<sup>3,32</sup>, Phe180<sup>ECL2</sup>, Phe277<sup>3,36</sup>, and Leu280<sup>7,39</sup> (**Figure 3I and Supplementary Figure 12**). Similarly, hydrophobic residues such as Leu30<sup>1,35</sup>, Trp91<sup>ECL1</sup>, Leu107<sup>3,32</sup>, and Phe180<sup>ECL2</sup> forms extensive interactions with GSK256073 (**Figure 3I**). Although niacin, acipimox and GSK256073 exhibit hydrophilic, hydrophobic, and charged properties that largely match with those of the ligand binding pocket (**Supplementary Figure 12**), slight conformational variation can be observed within the binding pocket for the niacin or acipimox and GSK256073 complexes. These conformational shifts can be attributed to the presence of the extra Cl moiety in GSK256073.

Like niacin, acipimox and GSK256073, the carboxyl group of MK6892 makes similar contacts with the surrounding polar and hydrophobic residues within the orthosteric binding pocket (**Figure 3G, H**). MK6892 has a relatively extended chemical structure compared to the other three agonists and therefore, it engages several additional residues in the receptor. For example, the extended moieties in MK6892 i.e., dimethyl, oxadiazole, and pyridyl groups interact with Gln112<sup>3,37</sup>, His189<sup>5,39</sup> and Met192<sup>5,42</sup> in an extended binding pocket in the receptor (**Figure 3J**). Interestingly, several conformational rearrangements in the side-chains of Arg111<sup>3,36</sup>, Gln112<sup>3,37</sup>, Ser179<sup>45,52</sup>, and Tyr284<sup>7,43</sup> are also observed compared to the other agonists in order to accommodate the bulky extended group of MK6892 (**Figure 3K**). Furthermore, an upward rotameric transition of His189<sup>5,39</sup> and an outward shift of Met192<sup>5,42</sup> is also observed within the extended binding pocket to prevent

steric clashes with the extended chain of MK6892 (**Figure 3J**). These additional interactions of MK6892 with the GPR109A are similar to those observed in a recent study<sup>25,26</sup>.

### **Agonist-induced activation of GPR109A**

When compared to the recently determined inactive state crystal structure of GPR109A, the niacin-activated GPR109A displayed the known conformational changes i.e., the cytoplasmic side of TM6 exhibits an outward movement of ~4 Å (measured from the Cα of Lys227) and 5.5 Å inward movement of TM5 towards the extracellular side (measured from the Cα of His189<sup>5,39</sup>) and about 4.5 Å outward movement towards the cytoplasmic side (measured from the Cα of Arg218<sup>ICL3</sup>) (**Figure 4A-C**). The agonist-bound structures of GPR109A exhibit the typical hallmark movements of receptor activation as reflected by the conserved motifs and microswitches. For example, the “P-I-F motif” consisting of Pro200<sup>5,50</sup>, Ile115<sup>3,40</sup> and Phe240<sup>6,44</sup> forms an interface at the base of the ligand binding pocket, and it undergoes conformational rearrangements upon receptor activation. The rearrangements include: (i) rotameric shift of Pro200<sup>5,50</sup>, (ii) rotameric flip of Ile115<sup>3,40</sup> and (iii) large transition of Phe240<sup>6,44</sup>, thus opening the cytoplasmic core of the receptor for the interaction with the α5 residues of Gαo (**Figure 4D**). Similar conformational changes can be observed with respect to the “D-R-Y” and “NPxxY” microswitches as well. Asp124<sup>3,49</sup>, Arg125<sup>3,50</sup> and Tyr126<sup>3,51</sup> in TM3 is a highly conserved motif where Asp124<sup>3,49</sup> forms a salt bridge with Arg125<sup>3,50</sup>, thus locking the receptor in an inactive conformation. A rotameric shift of Arg125<sup>3,50</sup> can be observed in the ligand-bound structures, facilitating the breaking of the salt-bridge/ionic lock and transition to its active conformation (**Figure 4D**). A variant of the “NPxxY” motif is present in GPR109A, where Asn290<sup>7,49</sup> is substituted with Asp290<sup>7,49</sup> in TM7. Upon activation, the lower portion of TM7 moves inwards towards the receptor core combined with a rotation of Tyr294<sup>7,53</sup> along the helical axis (**Figure 4D**).

### **The interface of GPR109A-G-protein interaction**

As mentioned earlier, significant movements of TM5, TM6 and TM7 create an opening on the cytoplasmic surface of the receptor that allows the docking of the α5 helix of Gαo leading to coupling of G-proteins with the activated receptor (**Figure 5A, C, E, G, I**). Expectedly, we observe a large buried surface area at the interface of the receptor and G-protein nearing almost 2000 Å<sup>2</sup> as typically observed in GPCR-G-protein complexes, and this is almost identical in all five

structures of GPR109A reported here (**Figure 5A, C, E, G, I**). The GPR109A-G-protein interface is stabilized by extensive hydrophobic and polar interactions between the TM2, TM3, ICL2, ICL3, TM6, TM7 and H8 in the receptor and the  $\alpha 5$  helix of Gao (**Figure 5B, D, F, H, J**). Specifically, Tyr354 at the carboxyl-terminus of Gao forms a key residue that is positioned in pocket on the cytoplasmic side of the receptor lined by Lys225<sup>6.25</sup>, Ile226<sup>6.30</sup> and Pro299<sup>8.48</sup> (**Figure 5B**). In addition, several hydrogen bonds between Asp341, Asn347 and Gly352 of Gao with Arg218<sup>ICL3</sup>, Arg128<sup>3.53</sup> and Ser297<sup>7.56</sup> of the GPR109A, respectively, further stabilize the interaction (**Figure 5B, D, F, H**). Finally, the stretch from Ala135 to Lys138 in the ICL2 of the receptor adopts a one-turn helix where His133 interacts with Thr340 and Ile342 which lie within a hydrophobic pocket formed by the residues from  $\alpha 5$  helix,  $\alpha N$ - $\beta 1$  loop and  $\beta 2$ - $\beta 3$  loop of Gao (**Figure 5F**). The receptor-G protein interface is further stabilized by residues of ICL3 with  $\alpha 5$  C-terminal loop and  $\alpha 4$ - $\beta 6$  loop of Gao, viz. Arg218<sup>ICL3</sup> forms extensive interactions with Thr340, Asp341 and Ile344 of Gao (**Figure 5B, D, F, H**). A list of ligands bound-GPR109A residues interacting with Gao is presented in **Figure 5K-L**, which underscores a largely conserved interface for G-protein interaction although some ligand-specific interactions are also observed. A comprehensive detail of the interactions between GPR019A and G-proteins are listed in **Supplementary Figure S13**.

### Structure-guided design of receptor inactivation and biased-agonism

As mentioned above, there were two key interactions in the ligand binding pocket namely the Arg111 in TM3 and Ser179 in ECL2 that appeared to be conserved in all five structures involved in a direct hydrogen bonding with the ligands (**Figure 6A**). Therefore, we generated Arg111<sup>3.36</sup>Ala and Ser179<sup>ECL2</sup>Ala mutants of the receptor and measured niacin-induced G-protein and  $\beta$ -arrestin-coupling vis-à-vis the wild-type receptor. These mutants expressed at comparable levels as the wild-type receptor (**Supplementary Figure 14**). Interestingly, we observed that R111<sup>3.36</sup>A mutant exhibited complete loss of G-protein activation as measured using G-protein dissociation and cAMP assay, and agonist-induced  $\beta$ -arrestin recruitment (**Figure 6B-D**). This may reflect a near-complete loss of niacin binding to the receptor mutant as reported previously using a radioligand binding assay<sup>2</sup>. On the other hand, S179<sup>ECL2</sup>A mutation resulted in a significant reduction in  $\beta$ -arrestin recruitment in terms of Emax (1.61 fold vs. 1.47 fold for WT and S179<sup>ECL2</sup>A) and EC50 (32.0 $\pm$ 4.90 nM vs. 3.05 $\pm$ 65 nM for WT and S179<sup>ECL2</sup>A) (**Figure 6D**), while exhibiting slightly



improved G-protein-coupling (Emax 38% vs. 53%, EC<sub>50</sub> 58.7±1.68 nM vs 9.90±1.42 nM for WT and S179<sup>ECL2A</sup> in Go dissociation assay, **Figure 6B**; Emax 25.78 vs. 64.95 in EC<sub>50</sub> 2.59±125 nM vs. 9.90±1.42 nM for WT and S179<sup>ECL2A</sup> in cAMP response assay, **Figure 6C**). Therefore, GPR109A<sup>Ser179Ala</sup> mutant represents a G-protein-biased construct (**Figure 6E-F**), and it may be a useful tool to further probe ligand-bias at this receptor. Taken together, these data demonstrate the feasibility of structure-guided engineering of receptor inactivation and biased agonism at the transducer-coupling response, which may facilitate a framework to better understand the mechanistic aspects of biased agonism going forward.

## Discussion

GPR109A continues to be an important drug target for developing therapeutic agents for dyslipidemia with properties superior to the commonly prescribed drug, niacin, especially in terms of reducing the side-effect of flushing response. While niacin, acipimox and MMF bind in a similar pose in the orthosteric binding pocket and make nearly-identical interactions, GSK256073 and MK6892 make additional contacts in the ligand binding pocket as expected based on their extended chemical structures. Interestingly, both GSK256073 and MK6892 appear to exhibit enhanced G-protein-coupling compared to niacin. However, GSK256073 is slightly more efficacious and potent in  $\beta$ -arrestin recruitment assay while MK6892 is slightly weaker than niacin in  $\beta$ -arrestin recruitment assay. Based on these data, it is tempting to speculate that additional contacts engaged by MK6892 make it a more potent and efficacious agonist of GPR109A compared to niacin, although follow-up experimentation would be required to test this hypothesis. As mentioned earlier, a previous study had reported the  $\beta$ -arrestin-mediated signaling to be the driver of flushing response for niacin while G-protein signaling is responsible for the lipid lowering effect. The G-protein-bias of MK6892 in  $\beta$ -arrestin recruitment assay therefore likely explains its non-flushing properties as reported earlier. However, the transducer-coupling profile of GSK256073 does not align with the same hypothesis, suggesting that the segregation of lipolysis vs. flushing response through GPR109A may involve additional fine-tuning that remains to be explored further.

Our mutagenesis studies guided by structural visualization of the key interactions between the agonists and the receptor yield a mutant that fails to elicit any transducer-coupling, and another mutant that maintains G-protein-coupling but loses  $\beta$ -arrestin binding. Although we have tested the

effects of these mutations on only niacin-induced receptor activation and signaling, considering their conserved nature in terms of interaction with other agonists, it is likely that such mutants will exhibit a similar pattern for other agonists as well. It is intriguing that Ser179Ala mutation in ECL2 results in impaired  $\beta$ -arrestin recruitment as the site is closer to the orthosteric binding pocket in the receptor and significantly away from the interface of  $\beta$ -arrestin coupling to prototypical GPCRs. Therefore, it is tempting to speculate that the effect observed for Ser179<sup>45,52</sup>Ala mutation imparts a reduction in  $\beta$ -arrestin interaction through an allosteric mechanism as reported for other GPCRs previously<sup>27,28</sup>. In addition, the experimental framework established here should also facilitate the structure determination of other subtypes namely, HCA1 and HCA3, and a complete structural coverage of all three receptor subtypes should allow us to better understand the sub-type selectivity for niacin and other ligands.

In summary, the structural snapshots of GPR109A presented here elucidate the molecular details of the interaction of chemically-diverse agonists, and also uncover the mechanism of activation of this therapeutically important receptor. Our findings should pave the way for rational therapeutic design targeting GPR109A, and they also provide a framework to impart signaling bias in GPCRs guided by structural insights that may help deconvolute the mechanism of biased agonism going forward.

### **Data availability statement**

All the data are included in the manuscript and any additional information required to reanalyze the data reported in this paper is available from the corresponding author upon reasonable request.

### **Acknowledgements**

Research in A.K.S.'s laboratory is supported by the Senior Fellowship of the DBT Wellcome Trust India Alliance (IA/S/20/1/504916) awarded to A.K.S., Science and Engineering Research Board (IPA/2020/000405), Young Scientist Award from Lady Tata Memorial Trust, and IIT Kanpur. Cryo-EM was performed at BioEM lab of the Biozentrum at the University of Basel, and we thank Carola Alampi and David Kalbermatter for their excellent technical assistance.

### **Authors' contribution**

MKY expressed and purified GPR109A, and reconstituted the receptor-G-protein complexes with help from VS and GM; PS carried out the pharmacological and cellular assays on GPR109A with

help from SM, AD and NZ; SS purified mini-Gao and ScFv16 with help from SaS, RB performed negative-staining EM, processed the cryo-EM data, prepared and deposited the coordinates with help from JM, and drafted the figures together with MG; MC screened the samples and collected cryo-EM data; AKS supervised and managed the overall project; all authors contributed to data analysis, interpretation and manuscript writing.

### **Conflict of interest**

The authors declare that they have no competing interests.

### **Accession number**

The cryo-EM maps and structures have been deposited in the EMDB and PDB with accession numbers 8IYP and EMD-35817 for niacin-GPR109A-Go, 8JER and EMD-36193 for acipimox-GPR109A-Go, 8IYW and EMD-35831 for GSK256073-GPR109A-Go, 8IYH and EMD-35822 for MK6892-GPR109A-Go and EMD-36280, PDB ID: 8JHN for MMF-GPR109A-Go complex.

### **Materials and methods**

#### **General reagents, plasmids, and cell culture**

The majority of standard reagents were purchased from Sigma Aldrich unless mentioned. Dulbecco's Modified Eagle's Medium (DMEM), Phosphate Buffer Saline (PBS), Trypsin-EDTA, Fetal-Bovine Serum (FBS), Hank's Balanced Salt Solution (HBSS), and Penicillin-Streptomycin solution were purchased from Thermo Fisher Scientific. HEK-293 cells were purchased from ATCC and maintained in 10% (v/v) FBS (Gibco, Cat. no. 10270-106) and 100U ml<sup>-1</sup> penicillin and 100 µg ml<sup>-1</sup> streptomycin (Gibco, Cat. no. 15140122) supplemented DMEM (Gibco, Cat. no. 12800-017) at 37 °C under 5% CO<sub>2</sub>. The cDNA coding region of GPR109A<sup>WT</sup>, GPR109A<sup>Y87A</sup>, GPR109A<sup>R111A</sup>, GPR109A<sup>S179A</sup>, and GPR109A<sup>Y284A</sup> with a HA signal sequence, a FLAG tag followed by the N-terminal region of M4 receptor (2-23 residues) at the N-terminus was cloned into pcDNA3.1 vector. For GloSensor assay, luciferase-based 22F cAMP biosensor construct was purchased from Promega. For the constructs used in NanoBiT assay, SmBiT was fused at the C-terminus of the receptor, and the LgBiT-βarr1/2 construct was the same as previously described<sup>29</sup>. All DNA constructs were verified by sequencing from Macrogen. Niacin was purchased from Himedia (Cat. no. TC157), acipimox and MMF were purchased from Sigma Aldrich (Cat. no: 92571 and Cat. no:

651419), respectively. GSK256073 and MK9862 were purchased from MedChemExpress (Cat. no: HY10680 and HY119222, respectively).

### **GPR109A purification**

Codon-optimized human GPR109A was cloned in the pVL1393 vector with an N-terminal HA signal sequence followed by a FLAG tag and M4 receptor N-terminal sequence for increased expression. The receptor was expressed and purified from *Spodoptera frugiperda* (Sf9) cells using a baculovirus-mediated expression system. For receptor purification, insect cells were infected with recombinant baculovirus for 72 hrs at 27 °C and harvested by high-speed centrifugation. Post-harvest, insect cells were sequentially dounced in hypotonic buffer (20 mM HEPES, pH 7.4, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM PMSF, and 2 mM Benzamidine), hypertonic buffer (20 mM HEPES, pH 7.4, 1 M NaCl, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM PMSF, and 2 mM Benzamidine) and lysis buffer (20 mM HEPES, pH 7.4, 450 mM NaCl, 1 mM PMSF, and 2 mM Benzamidine). Lysed cells were solubilized by continuous stirring in 1% L-MNG (Anatrace, Cat. no. NG310) for two hours at 4 °C in the presence of 0.01% cholesteryl hemisuccinate (Sigma, Cat. no. C6512). To prevent receptor aggregation, 2 mM Iodoacetamide was added to the solution. Post-solubilization, salt concentration was lowered to 150 mM with dilution buffer (20 mM HEPES, pH 7.4, 2 mM CaCl<sub>2</sub>, 1 mM PMSF, and 2 mM Benzamidine), and cell debris was separated by high-speed centrifugation, and the receptor was enriched on M1-anti FLAG columns. Non-specific proteins were removed by three washes of low salt buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 2 mM EDTA, 0.01% MNG, 0.01% CHS, and 2 mM CaCl<sub>2</sub>) alternating with two washes of high salt buffer (20 mM HEPES, pH 7.4, 350 mM NaCl, 0.01% MNG, and 2 mM CaCl<sub>2</sub>). The bound receptor was eluted with FLAG peptide-containing buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 2 mM EDTA, and 250 µg ml<sup>-1</sup> flag peptide). The purified receptor was treated with two rounds of 2 mM Iodoacetamide, followed by one round of 2 mM cysteine treatment to remove free Iodoacetamide. The purified receptor was concentrated using a 30 kDa MWCO concentrator (Cytiva, Cat no. 28932361) and stored at -80 °C with a 10% final glycerol concentration. 1 µM of niacin, acipimox, MK6892, GSK256073 or MMF was kept throughout the purification.

### **Purification of Gβ1γ2 dimer**



N terminal 6X His tagged Gβ1 and Gy2 subunits were cloned in the Dual pVL1392 vector and expressed in the *Sf9* cell using a baculovirus-based expression system. For purification, insect cells were infected with the recombinant virus for 72 hrs and harvested with high-speed centrifugation. Cells were lysed by douncing in lysis buffer (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 1 mM PMSF, and 2 mM Benzamidine) and pelleted by centrifugation at 20,000 rpm for 20 min. The cell pellet was re-dissolved and dounced in solubilization buffer (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, 1% DDM (Anatrace, Cat. no. D310), 1 mM PMSF, and 2 mM Benzamidine). Lysed cells were solubilized for two hrs at 4 °C with constant stirring. Cell debris was separated by high-speed centrifugation, and protein was passed through the Ni-NTA column using gravity flow. Non-specific proteins were removed by a one-column wash with buffer (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 0.01% MNG), and protein was eluted with 250 mM Imidazole (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 0.01% MNG, 250 mM Imidazole). Eluted protein was concentrated with a 10 kDa MWCO concentrator (Cytiva Cat. no. 28932360) and stored with 10% final glycerol concentration.

### **Mini Gao purification**

The gene encoding miniGao was designed as described previously<sup>30,31</sup> and cloned into pET-15b (+) vector with 6X His tag at the N-terminal followed by TEV protease cleavage site. The recombinant construct was transformed into *E. coli* BL21(DE3) cells. A 5 ml starter culture, grown for 6-8h at 37 °C, was inoculated into a 50 ml primary culture media supplemented with 0.2% glucose and allowed to grow at 30 °C for 16-18 hrs. 1.5 litre of TB (Terrific Broth) media was inoculated with 15 ml of primary culture and grown at 30 °C. At O.D<sub>600</sub> 0.8, cells were induced with 50 μM IPTG (isopropylthio-β-galactoside) and allowed to grow for an additional 18–20 hrs. Cells were pelleted down and first lysed by lysozyme in lysis buffer (40 mM HEPES, pH 7.4, 100 mM NaCl, 10 mM Imidazole, 10% Glycerol, 5 mM MgCl<sub>2</sub>, 1 mM PMSF, 2 mM Benzamidine, 50 μM GDP, 100 μM DTT, and 1 mg ml<sup>-1</sup> lysozyme), followed by disruption by ultrasonication. Cell debris was pelleted by high-speed centrifugation at 4 °C, and protein was enriched on the Ni-NTA column. Non-specifically bound proteins were removed by extensive washing with wash buffer (20 mM HEPES, pH 7.4, 500 mM NaCl, 40 mM Imidazole, 10% Glycerol, 50 μM GDP, and 1 mM MgCl<sub>2</sub>), and protein was eluted with 500 mM Imidazole (in 20 mM HEPES, pH 7.4, 100 mM NaCl, and 10%

Glycerol). His tag was cleaved by overnight TEV treatment at room temperature (1:20, TEV: protein), and untagged protein was recovered by size exclusion chromatography on Hi-Load Superdex 200 PG 16/600 column (Cytiva, Cat. no. 17517501). Fractions corresponding to cleaved protein were pooled, analyzed on SDS-PAGE, and stored at -80 °C with 10% glycerol.

### **ScFv16 purification**

The gene encoding ScFv16 was cloned in pET-42a (+) vector downstream of 10X His tagged MBP gene with a TEV protease cleavage site between them and overexpressed in the *E. coli* Rosetta (DE3) strain<sup>32</sup>. A single colony from a freshly transformed plate was inoculated in 50 ml of 2XYT media and allowed to grow overnight at 37 °C. 1-litre 2XYT media supplemented with 0.5% glucose and 5 mM MgSO<sub>4</sub> was inoculated with overnight primary culture and induced with 250 µM IPTG at O. D<sub>600</sub> of 0.8–1.0. The culture was allowed to grow for 16–18 hrs at 18 °C. Post-harvest, cells were resuspended in 20 mM HEPES, pH 7.4, 200 mM NaCl, 30 mM Imidazole, 1 mM PMSF, and 2 mM Benzamidine buffer and incubated at 4 °C for 40 min with constant stirring. Cells were lysed by sonication, and cell debris was removed with high-speed centrifugation at 4 °C. Protein was captured on the Ni-NTA column using gravity flow, and non-specific proteins were removed by extensive washing with 20 mM HEPES, pH 7.4, 200 mM NaCl, and 50 mM Imidazole. Bound ScFv16 was eluted with 300 mM imidazole-containing buffer (20 mM HEPES, pH 7.4, 200 mM NaCl) and was re-passed through amylose resins, and after one column wash with 20 mM HEPES, pH 7.4, 200 mM NaCl buffer, bound protein was eluted with 10 mM maltose (prepared in 20 mM HEPES, pH 7.4, 200 mM NaCl). To obtain tag-free ScFv16, the eluted protein was overnight digested with TEV, and His-MBP was removed by passing the digested protein through the Ni-NTA column. Eluted protein was further cleaned by size exclusion chromatography on Hi-Load Superdex 200 PG 16/600 column. Eluted protein was analyzed on SDS-PAGE and stored at -80 °C with 10% glycerol.

### **Reconstitution of GPR109A-G protein-ScFv16 complexes**

Purified GPR109A was mixed with a 1.2 molar excess of miniGo, Gβγ, and ScFv16 in the presence of 25 mU ml<sup>-1</sup> apyrase (NEB, Cat. no. M0398S) and 1 µM of individual ligand, and complexing was allowed for two hours at room temperature. The receptor complex was concentrated with a 100 kDa MWCO (Cytiva, Cat. no. GE28-9323-19) concentrator and separated from the unbound

component by size exclusion chromatography on Superose 6 increase 10/300 GL column (Cytiva, Cat. no. 29091596). The SEC eluate was analyzed on 12% SDS-PAGE, and complex fractions were concentrated to  $\sim 10 \text{ mg ml}^{-1}$  and stored at  $-80^\circ\text{C}$ .

### **Negative stain electron microscopy**

Homogeneity of the purified protein complexes was determined through negative staining with uranyl formate prior to data collection under cryogenic conditions following the protocols described previously<sup>33,34</sup>. 3.5  $\mu\text{l}$  of the purified complexes were dispensed onto fresh glow discharged carbon/formvar coated 300 mesh Cu grids (PELCO, *Ted Pella*) at a concentration of  $0.02 \text{ mg ml}^{-1}$  and incubated for 1 min at room temperature. This was followed by blotting of the excess samples from the grids using filter paper. The grid containing the adhered sample was touched onto a first drop of freshly prepared 0.75% uranyl formate stain and immediately blotted off by touching the edge of the grid onto a filter paper. The grid was then touched and incubated on a second drop of uranyl formate for 30s and left on the bench in a Petri plate for air drying. Data collection was performed on a FEI Tecnai G2 12 Twin TEM (LaB6) operating at 120 kV and equipped with a Gatan CCD camera (4k x 4k) at 30,000x magnification. Data processing of the collected micrographs was performed with Relion<sup>35</sup> 3.1.2 version. More than 10,000 particles were autopicked with the gaussian blob picker, extracted and subjected to reference-free 2D classification.

### **Cryo-EM sample preparation and data collection**

3  $\mu\text{l}$  of the individual complexes were applied onto glow-discharged Quantifoil holey carbon grids (Cu R2/1 or R2/2) and vitrified in liquid ethane ( $-181^\circ\text{C}$ ) using a Leica GP plunger (Leica Microsystems, Austria) maintained at 90% humidity and  $10^\circ\text{C}$ . CryoEM movies were acquired on a TFS Glacios microscope operating at 200 kV and equipped with Gatan K3 direct electron detector (Gatan Inc.). Images were collected automatically with SerialEM software in counting mode at a nominal magnification of 46,000x and pixel size of  $0.878 \text{ \AA}$  over a defocus range of  $0.5\text{-}2.5 \mu\text{m}$ . An accumulated dose of  $55 \text{ e}^-/\text{\AA}^2$  was fractionated into a movie stack consisting of 40 frames.

### **Cryo-EM data processing**

459 All data processing steps were performed with cryoSPARC<sup>36</sup> v4.0 unless otherwise stated. Dose  
460 fractionated movie stacks were subjected to beam-induced motion correction using Patch motion  
461 correction (multi) followed by estimation of contrast transfer function parameters with Patch CTF  
462 estimation (multi).

463 For the Niacin-GPR109A-Go dataset, 11,070 dose weighted, motion-corrected micrographs  
464 were selected for downstream processing. Auto-picking yielded 7,027,107 particles which were  
465 subjected to several rounds of reference-free 2D classification to eliminate particles with poor  
466 features. 1,737,725 particle projections corresponding to the 2D averages with clear secondary  
467 features were selected and subjected to Ab-initio reconstruction with 3 classes. Subsequent  
468 heterogeneous refinement yielded a model with features of a typical GPCR-G protein complex  
469 containing 1,011,301 particle projections which accounted for 75% of the particles used for  
470 heterogeneous refinement. This particle stack was subjected to non-uniform refinement, followed  
471 by local refinement with mask excluding the noise outside the molecule, yielding a coulombic map  
472 with an indicated global resolution of 3.37 Å at 0.143 FSC cut-off.

473 For the MK6892-GPR109A-Go dataset, 90,683,101 particles were autopicked from 10,753  
474 motion corrected micrographs which were extracted with a box size of 360 px (fourier cropped to  
475 64 px) and subjected to multiple rounds of reference-free 2D classification. 2D class averages  
476 consisting of 1,829,840 particles with clear secondary features and resembling conformation of  
477 protein complexes were re-extracted with a box size of 360 px and fourier cropped to 288 px.  
478 These particle stacks from the extraction job were subsequently subjected to Ab-initio  
479 reconstruction, followed by heterogeneous refinement yielding 4 models. 1,200,513 particle  
480 projections (accounting for 66% of the total particles) from the best 3D class were selected and  
481 subjected to non-uniform refinement, followed by local refinement, which yielded a map with an  
482 overall resolution of 3.26 Å using the 0.143 FSC criterion.

483 For the GSK256073-GPR109A-Go dataset, 5,761,414 particles were automatically picked  
484 from 10,574 motion-corrected micrographs. These particle projections were extracted with a box  
485 size of 360 px (fourier cropped to 64 px) and subjected to iterative rounds of reference-free 2D  
486 classification to discard noisy particles. 1,601,694 particles corresponding to the 2D classes with



evident features of protein complexes were selected and re-extracted with a box size of 360 px and fourier cropped to 288 px. These selected particle projections were used to generate 3 maps for heterogeneous refinement. One of the 3D classes with 523,816 particles showing all the features of a GPCR-G protein complex was subjected to 3D non-uniform refinement, reaching a nominal resolution of 3.45 Å.

For the Acipimox-GPR109A-Go dataset, 9,115,816 particles were autopicked from 11,263 micrographs, extracted with a box size of 360 px (fourier cropped to 64 px) followed by 2D classification to obtain classes with clear secondary features. 166,548 particles corresponding to the clean classes were re-extracted with a box size of 360 px (fourier cropped to 288) and subjected to ab-initio reconstruction and heterogenous refinement to generate 4 classes. 1,059,994 particles from the best 3D class were selected and subjected to non-uniform refinement followed by local refinement with mask to yield a final reconstruction at a resolution of 3.45 Å.

For the MMF-GPR109A-Go dataset, autopicking was performed with the blob-picker subprogram which yielded 9,029,435 particles. Particles were extracted with a box size of 360 px (fourier cropped to 64 px) and pared down to 183,241 particles after reference-free 2D classification. The clean particle stack was then re-extracted with a box size of 360 px (fourier cropped to 288 px). Two rounds of ab initio and subsequent hetero-refinement (using four models) were then performed to further refine the particle stack to 678,286 particles. Non-uniform refinement and successive local refinement resulted in a map with an estimated global resolution of 3.56 Å.

Local resolution estimation of all maps was determined using the Blocres subprogram within cryoSPARC with the corresponding half maps. Sharpening of all maps was performed with “Autosharpen maps” within the Phenix suite<sup>37,38</sup> to enhance features for model building.

## **Model building and refinement**

Coordinates from an AlphaFold model of GPR109A (AF-Q8TDS4-F1) was used to dock into the EM density map of niacin-GPR109A-Go using Chimera<sup>39,40</sup>. Similarly, coordinates of Gao, Gβγ and ScFv16 were obtained from a previously solved structure of C5aR1 in complex with Gao (PDB: 8HPT). The combined model so obtained containing all the components was subjected to “all atom

refine” sub-module within COOT<sup>41</sup>, followed by manual rebuilding of the residues and the ligand. The rebuilt model was subjected to real space refinement in Phenix<sup>37,38</sup> to obtain a model with 97.13% of the residues in the most favoured region and 2.77% in the allowed region of the Ramachandran plot. Validation of all the models was performed with Molprobrity<sup>42</sup> within Phenix.

The ligand free model of niacin-GPR109A-Go complex (PDB: 8IY9) was fitted into the density maps of acipimox-GPR109A-Go and MMF-GPR109A-Go in Chimera followed by flexible fitting of the coordinates with the “all atom refine” module in COOT. After several rounds of manual adjustments, the generated model was automatically refined with Phenix\_refine. The final refined models of acipimox-GPR109A-Go and MMF-GPR109A-Go showed good Ramachandran statistics with 96.86% and 97.31% in the most favored regions of the Ramachandran plot, respectively.

The ligand-free model of niacin-GPR109A-Go complex (PDB: 8IY9) was fitted into the density maps of acipimox-GPR109A-Go and MMF-GPR109A-Go in Chimera followed by flexible fitting of the coordinates with the “all atom refine” module in COOT. After several rounds of manual adjustments, the generated model was automatically refined with Phenix\_refine. The final refined models of acipimox-GPR109A-Go and MMF-GPR109A-Go showed good Ramachandran statistics with 96.86% and 97.31% in the most favored regions of the Ramachandran plot, respectively.

Likewise, the ligand-free model of niacin-GPR109A-Go complex (PDB: 8IY9) was used to dock into the coulombic maps of MK6892-GPR109A-Go and GSK256073-GPR109A-Go using Chimera. The docked model and the corresponding maps were then imported into COOT and fitted into the respective maps with the “all atom refine” module. The poorly fitted regions were manually adjusted in COOT followed by iterative refinement of the coordinates against the maps using Phenix\_refine. The final refined models of MK6892-GPR109A-Go and GSK256073-GPR109A-Go contained residues in 97.67% and 97.14% of the most favored regions of the Ramachandran plot with no outliers. Data collection, processing and model refinement statistics are provided in **Supplementary Figure 8**. All figures included in the manuscript have been prepared with Chimera and ChimeraX software.

### **GloSensor-based cAMP assay**

cAMP response upon ligand stimulation was measured by GloSensor assay<sup>43</sup>. Briefly, HEK-293 cells were transiently transfected with 2 µg of GPR109A construct together with 5 µg 22F cAMP

plasmid using the transfection reagent polyethylenimine (PEI) linear (Polysciences, Cat. no. 23966) at DNA: PEI ratio of 1:3. After 16-18hrs of transfection, cells were harvested followed by resuspension of the cell pellet in assay buffer composed of 1X HBSS, 20 mM of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.4 and D-luciferin (0.5 mg ml<sup>-1</sup>) (GoldBio, Cat. no.: LUCNA-1G). Harvested cells were then seeded in an opaque flat bottom white 96 well cell culture plate (SPL life sciences, Cat. no. 30196) at a density of 2x10<sup>5</sup> cells well<sup>-1</sup>. After seeding, cells were incubated at 37 °C for 90 min and 30 min at room temperature. After 120 min of incubation basal level luminescence was recorded using a multi-mode plate reader (Lumistar/Fluostar microplate reader, BMG Labtech). In order to record ligand-induced cAMP decrease as a readout of Gi activation, cellular cAMP level was increased by adding 5 μM forskolin, and 7-8 cycles luminescence was recorded until the signal got saturated. Once the luminescence signal got stabilized cells were stimulated with corresponding ligands, and luminescence values were recorded for 20 cycles. For stimulation, ligand concentrations ranging from 100 pM to 10 μM were prepared by serial dilution in the buffer constituted of 1X HBSS, 20 mM HEPES pH 7.4. Nicotinic acid, acipimox, GSK256073, MK6892 and MMF of different concentrations were added to the corresponding wells. Baseline corrected data were normalized with respect to the luminescence signal of minimal concentration of each ligand as 100% and plotted using nonlinear regression analysis in GraphPad Prism v 9.5.0 software.

### **Surface expression assay**

Plasma membrane expression of receptors in respective assays was measured by whole cell-based surface ELISA as previously discussed<sup>44</sup>. Briefly, transfected cells were seeded at a density of 2x10<sup>5</sup> cells well<sup>-1</sup> in 0.01% poly-D-Lysine pre-treated 24-well plate and incubated for 24 h at 37 °C. Post incubation, growth media was aspirated, and cells were washed with ice-cold 1X TBS for once, followed by fixation with 4% PFA (w/v in 1X TBS) on ice for 20 min. Post fixation, cells were washed three times with 1X TBS (400 μl in each wash) followed by blocking with 1% BSA (w/v in 1X TBS) at room temperature for 90 min. After blocking with 1% BSA, 200 μl anti-FLAG M2-HRP was added and incubated for 90 min (prepared in 1% BSA, 1:10,000) (Sigma, Cat. no. A8592). Post antibody incubation, to remove unbound antibodies, cells were washed with 1% BSA (prepared in 1X TBS) three times, followed by the development of signal by treating cells with 200

µl TMB-ELISA (Thermo Scientific, Cat no. 34028) until the light blue colour appeared. Signal was quenched by transferring the light blue-coloured solution to a 96-well plate containing 100 µl 1M H<sub>2</sub>SO<sub>4</sub>. The absorbance of the signal was measured at 450 nm using a multi-mode plate reader. Next, cells were incubated with 0.2% Janus Green (Sigma; Cat. no. 201677) w/v for 15 min after removal of TMB-ELISA by washing once with 1X TBS. Afterwards, Janus Green was aspirated followed by washing with distilled water to remove the excess stain. After washing, 800 µl of 0.5 N HCl was added to elute the stain. 200 µl of the eluate was transferred to a 96-well plate, and at 595 nm absorbance was recorded. For analysis, data were analyzed by calculating the ratio of absorbance at 450/595 followed by normalizing the value of pcDNA transfected cells reading as 1. Normalized values were plotted using GraphPad Prism v 9.5.0 software.

### **NanoBiT-based βarr recruitment assay**

Plasma membrane localization of βarr upon stimulation of GPR109A with respective ligands was measured by luminescence-based enzyme-linked complement assay (NanoBiT-based assay) following the protocol described earlier<sup>29,33,45</sup>. Briefly, a receptor harbouring SmBiT at the carboxy-terminus (3.5 µg) and βarr1/2 constructs (3.5 µg) with N-terminally fused LgBiT were co-transfected in HEK-293 cells using the transfection reagent polyethyleneimine (PEI) linear at DNA: PEI ratio of 1:3. Post 16-18 hr of transfection, cells were trypsinized, and resuspended in the NanoBiT assay buffer containing 1X HBSS, 0.01% BSA, 5 mM HEPES pH 7.4, and 10 µM coelenterazine (GoldBio, Cat. no. CZ05). Cells were then seeded in opaque flat bottom white 96 well plate at a density of 1x10<sup>5</sup> cells well<sup>-1</sup> and incubated for 120 min (90 min at 37 °C, followed by 30 min at room temperature). Post incubation, basal level luminescence readings were taken, followed by ligand addition. A series of ligand concentrations, spanning from 10 pM to 10 µM, were prepared using a buffer solution composed of 1X HBSS and 5 mM HEPES at pH 7.4. Subsequently, cells were stimulated with different doses of the specified ligands. Luminescence upon stimulation was recorded up to 20 cycles by a multi-mode plate reader. For analysis, stimulated readings were normalized with respect to the signal of minimal ligand concentration as 1 and plotted using nonlinear regression analysis in GraphPad Prism v 9.5.0 software.

### **NanoBiT-based G-protein dissociation assay**

Agonist-induced G-protein activation was measured by a nanoBiT-based G-protein dissociation assay described previously<sup>33</sup>. Briefly, HEK-293 cells were transfected with 1 µg of LgBiT-tagged Gα subunit, 4 µg of SmBiT-tagged Gγ2 subunit, 4 µg of untagged Gβ1 subunit along with 1 µg of untagged receptor construct using transfection reagent PEI at DNA: PEI ratio of 1:3. Post transfection, cells were harvested and seeded in a 96 well plate at a density of 1x10<sup>5</sup> cells well<sup>-1</sup>. Cells were seeded in buffer containing 1X HBSS, 0.01% BSA, 5 mM HEPES pH 7.4, and 10 µM coelenterazine and incubated for 120 min (90 min at 37 °C and 30 min at room temperature). Post incubation, 3 cycles of basal level luminescence readings were recorded using a multi-mode plate reader. After that, cells were stimulated with varying ligand concentrations ranging from 10 pM to 10 µM. After stimulation, 20 cycles of luminescence were recorded. For data analysis, values after 15 min of stimulation were used and normalized with respect to the signal at the minimal ligand concentration of 100%. Normalized values were plotted using nonlinear regression analysis in GraphPad Prism v 9.5.0 software.

## References

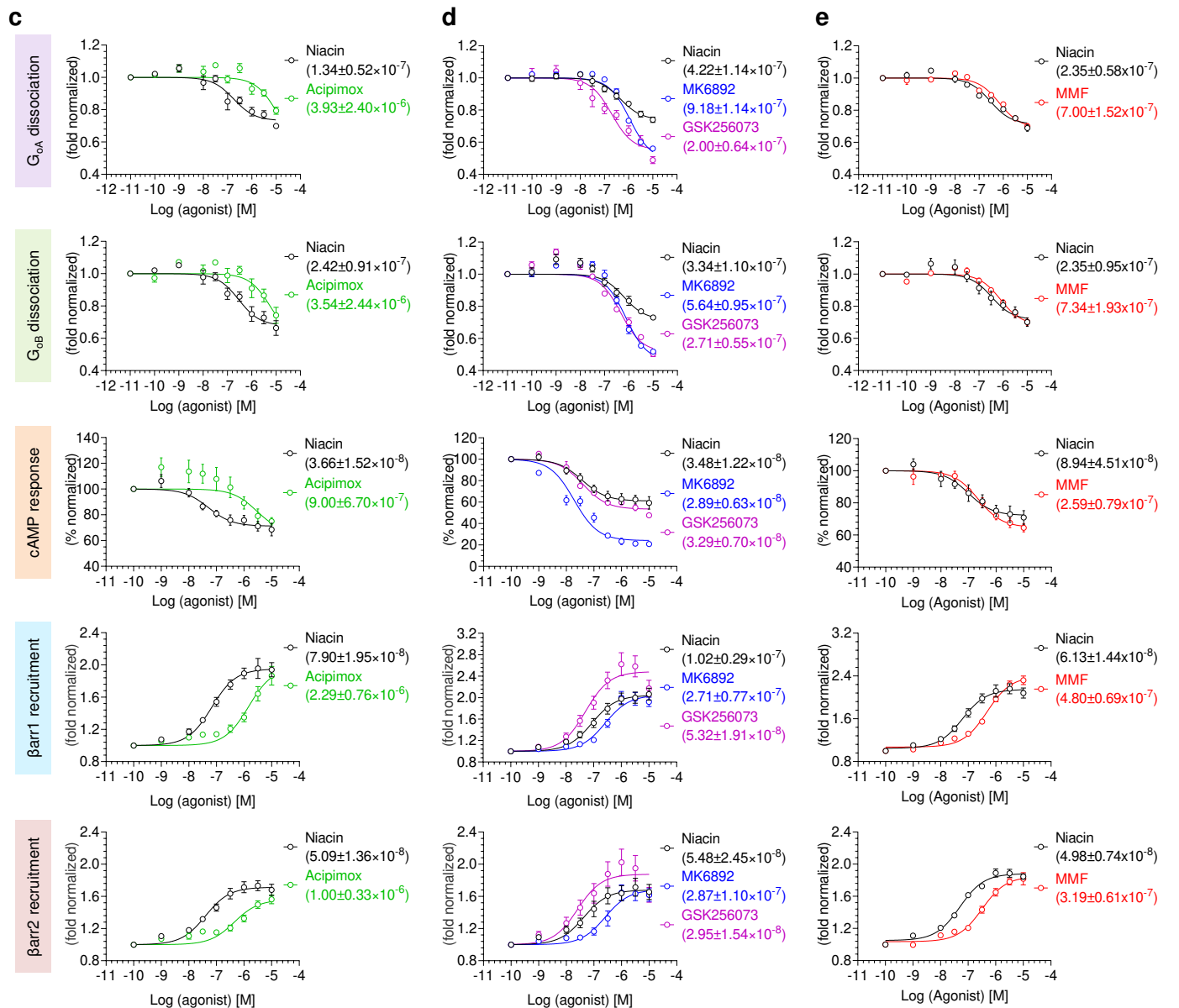
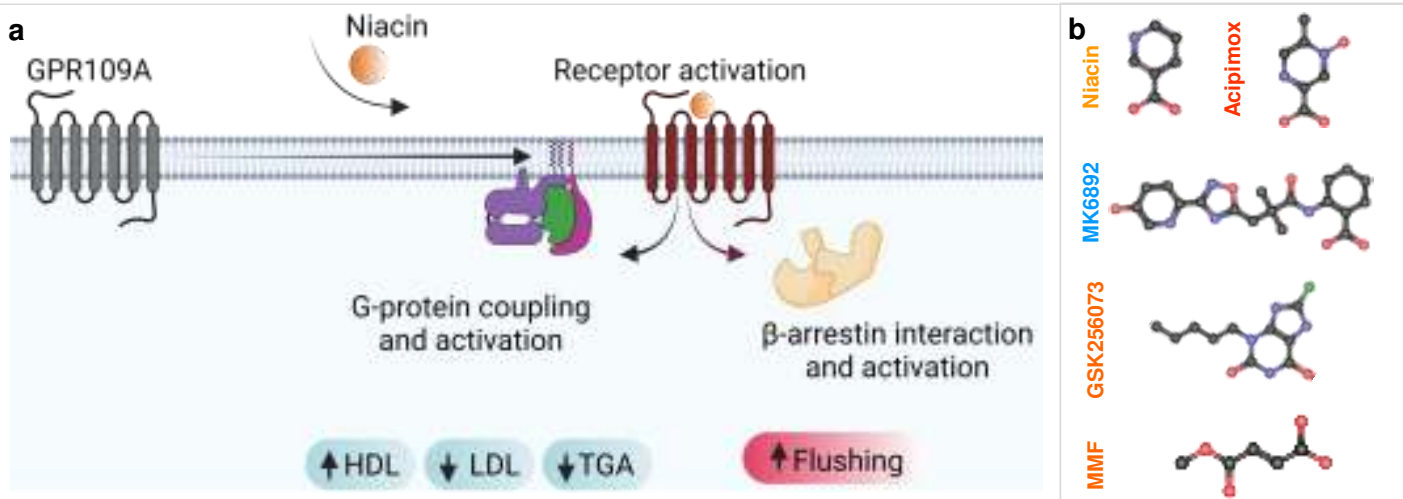
1. Soga, T. et al. Molecular identification of nicotinic acid receptor. *Biochem Biophys Res Commun* **303**, 364–9 (2003).
2. Tunaru, S., Lattig, J., Kero, J., Krause, G. & Offermanns, S. Characterization of determinants of ligand binding to the nicotinic acid receptor GPR109A (HM74A/PUMA-G). *Mol Pharmacol* **68**, 1271–80 (2005).
3. Wise, A. et al. Molecular identification of high and low affinity receptors for nicotinic acid. *J Biol Chem* **278**, 9869–74 (2003).
4. Hanson, J. et al. Nicotinic acid- and monomethyl fumarate-induced flushing involves GPR109A expressed by keratinocytes and COX-2-dependent prostanoid formation in mice. *J Clin Invest* **120**, 2910–9 (2010).
5. Maciejewski-Lenoir, D. et al. Langerhans cells release prostaglandin D-2 in response to nicotinic acid. *Journal of Investigative Dermatology* **126**, 2637–2646 (2006).
6. Offermanns, S. et al. International Union of Basic and Clinical Pharmacology. LXXXII: Nomenclature and Classification of Hydroxy-carboxylic Acid Receptors (GPR81, GPR109A, and GPR109B). *Pharmacol Rev* **63**, 269–90 (2011).
7. Blad, C.C., Ahmed, K., AP, I.J. & Offermanns, S. Biological and pharmacological roles of HCA receptors. *Adv Pharmacol* **62**, 219–50 (2011).
8. Richman, J.G. et al. Nicotinic acid receptor agonists differentially activate downstream effectors. *Journal of Biological Chemistry* **282**, 18028–18036 (2007).



- 641 9. Tang, Y.T. et al. Enhancement of arachidonic acid signaling  
642 pathway by nicotinic acid receptor HM74A. *Biochemical and*  
643 *Biophysical Research Communications* **345**, 29–37 (2006).
- 644 10. Walters, R.W. et al. beta-Arrestin1 mediates nicotinic acid-  
645 induced flushing, but not its antilipolytic effect, in mice.  
646 *J Clin Invest* **119**, 1312–21 (2009).
- 647 11. Reiter, E. & Lefkowitz, R.J. GRKs and beta-arrestins: roles  
648 in receptor silencing, trafficking and signaling. *Trends*  
649 *Endocrinol Metab* **17**, 159–65 (2006).
- 650 12. Tunaru, S. et al. PUMA-G and HM74 are receptors for nicotinic  
651 acid and mediate its anti-lipolytic effect. *Nat Med* **9**, 352–5  
652 (2003).
- 653 13. Pike, N.B. Flushing out the role of GPR109A (HM74A) in the  
654 clinical efficacy of nicotinic acid. *Journal of Clinical*  
655 *Investigation* **115**, 3400–3403 (2005).
- 656 14. Tang, H., Lu, J.Y.L., Zheng, X.M., Yang, Y.H. & Reagan, J.D.  
657 The psoriasis drug monomethylfumarate is a potent nicotinic  
658 acid receptor agonist. *Biochemical and Biophysical Research*  
659 *Communications* **375**, 562–565 (2008).
- 660 15. Berger, A.A. et al. Monomethyl Fumarate (MMF, Bafiertam) for  
661 the Treatment of Relapsing Forms of Multiple Sclerosis (MS).  
662 *Neurol Int* **13**, 207–223 (2021).
- 663 16. Hanson, J., Gille, A. & Offermanns, S. Role of HCA(2)  
664 (GPR109A) in nicotinic acid and fumaric acid ester-induced  
665 effects on the skin. *Pharmacol Ther* **136**, 1–7 (2012).
- 666 17. Benyo, Z. et al. GPR109A (PUMA-G/HM74A) mediates nicotinic  
667 acid-induced flushing. *J Clin Invest* **115**, 3634–40 (2005).
- 668 18. Benyo, Z., Gille, A., Bennett, C.L., Clausen, B.E. &  
669 Offermanns, S. Nicotinic acid-induced flushing is mediated by  
670 activation of epidermal langerhans cells. *Mol Pharmacol* **70**,  
671 1844–9 (2006).
- 672 19. Offermanns, S. Heating up the cutaneous flushing response.  
673 *Arterioscler Thromb Vasc Biol* **34**, 1122–3 (2014).
- 674 20. Dubrall, D. et al. Do dimethyl fumarate and nicotinic acid  
675 elicit common, potentially HCA(2) -mediated adverse  
676 reactions? A combined epidemiological-experimental approach.  
677 *Br J Clin Pharmacol* **87**, 3813–3824 (2021).
- 678 21. Bodor, E.T. & Offermanns, S. Nicotinic acid: an old drug with  
679 a promising future. *Br J Pharmacol* **153 Suppl 1**, S68–75  
680 (2008).
- 681 22. Shen, H.C. et al. Discovery of a Biaryl Cyclohexene  
682 Carboxylic Acid (MK-6892): A Potent and Selective High  
683 Affinity Niacin Receptor Full Agonist with Reduced Flushing  
684 Profiles in Animals as a Preclinical Candidate. *Journal of*  
685 *Medicinal Chemistry* **53**, 2666–2670 (2010).
- 686 23. Sprecher, D. et al. Discovery and characterization of  
687 GSK256073, a non-flushing hydroxy-carboxylic acid receptor 2  
688 (HCA2) agonist. *European Journal of Pharmacology* **756**, 1–7  
689 (2015).
- 690 24. van Veldhoven, J.P.D. et al. Structure-activity relationships  
691 of trans-substituted-propenoic acid derivatives on the  
692 nicotinic acid receptor HCA2 (GPR109A). *Bioorganic &*  
693 *Medicinal Chemistry Letters* **21**, 2736–2739 (2011).

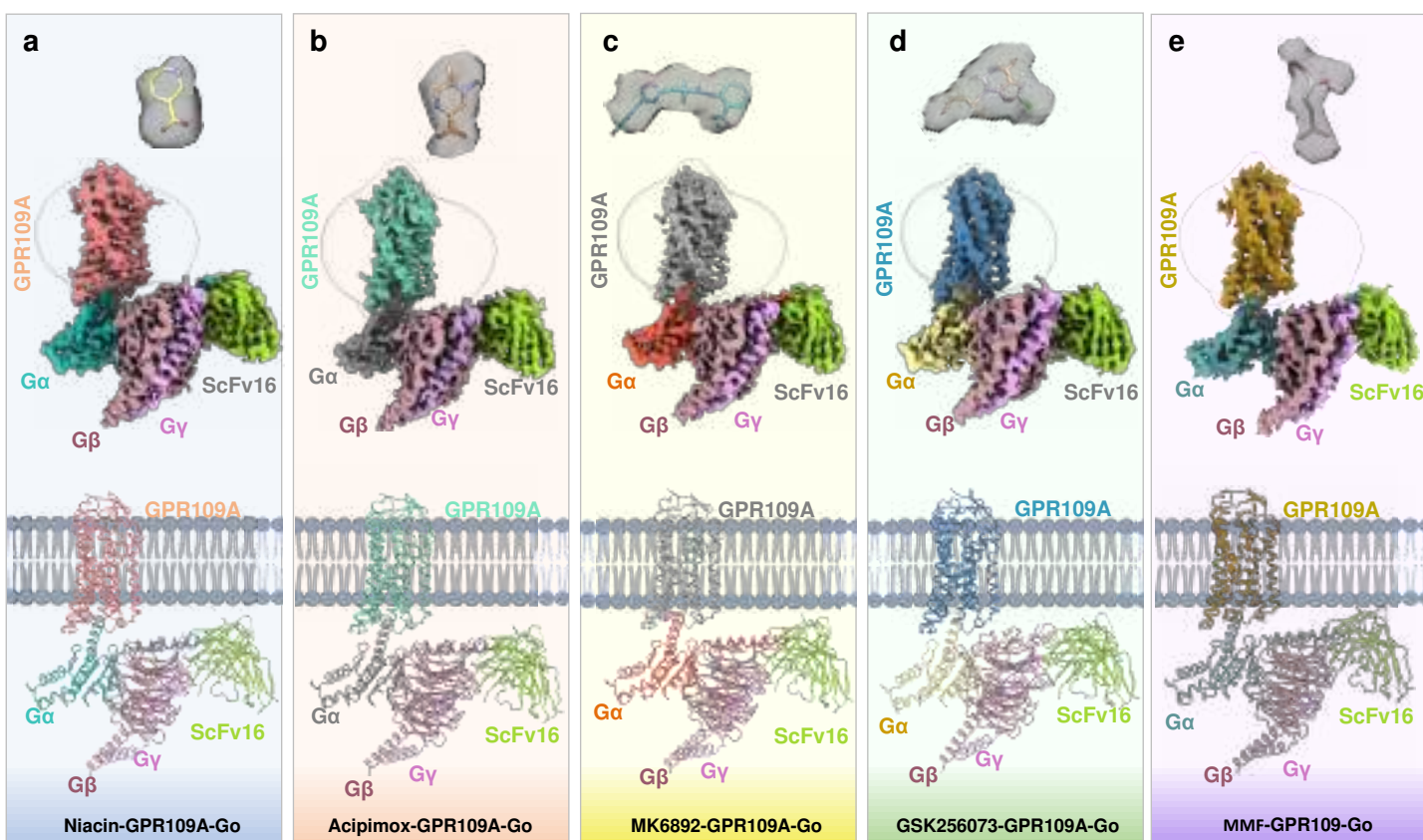
- 694 25. Yang, Y. et al. Structural insights into the human niacin  
695 receptor HCA2-G(i) signalling complex. *Nat Commun* **14**, 1692  
696 (2023).
- 697 26. Xin Pan, F.Y., Peiruo Ning, Zhiyi Zhang, Binghao Zhang, Geng  
698 Chen, Wei Gao, Chen Qiu, Zhangsong Wu, Kaizheng Gong,  
699 Jiancheng Li, Jiang Xia, Yang Du. Structural insights into  
700 ligand recognition and selectivity of the human  
701 hydroxycarboxylic acid receptor HCAR2. *bioRxiv*  
702 2023.03.28.534513 (2023).
- 703 27. Ma, N., Nivedha, A.K. & Vaidehi, N. Allosteric communication  
704 regulates ligand-specific GPCR activity. *FEBS J* **288**, 2502-  
705 2512 (2021).
- 706 28. Nivedha, A.K. et al. Identifying Functional Hotspot Residues  
707 for Biased Ligand Design in G-Protein-Coupled Receptors. *Mol*  
708 *Pharmacol* **93**, 288-296 (2018).
- 709 29. Kawakami, K. et al. Heterotrimeric Gq proteins act as a  
710 switch for GRK5/6 selectivity underlying beta-arrestin  
711 transducer bias. *Nat Commun* **13**, 487 (2022).
- 712 30. Carpenter, B. & Tate, C.G. Expression, Purification and  
713 Crystallisation of the Adenosine A(2A) Receptor Bound to an  
714 Engineered Mini G Protein. *Bio Protoc* **7**(2017).
- 715 31. Nehme, R. et al. Mini-G proteins: Novel tools for studying  
716 GPCRs in their active conformation. *PLoS One* **12**, e0175642  
717 (2017).
- 718 32. Hong, C. et al. Structures of active-state orexin receptor 2  
719 rationalize peptide and small-molecule agonist recognition  
720 and receptor activation. *Nat Commun* **12**, 815 (2021).
- 721 33. Pandey, S. et al. Intrinsic bias at non-canonical, beta-  
722 arrestin-coupled seven transmembrane receptors. *Mol Cell* **81**,  
723 4605-4621 e11 (2021).
- 724 34. Ghosh, E. et al. Conformational Sensors and Domain Swapping  
725 Reveal Structural and Functional Differences between beta-  
726 Arrestin Isoforms. *Cell Rep* **28**, 3287-3299 e6 (2019).
- 727 35. Zivanov, J. et al. New tools for automated high-resolution  
728 cryo-EM structure determination in RELION-3. *Elife* **7**(2018).
- 729 36. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A.  
730 cryoSPARC: algorithms for rapid unsupervised cryo-EM  
731 structure determination. *Nat Methods* **14**, 290-296 (2017).
- 732 37. Liebschner, D. et al. Macromolecular structure determination  
733 using X-rays, neutrons and electrons: recent developments in  
734 Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
- 735 38. Adams, P.D. et al. PHENIX: a comprehensive Python-based  
736 system for macromolecular structure solution. *Acta*  
737 *Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
- 738 39. Pettersen, E.F. et al. UCSF ChimeraX: Structure visualization  
739 for researchers, educators, and developers. *Protein Sci* **30**,  
740 70-82 (2021).
- 741 40. Pettersen, E.F. et al. UCSF Chimera--a visualization system  
742 for exploratory research and analysis. *J Comput Chem* **25**,  
743 1605-12 (2004).
- 744 41. Emsley, P. & Cowtan, K. Coot: model-building tools for  
745 molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**,  
746 2126-32 (2004).

747 42. Chen, V.B. et al. MolProbity: all-atom structure validation  
748 for macromolecular crystallography. *Acta Crystallogr D Biol*  
749 *Crystallogr* **66**, 12-21 (2010).  
750 43. Baidya, M. et al. Allosteric modulation of GPCR-induced beta-  
751 arrestin trafficking and signaling by a synthetic intrabody.  
752 *Nat Commun* **13**, 4634 (2022).  
753 44. Pandey, S., Roy, D. & Shukla, A.K. Measuring surface  
754 expression and endocytosis of GPCRs using whole-cell ELISA.  
755 *Methods Cell Biol* **149**, 131-140 (2019).  
756 45. Maharana, J. et al. Structural snapshots uncover a key  
757 phosphorylation motif in GPCRs driving beta-arrestin  
758 activation. *Mol Cell* (2023).  
759

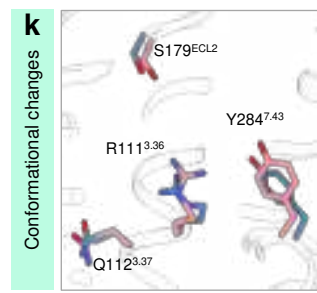
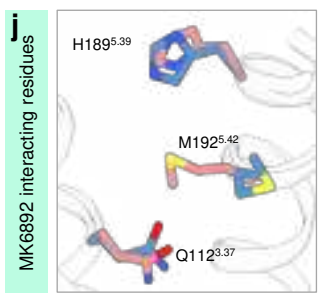
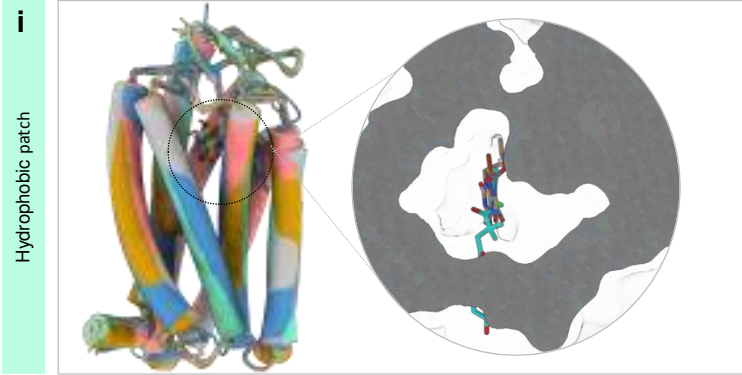
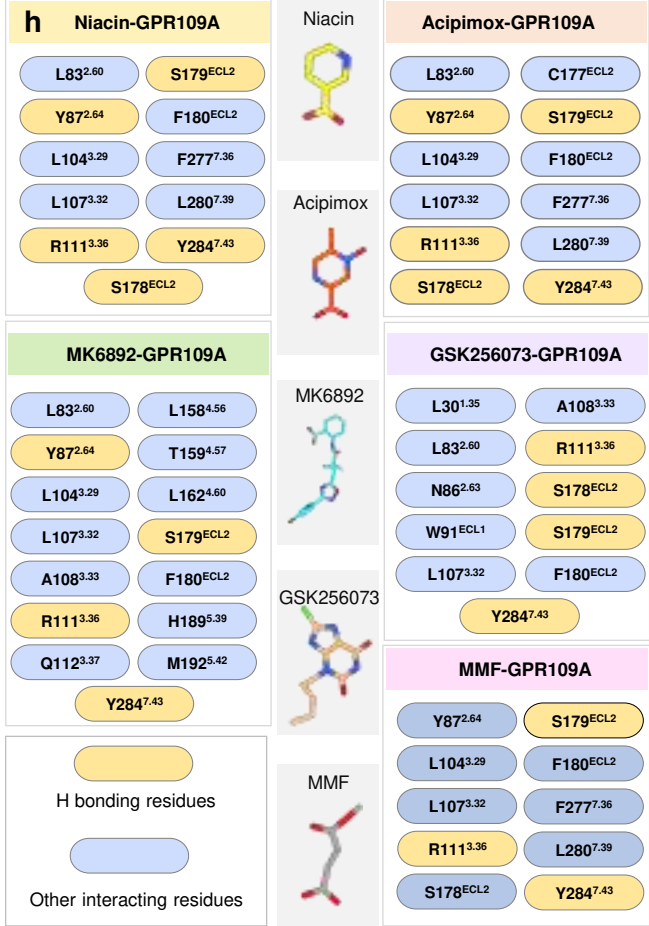
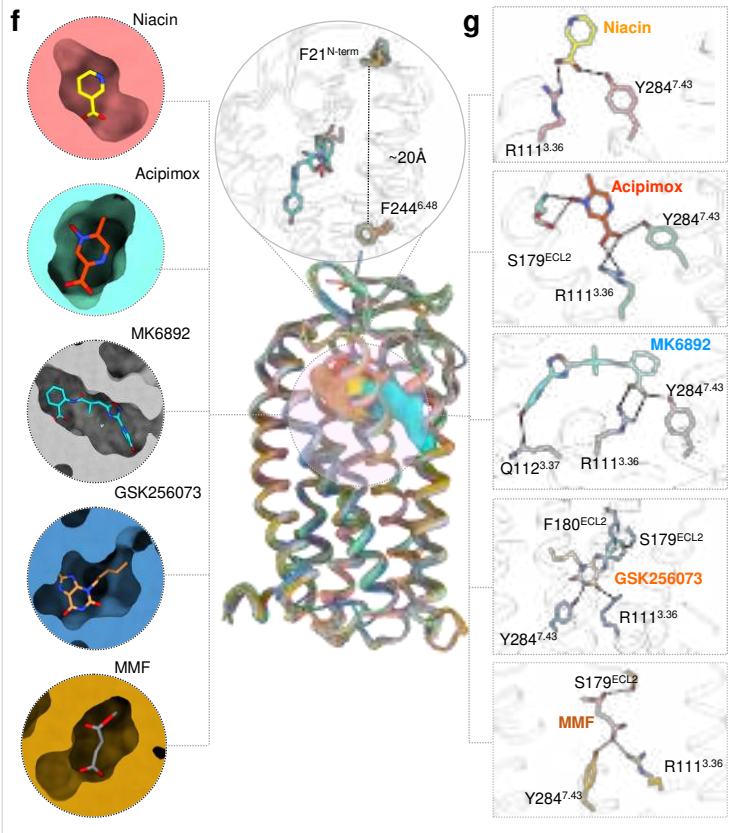
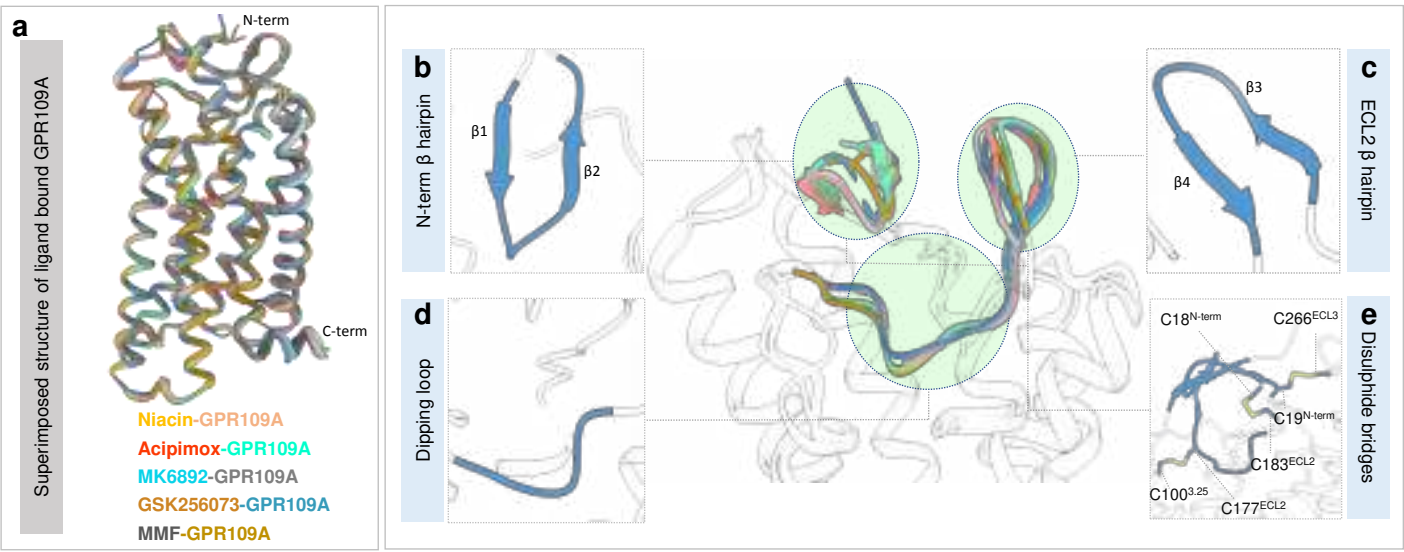


**Figure 1: Pharmacological profiling of niacin, acipimox, MK6892, GSK256073 and MMF on GPR109A.** **a**, Chemical structure of four GPR109A ligands used in the current study. **b**, Diagrammatic illustration of GPR109A activation and downstream signalling outcomes. **c**, G-protein activation and  $\beta$ arr recruitment downstream of GPR109A in response to acipimox with niacin as a reference ligand. First and second panel showing  $G_{oA}$  and  $G_{oB}$  dissociation studied by nanoBiT-based G-protein dissociation assay (Receptor+LgBiT- $G\alpha_{oA}/G\alpha_{oB}+G\beta+SmBiT-G\gamma$ ) (mean $\pm$ SEM ; n=4 ; fold normalized with the minimum concentration for each ligand as 1). Forskolin induced cAMP decrease measured by GloSensor assay is shown in the third panel (mean $\pm$ SEM ; n=4 ; % normalized with the minimum concentration for each ligand as 100).  $\beta$ arr recruitment was studied by nanoBiT-based assay (Receptor-SmBiT+LgBiT- $\beta$ arr) and is shown in fourth and fifth panel (mean $\pm$ SEM ; n=4 and n=5 for fourth and fifth panel respectively ; fold normalized with the minimum concentration for each ligand as 1). **d**, G-protein activation and  $\beta$ arr recruitment downstream of GPR109A in response to MK6892 and GSK256073 with niacin as a reference ligand. First and second panel showing  $G_{oA}$  and  $G_{oB}$  dissociation (mean $\pm$ SEM ; n=4 ; fold normalized with the minimum concentration for each ligand as 1). Forskolin induced cAMP response is shown in the third panel (mean $\pm$ SEM ; n=4 ; % normalized with the minimum concentration for each ligand as 100). Fourth and fifth panel showing  $\beta$ arr1 and 2 recruitment respectively (mean $\pm$ SEM ; n=6 ; fold normalized with the minimum concentration for each ligand as 1). **e**, G-protein activation and  $\beta$ arr recruitment downstream of GPR109A in response to monomethyl fumarate (MMF) with niacin as a reference ligand. First and second panel showing  $G_{oA}$  and  $G_{oB}$  dissociation (mean $\pm$ SEM ; n=3 ; fold normalized with the minimum concentration for each ligand as 1). Forskolin induced cAMP response is shown in the third panel (mean $\pm$ SEM ; n=4 ; % normalized with the minimum concentration for each ligand as 100). Fourth and fifth panel showing  $\beta$ arr1 and 2 recruitment respectively (mean $\pm$ SEM ; n=4 ; fold normalized with the minimum concentration for each ligand as 1). **f**, Bias factor was calculated using the software <https://biasedcalculator.shinyapps.io/calc/>. During bias factor calculation Niacin stimulated response was considered as reference and observed G-protein biased with MK6892.





**Figure 2: Overall architecture of Niacin, Acipimox, MK6892, GSK256073 and MMF bound GPR109A-G protein complexe.** Map and ribbon diagram of the ligand-bound GPR109A-Go complexes (front view) and the cryo-EM densities of the ligands (sticks) are depicted as transparent surface representations. **a, niacin-GPR109A-Go:** Light coral: GPR109A, light sea green: miniGao, rosy brown: Gβ1, orchid: Gγ2, yellow green: ScFv16, **b, acipimox-GPR109A-Go:** medium aquamarine: GPR109A, gray: miniGao, rosy brown: Gβ1, orchid: Gγ2, yellow green: ScFv16, **c, MK6892-GPR109A-Go:** Dark gray: GPR109A, tomato: miniGao, rosy brown: Gβ1, orchid: Gγ2, yellow green: ScFv16, **d, GSK256073-GPR109A-Go:** Steel blue: GPR109A, khaki: miniGao, rosy brown: Gβ1, orchid: Gγ2, yellow green: ScFv16, **e, MMF-GPR109A-Go:** Dark golden rod: GPR109A, cadet blue: miniGao, rosy brown: Gβ1, orchid: Gγ2, yellow green: ScFv16.

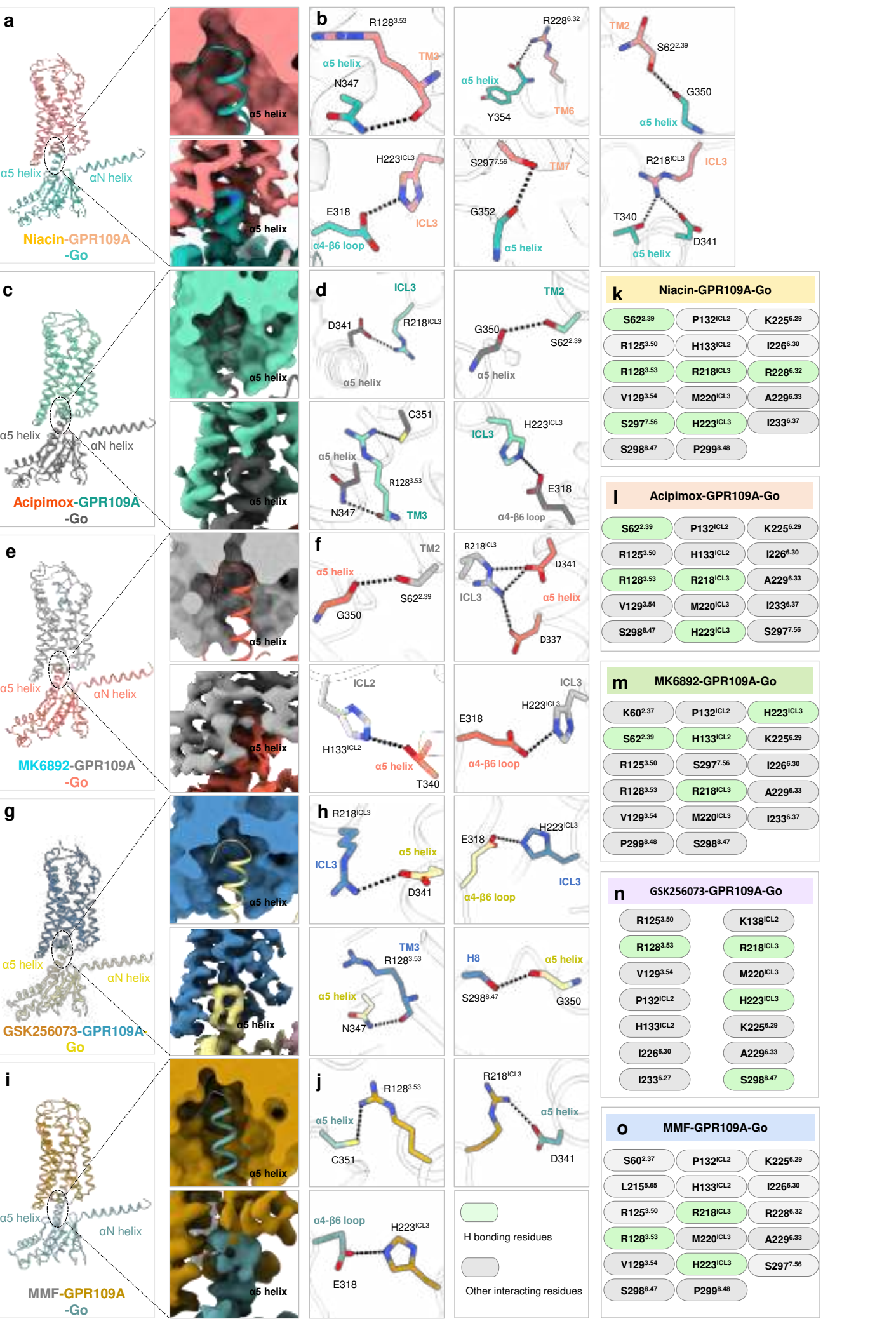


**Figure 3:Ligand binding pocket of GPR109A.**

**a**, Structure superposition of niacin, acipimox, MK6892, GSK256073 and MMF bound GPR109A. **b**, Structural features of Ligand(GSK256073) bound-GPR109A, N terminal  $\beta$ -hairpin (left upper panel), **c**, Close-up view of ECL2 dipping into the orthosteric pocket (left lower panel), **d**, ECL2  $\beta$ -hairpin (upper right panel) and **e**, Ribbon diagram of disulfide bridges. **f**, Superposed niacin, acipimox, MK6892, GSK256073 and MMF bound GPR109A structures highlighting the orthosteric binding pocket (Left panel, cross-sections of GPR109A bound to the individual ligand. **g**, GPR109A ligand binding pocket highlighting the major interactions of the individual ligand. **h**, List of GPR109A residues interaction with ligands. **i**, Cross-section of GPR109A orthosteric sites depicting the hydrophobic patch surrounding the individual ligand. **j**, Interacting residues of GPR109A with the extended part of MK6892. **k**, Conformational changes of GPR109A residues interacting with MK6892 with respect to niacin.



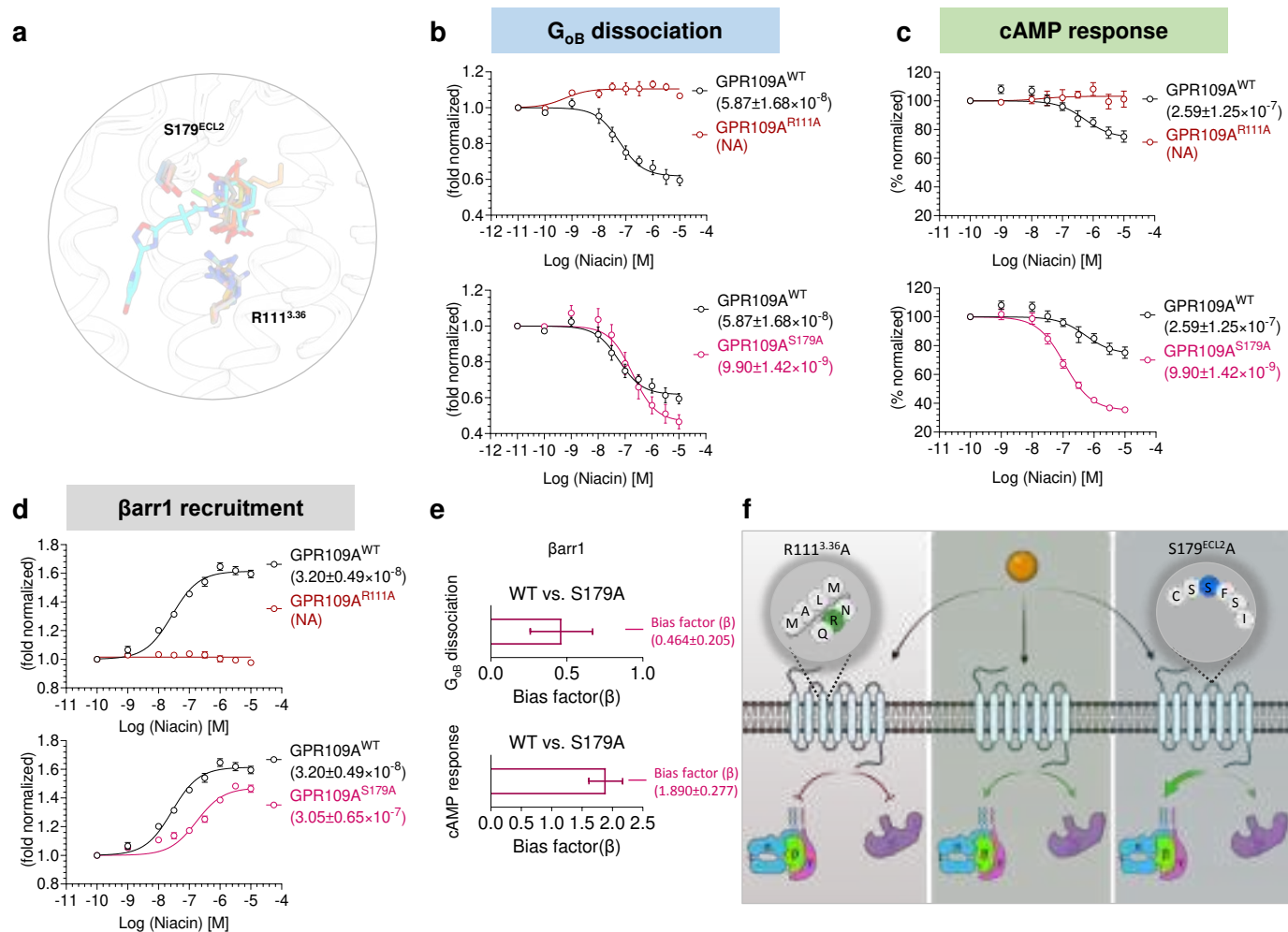






**Figure 5: GPR109A-G-protein interacting interface.**

**a, c, e, g, i,** Representation of  $\alpha 5$  helix of Gao docking into the cytoplasmic core of GPR109A bound to niacin, acipimox, MK6892, GSK256073 and monomethyl fumarate (MMF) respectively. **b, d, f, h, j,** Key interactions between Gao residues and residues of the cytoplasmic core of GPR109A. Black dotted line represents the H-bond. **k-o,** Illustration of residues contact between GPR109A and Gao in niacin, acipimox, MK6892, GSK256073, and monomethyl fumarate (MMF) bound structures.



**Figure 6. Structure guided bias signaling.**

**a**, Cartoon representation of residues interacting via H-bond with niacin (yellow), acipimox (orange red), MK6892 (cyan), GSK256073 (sandy brown), and monomethyl fumarate (MMF) (grey). **b,c**, G-protein activation downstream of GPR109A<sup>WT</sup>, GPR109A<sup>R111A</sup>, and GPR109A<sup>S179A</sup> in response to niacin was studied by nanoBiT-based G-protein dissociation assay (Receptor+LgBiT-Gα<sub>oB</sub>+Gβ+SmBiT-Gγ) (panel b) (mean±SEM ; n=3 ; fold normalized with the minimum concentration for each ligand as 1) and forskolin-induced cAMP level decay by GloSensor assay (panel c) (mean±SEM ; n=3 ; % normalized with the minimum concentration for each ligand as 100). **d**, βarr recruitment downstream of GPR109A<sup>WT</sup>, GPR109A<sup>R111A</sup>, and GPR109A<sup>S179A</sup> in response to niacin was studied by nanoBiT-based assay (Receptor-SmBiT+LgBiT-βarr) (mean±SEM; n=4 ; fold normalized with the minimum concentration for each ligand as 1). **e**, Bias factor for the mutant GPR109A<sup>S179A</sup> was calculated using the software <https://biasedcalculator.shinyapps.io/calc/>. During bias factor calculation GPR109A<sup>WT</sup> was considered as reference and observed G-protein biased with GPR109A<sup>S179A</sup> upon stimulation with niacin. **f**, Schematic depicting the effect of the two mutants GPR109A<sup>R111A</sup> and GPR109A<sup>S179A</sup> on G-protein activation and βarr recruitment in response to niacin.