Computationally designed GPCR quaternary structures bias signaling pathway activation

Paradis, J. S.[¶,1] , Feng, X. [¶,2,$], Murat, B.[#1], Jefferson R.[#2], Szpakowska, M.[3], Hogue, M.[1], Bergkamp, N. D.[4], Heydenreich, F.M.[1], Smit, M. J.[4], Chevigne, A.[3], Bouvier, M.[§,1], Barth, P.[§,2]

Affiliations

1: Department of Biochemistry and Molecular Medicine, Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, Montréal, Quebec, Canada H3T 1J4

2: Interfaculty Institute of Bioengineering, Ecole Polytechnique Federale de Lausanne, Lausanne, CH-1015, Switzerland

3: Department of Infection and Immunity, Immuno-Pharmacology and Interactomics, Luxembourg Institute of Health, Esch-sur-Alzette, Luxembourg

4: Amsterdam Institute for Molecules, Medicines and Systems (AIMMS), Division of Medicinal Chemistry, Faculty of Sciences, Vrije Universiteit, Amsterdam, The Netherlands

$: Current address: Structural Biology Program, Van Andel Institute, Grand Rapids, MI, USA

¶, # Contributed equally

§ Correspondences should be addressed to:

Patrick Barth, Interfaculty Institute of Bioengineering, Ecole Polytechnique Federale de Lausanne, Lausanne, CH-1015, Switzerland, patrick.barth@epfl.ch

Michel Bouvier, Institute for Research in Immunology and Cancer, Université de Montréal

2900, Boulevard Édouard-Montpetit; email: michel.bouvier@umontreal.ca

Abstract

Communication across membranes controls critical cellular processes and is achieved by receptors translating extracellular signals into selective cytoplasmic responses. While receptor tertiary structures can now be readily characterized, receptor associations into quaternary structures are very challenging to study and their implications in signal transduction remain poorly understood. Here, we report a computational approach for predicting membrane receptor self-associations, and designing receptor oligomers with various quaternary structures and signaling properties. Using this approach, we designed chemokine receptor CXCR4 dimers with reprogrammed stabilities, conformations, and abilities to activate distinct intracellular signaling proteins. In agreement with our predictions, the designed CXCR4s dimerized through distinct conformations and displayed different quaternary structural changes upon activation. Consistent with the active state models, all engineered CXCR4 oligomers activated the G protein Gi, but only a few specific dimer structures also recruited β-arrestins. Overall, we demonstrate that quaternary structures represent an important unforeseen mechanism of receptor biased signaling and reveal the existence of a conformational switch at the dimer interface of several G protein-coupled receptors including CXCR4, mu-Opioid and type-2 Vasopressin receptors that selectively control the activation of G proteins vs β-arrestin-mediated pathways. The approach should prove useful for predicting and designing receptor associations to uncover and reprogram selective cellular signaling functions.

Introduction

A wide range of membrane proteins, including single-pass receptor tyrosine kinases, cytokine receptors and ion channels, function through the folding and association of several polypeptide chains into specific quaternary structures. The functional role of oligomerization in other membrane protein classes remains controversial as the observation of receptor associations is very sensitive to the experimental conditions and techniques[1-4]. Receptors from the largest class of G protein-coupled receptors (GPCRs) were often observed as oligomers in electron microscopy, X-ray crystallography and BRET studies[5-11]. However, when trapped as monomers in nanolipid disks, GPCRs, such as rhodopsin and β2 adrenergic receptor, remained functional, binding and activating their primary intracellular signaling G proteins[12,13]. Structural studies suggested that different GPCRs can self-associate through distinct transmembrane helical (TMH) interfaces. Computational modeling approaches based on molecular dynamics simulations have also identified different possible modes and lifetimes of GPCR associations[14,15] but the functional relevance of these oligomeric forms remain poorly understood[5-7,16-22]. For example, chemokine receptor CXCR4 signaling is linked to the formation of nanoclusters at the cell membrane[23]. Such nanoclusters are controlled by key structural motifs at the receptor TMH surface but do not involve the receptor dimeric interface observed in X-ray structures[10].

In principle, computational design techniques can probe and decipher the importance of protein associations by reprogramming protein-protein interactions or designing competitive binding inhibitors, but these approaches have mostly been applied to soluble proteins[24-26]. Applications to membrane proteins have been limited to the design of single-pass TMH associations[27-29].

Here, we developed a computational approach for modeling and designing quaternary structures of multi-pass membrane receptors. Using the method, we engineered the chemokine receptor CXCR4 to associate into distinct oligomeric structures that recruited and activated

intracellular signaling proteins differently. Altogether, our study reveals that quaternary structures constitute important unforeseen structural determinants of GPCR biased signaling and identified a common conformational switch at the dimer interface of several GPCRs that control G protein versus β-arrestin signaling. The approach is general and should prove useful for reprogramming cellular functions through designed receptor associations.

Results

Computational approach for modeling and designing multi-pass receptor oligomers

We developed an approach to model and design multi-pass membrane protein associations with precise quaternary structures, stabilities and signaling functions (**Fig.1a,b, Supplementary Fig.S1**). The method builds GPCR monomeric structures in distinct active and inactive states, docks them to identify possible modes of protomer associations in homodimers and design the binding interfaces to generate quaternary structures with distinct dimer stabilities, conformations and propensity to recruit and activate specific intracellular signaling proteins. In this study, an active state model refers to a GPCR in an active state conformation modelled by using a structure bound to G-protein or β-arrestin as template.

We applied the approach to reprogram the homo-dimeric structure and function of CXCR4, a GPCR from the chemokine receptor family. We chose CXCR4 because it is a critical signaling hub involved in immune responses[7,30] and HIV infection, as well as a receptor for which multiple experimental lines of evidence supporting the formation of constitutive homo-oligomers and its regulation by ligands exists[10,11].

We designed CXCR4 receptor oligomers with different binding affinities and quaternary structures to elucidate the role of self-association in distinct CXCR4 intracellular signaling functions. Specifically (**Fig.1a,b, Supplementary Fig.S1**), we modeled CXCR4 WT monomers in inactive and

active signaling states. For instance, the active state model of CXCR4 was obtained from the active state structure of the homologous viral GPCR US28 (PDB: 4XT1) using the method IPHoLD which integrates homology modeling and ligand docking [31]. The CXCR4 WT monomers in the inactive state were taken from the antagonist-bound CXCR4 WT structure (PDB: 3ODU) after energy minimization of the X-ray coordinates. The CXCR4 WT monomers were assembled into inactive or active state dimers along different dimer binding interfaces involving TMHs 4, 5 and 6. We found that, in both the inactive and active states, the dimer WT models populated primarily an open-dimer conformation similar to that observed in the antagonist-bound receptor X-ray structure but also, to a lesser extent, a distinct closed-dimer conformation (**Fig.1c-f, Supplementary Table S1, Supplementary Fig.S2, Extended Data Set 1**). The distribution between dimer conformations can be deduced from the difference in binding energy (strength of association) at the distinct dimer interfaces (**Supplementary Table S1**). Interestingly, while the major open dimer conformation remains very similar in both signaling states, the minor closed form differs by a slight rotation around TMH5 between the inactive and active state conformations of the receptor (**Supplementary Fig.S2**).

To elucidate the function of these different quaternary structures, we then designed TMH and loop binding surfaces to selectively stabilize either the open-dimer or the closed-dimer conformation of the inactive and active state dimer models. The method first searches for combinations of mutations and conformations that modulate the intermolecular interactions between the monomers without affecting the monomer's intrinsic conformational stability and functions. Any design that modifies the dimer binding energies as intended but significantly affect monomer stability is systematically discarded (see **Methods**). After each round of design, the designed CXCR4 monomers are assembled into dimers to predict the effects of the designed sequence-structure features on the distribution of quaternary structures in distinct signaling states. Lastly, the G protein Gi and β-arrestin are docked and assembled onto the designed CXCR4 active state

dimers to predict whether the engineered receptors would effectively recruit and activate these intracellular signaling proteins. The cycles of design, flexible docking and ternary complex assembly are repeated until the calculations converge to significant predicted reprogramming of the quaternary structure and functional selectivity of the designed CXCR4 oligomers (**Fig.1a,b, Supplementary Fig.S1**).

Designing CXCR4 dimers with selective conformations and intracellular functions

From our *in silico* design screen, we first selected three engineered CXCR4s predicted to dimerize with greater propensity than CXCR4 WT in the open-dimer conformation (**Fig.1c,e,f, Supplementary Table S1**). The designs involved key conformational lock motifs stabilizing the open-dimer conformation (**Fig.1c,e,f**). The L194$^{5.33}$K and the L194$^{5.33}$R design introduced a set of strong and conformationally selective polar contacts between the extracellular sides of TMH5s of two protomers (**Fig.1c**). The triplet design, formed by the V198$^{5.37}$F-V197$^{5.36}$M mutation on one protomer and the V198$^{5.37}$W on the other protomer, encoded a new network of optimal hydrophobic contacts bridging the membrane-embedded core of the dimer-binding interface between TMH5s (**Fig.1c**). Similarly to WT, when modeled in the active state, these designs primarily dimerized in an open conformation that could readily form tight active state complex structures with both Gi and β-arrestin (**Fig.1f, Supplementary Table S1, Supplementary Fig.S3**).

Conversely, we also engineered two binding surfaces predicted to instead stabilize the closed-dimer conformation (**Fig.1d,e,f, Supplementary Table S1**). We selected these "closed-dimer-stabilizing" designs, because, unlike WT, they preferentially assemble into closed-dimer conformations that form tight active-state complex structures with Gi but not with β-arrestin (**Fig.1b, f, Supplementary Fig.S3**). We found that steric hindrance prevents the optimal interaction of β-arrestin's finger loop in the intracellular binding groove of CXCR4 when the receptor occupies the

closed-dimer conformation (**Supplementary Fig.S4**). Both designed interfaces (that we name W195$^{5.34}$L and N192^{ECL2}W design switches) involved several conformational switch motifs stabilizing the closed-dimer conformation, especially when the receptor occupies the signaling active state (**Fig.1f, Supplementary Table S1**). The W195$^{5.34}$L design switch increased the packing of TMH4 and 5 across the extracellular side of the binding interface, stabilizing the closed dimer conformation through additional van der Waals contacts (**Fig.1d**). The N192^{ECL2}W design switch induced several conformational changes in a neighboring layer of residues buried at the dimer interface, creating new key hydrophobic interactions stabilizing the closed form (**Fig.1d**).

By simulating the association for the WT and the designed CXCR4 monomers, we identified important changes in the stability and hence distribution of the dimer conformations between the inactive and active states of the receptor (**Fig.1e, f, Supplementary Table S1**). Concerning the WT receptor, we observed that the closed dimer conformation was significantly more stable in the active state, indicating a relative shift toward the closed form in that state. By contrast, virtually no difference in the closed dimer conformation stability between the inactive and active states was observed for the "open-dimer-stabilizing" designs (L194$^{5.33}$R switch). The largest changes in dimer populations between inactive and active signaling states were observed for the "closed-dimer-stabilizing" designs (W195$^{5.34}$L switch). Despite a significant stabilization of the closed dimer conformation, the open dimer remained the most stable form in the inactive state and the W195$^{5.34}$L variant still predominantly populated the open-dimer structure in that state. However, the distribution between open and closed conformation of the W195$^{5.34}$L variant was reversed in the active state and the closed form became the most stable and dominant structure. Overall, the W195$^{5.34}$L design was found to be most stabilized in the active state closed-dimer form (**Supplementary Table S1**). In addition to the dimer conformation distribution, our calculations provided insights into the dimerization propensity of the different CXCR4 variants (**Supplementary Table S2**). In the inactive state, the "closed-dimer-stabilizing" N192^{ECL2}W and W195$^{5.34}$L designs formed weaker dimers while

the "open-dimer-stabilizing" L194$^{5.33}$R design formed stronger dimers than WT, suggesting that the "closed-dimer-stabilizing" designs would occupy more often the monomeric state. The reverse scenario was observed in the active state. While the dimerization propensity of the L194$^{5.33}$R design was lower than WT, the W195$^{5.34}$L design formed the most stable active state dimers among all variants.

Designed CXCR4 receptors dimerize in distinct conformations

We validated the predicted designed oligomeric CXCR4 structures and functions using an ensemble of cell-based experiments.

We first measured constitutive and CXCL12 agonist-promoted CXCR4 dimerization in living HEK293T cells by BRET using CXCR4-RLuc and CXCR4-YFP constructs (**Fig.2a**). A large constitutive BRET signal was observed for the WT receptor which, as previously reported[10,11], further increased upon activation by agonist (**Fig.2a, Supplementary Fig.S7**) (**Supplementary Fig.S6** for CXCR4 cell surface expression levels). This increase in BRET can be interpreted as a change in conformation within dimers and a shift toward the closed dimer form, as suggested by our calculations, or as an increase in dimer population upon activation. Consistent with the "open-dimer-stabilizing" designs associating in a similar open conformation than WT, the constitutive BRET signals measured for these designs (L194$^{5.33}$K/L194$^{5.33}$K, L194$^{5.33}$R/L194$^{5.33}$R, V198$^{5.37}$F/V197$^{5.36}$M-V198$^{5.37}$W), were either similar to or slightly larger than WT. However, unlike what is seen for the WT receptor, we did not observe any significant BRET increase upon agonist stimulation. These results suggest that stabilization of the open-dimer conformation prevents further agonist-induced conformational changes across the binding interface and locks the receptor dimer in a constitutive open-dimer conformation, consistent with the lack of stabilization of the closed dimer form upon receptor activation in our simulations. In the specific case of the L194$^{5.33}$K design, the lack of BRET increase upon stimulation could also result in part from the designed receptors

196 occupying more frequently the dimer state than the WT receptor, even without stimulus, as
197 suggested by the significantly increased constitutive BRET signals measured for that design.
198
199 The BRET signals measured for the "closed-dimer-stabilizing" designs
200 (N192^{ECL2}W/N192^{ECL2}W, W195$^{5.34}$L/W195$^{5.34}$L) without ligand stimulus were significantly lower than
201 WT. These observations are consistent with the designs still predominantly occupying the open
202 conformation in the inactive state (**Supplementary Table S1**) and forming overall weaker dimers
203 than WT (**Supplementary Table S2**) that may result in a greater proportion of receptor in the
204 monomeric state. Upon agonist stimulation, however, we observed a larger increase in net BRET
205 signal compared to WT, especially for W195$^{5.34}$L in agreement with the large predicted changes in
206 dimer conformation and dimerization propensity upon receptor activation (**Supplementary Tables**
207 **S1 and S2**). These results suggest that the "closed-dimer-stabilizing" receptors constitutively
208 dimerize less stably than WT and display stronger propensity to associate in the closed form upon
209 agonist stimulation.
210
211 Overall, we observed a strong correlation across receptor variants between the predicted
212 change in closed-dimer stability and the increase in BRET upon receptor activation
213 (**Supplementary Fig.S5**). These results suggest that major conformational changes and population
214 shifts towards the closed dimer form can readily occur in the active state when triggered by strong
215 switching mutations such as W195$^{5.34}$L.
216
217 In summary, the BRET measurements validate the designed CXCR4 dimer structures and indicate
218 that receptor dimers with distinct strengths of associations and quaternary conformations can be
219 rationally engineered using our approach.
220
221 **Designed CXCR4 receptors activate distinct intracellular signaling proteins**

222 According to our calculations, the two classes of designed receptors should display distinct
223 propensity to bind and activate intracellular signaling proteins. While the receptors dimerizing in the
224 open conformation should recruit both Gi and β-arrestin, the receptors preferentially dimerizing in a
225 closed conformation should couple strongly to Gi only.
226
227 To validate these predictions, we measured Gi activation and β-arrestin recruitment to
228 CXCR4 using BRET-based assays in HEK293 cells. Consistent with the active state modeling, both
229 classes of designed CXCR4 dimers were able to activate Gi similarly to WT, as measured by the
230 agonist-induced dissociation of the heterotrimeric Gi protein (**Fig.2b**) and the inhibition of cAMP
231 production (**Fig.2c**). As shown in **Supplementary Fig.S6d**, HEK293 cells endogenously express a
232 low level of CXCR4 that result in a background CXCL12-promoted cAMP inhibition that can easily
233 be distinguished from the signal generated by the transfected WT or mutant receptors (**Fig.2c**). No
234 such background signal could be observed in the BRET-based Gi activation or β-arrestin
235 recruitment assays due to their lower level of amplification. Both assays clearly indicated that the
236 mutations did not affect the ability of the receptor to activate Gi.
237
238 β-arrestin recruitment was measured using BRET reporting directly the interaction between
239 CXCR4-RLuc and β-arrestin2-YFP (in HEK293 cells)[32] or ebBRET[33] monitoring the interaction
240 between β-arrestin2-RLuc and the lipid-modified rGFP-CAAX at the cell membrane. Both assays
241 consistently showed that the "open-dimer-stabilizing" designs recruited β-arrestin very effectively
242 and similarly to WT upon agonist stimulus (**Fig.3a, Supplementary Fig.S8**). On the contrary, and in
243 agreement with our predictions, the "closed-dimer-stabilizing" designs had largely impaired β-
244 arrestin recruitment abilities. Specifically, while β-arrestin2 coupling to the $N192^{ECL2}W$ design was
245 considerably reduced compared to WT, virtually no recruitment signals could be measured for the
246 $W195^{5.34}L$ design (**Fig.3a, Supplementary Fig.S8**). The differences in β-arrestin recruitment were

not due to difference in the expression levels of the different mutants as they showed similar cell surface expression as assessed by ELISA (**Supplementary Fig.S6**).

Consistent with a role for β-arrestin in ERK activation[34],the W195$^{5.34}$L design showed a reduced level of ERK phosphorylation compared to WT, suggesting that the scaffolding function supported by β-arrestin was affected (**Fig.3b**). Because of a high background CXCL12-promoted ERK activity in HEK293 cells, the ERK assay was performed in U87.GM cells that lack endogenous CXCR4 in which WT- and W195$^{5.34}$L-CXCR4 were heterologously expressed at equivalent expression levels (**Supplementary Fig.S6e).** These results validated our designed functional quaternary switch and revealed that selective receptor signaling functions can be modulated by specific oligomeric structures.

New structural mechanism of GPCR-mediated biased signaling

Overall, our designs reveal an unforeseen structural mechanism of GPCR-mediated biased signaling. Molecular determinants of biased signaling identified so far were primarily encoded by specific sequence motifs and conformations of receptor monomers[35]. However, Gi-mediated CXCR4 signaling triggering important functions such as chemotaxis was recently found to depend on the formation of specific receptor nanoclusters at the cell surface[23]. These oligomers are controlled by specific structural motifs on the lipid-exposed intracellular surface of TMH6, that is remote from the dimer interface studied here (**Fig.4a**). Mutations of the corresponding residues on TMH6 resulted in nanocluster-defective receptor variants with severely impaired Gi-mediated signaling, suggesting that this CXCR4 oligomerization surface constitutes a Gi bias signaling switch. On the other hand, our study demonstrates that the extracellular dimerization surface primarily constituted by TMH5 residues controls the selective recruitment of the other main class of GPCR signaling and regulating partners, β-arrestin. Since Gi coupling remains insensitive to the

precise dimer structure mediated by TMH5 contacts, we propose that this binding surface constitutes a β-arrestin bias signaling switch.

Although we cannot exclude the possibility that the compromised ability of the closed-dimer stabilized receptor to recruit β-arrestin may be linked to a defect in their phosphorylation, our modeling suggest that the active close-dimer conformation prevent the engagement of the β-arrestin finger loop of the receptor by the cradle core of the receptor through steric hindrance (**Supplementary Fig.S4**).

Since the structural motifs identified at the surface of CXCR4 monomers control a key signaling pathway conserved in most GPCRs, we wondered whether similar binding surfaces could be identified in other receptors. We first performed a sequence alignment of CXCR4s from various organisms and found that the native residues at the designed dimerization hotspot positions were highly conserved in CXCR4s through evolution, supporting an important functional role for this region of the receptor (**Fig.4b**). Strikingly, a similar analysis revealed that these positions are poorly conserved in other human chemokine receptors with the exception of W195$^{5.34}$ (**Fig.4b**). Interestingly, while no other chemokine receptors have been crystallized in a dimeric form involving TMH5-mediated contacts, the position of W5.34 in CXCR1 (2LNL), CCR2 (5T1A), CCR5 (4MBS) and Y5.34 in chemokine-related US28 (4XT1) was found to be superimposable to that in CXCR4 (**Fig.4c**). We also found conserved aromatic residues at position 5.34 in P2Y and other peptide-binding receptors which are known to dimerize (**Fig.4b**).

Since a single mutation at that position is sufficient to disrupt β-arrestin recruitment, these findings suggest that this particular position may constitute a general β-arrestin bias signaling switch for several class A GPCRs. To validate this hypothesis, we investigated the effect of

297 mutating the native aromatic residue at position 5.34 in two additional strongly dimerizing GPCRs,
298 the mu opioid receptor (μOR) and the type-2 vasopressin receptor (V2R), that are amongst the
299 peptide-binding receptor subfamily with an aromatic residue at position 5.34 and 5.33 respectively.
300 The native tryptophan was mutated to an alanine to assess its role on the receptor signaling
301 functions. Receptor signaling was measured using distinct BRET sensors monitoring β-arrestin and
302 Gαo for μOR, and β-arrestin, and Gs for V2R (**Fig.5**). While signaling through the G proteins was
303 affected by the W5.33A mutation in V2R and W5.34A in μOR (V2R: 85% and μOR: 61% of WT
304 efficacy, **Fig.5**), μOR and V2R receptor mutants were more strongly impaired in β-arrestin signaling
305 (< 25% of WT efficacy, **Fig.5**). These results indicate that this aromatic residue largely controls the
306 signaling efficacy through the β-arrestin pathway in distinct peptide-binding GPCRs and may
307 constitute a general signaling switch in several GPCRs.
308
309 To better understand the structure-function underpinnings of the mutational effects and
310 assess whether a common structural mechanism underlies W5.34's function in the studied peptide-
311 binding receptors, we investigated the structural impact of the tryptophan to alanine substitution.
312 We focused our analysis on μOR, as a broad range of structural and functional evidence indicate
313 that this receptor strongly homodimerizes in cell membranes[36,37]. In particular, a high-resolution
314 structure of the murine μOR in the inactive state revealed a homodimer stabilized by an extensive
315 binding interface between TMH5 and TMH6. Using our computational quaternary structure
316 modeling approach, we modeled WT and W5.34A μOR homodimers in active signaling complexes
317 bound to either G-protein Gi or β-arrestin. Our simulations revealed that μOR in the active signaling
318 state mainly adopts a major "open" and a minor "wide-open" homodimer conformational state
319 (**Supplementary Fig.S9**). The open dimer form of μOR was found to strongly bind to β-arrestin,
320 while the wide-open dimer interacted considerably less well with that protein (**Supplementary**
321 **Table S3**). By contrast, both homodimer conformations were able to strongly recruit Gi. W5.34 was

found at the dimer interface of all μOR homodimers but involved in different sets of interactions. Consequently, the W5.34A mutation displayed distinct effects on the dimer structures, destabilizing the major open dimer and stabilizing the minor wide-open form (**Supplementary Table S3**). The simulations corroborate the experimental observations and provide a structural explanation as to why W5.34A preferentially affects signaling through β-arrestin.

Overall, our findings imply that W5.34 controls the signaling of CXCR4 and μOR through the β-arrestin signaling pathway by acting as a conformational switch of quaternary structure and suggest that this structural mechanism may be common to several GPCRs. However, this may not be a universal rule since other GPCRs (e.g. rhodopsin, beta 1 adrenergic, EP3, kappa opioid receptors) may dimerize through alternative dimer interfaces involving TMHs 1,2,7 and 8 (e.g. PDB: 6OFJ, 3CAP, 4GPO, 6AK3, 4DJH) and use other functional selectivity switches.

Conclusion

Membrane protein oligomers are ubiquitously observed in cell membranes and have been widely investigated using structural, spectroscopic and mutagenesis approaches[38]. However, how specific self-associations and quaternary structures control selective protein functions has remained elusive for many classes of multi-pass membrane proteins. We developed a general computational modeling and design approach that enabled the precise design of binding surfaces and interactions to perturb native or create novel receptor oligomeric structures and associated functions.

A large fraction of GPCRs can activate multiple signaling pathways. This promiscuity has proven a challenge for the development of selective therapeutics since drugs targeting the canonical extracellular ligand binding site of GPCRs often trigger several intracellular functions leading to undesirable side-effects[39]. In this study, we uncovered and engineered hotspot

dimerization conformational switches on the extracellular side of CXCR4 and μOR that controlled the precise receptor dimeric structure and the selective activation of intracellular signaling pathways. Interestingly, we also identified a biased signaling hotspot at the same location in another strongly homodimerizing peptide-binding GPCR, V2R, but for which a high-resolution structure has not yet been determined. Altogether, the results suggest that the position 5.34 may act as a general dimerization conformational switch that controls biased signaling in several GPCRs.

The extracellular locations of these biased signaling switches suggest that the sites are druggable. The signaling regulatory mechanism controlled by specific receptor oligomeric structures emerging from our study opens new avenues for selective pharmacological treatments that do not perturb receptor monomeric structures and associated signaling functions.

Overall, our approach should prove useful for designing multi-pass membrane protein associations with novel structures and functions, and expand protein design toolkits for engineered cell-based therapies and synthetic biology applications.

References

1. Felce, J.H., Davis, S.J. & Klenerman, D. Single-Molecule Analysis of G Protein-Coupled Receptor Stoichiometry: Approaches and Limitations. Trends Pharmacol Sci 39, 96-108 (2018).

2. Felce, J.H. et al. Receptor Quaternary Organization Explains G Protein-Coupled Receptor Family Structure. Cell Rep 20, 2654-2665 (2017).

3. Bouvier, M., Heveker, N., Jockers, R., Marullo, S. & Milligan, G. BRET analysis of GPCR oligomerization: newer does not mean better. Nat Methods 4, 3-4; author reply 4 (2007).

4. Milligan, G. & Bouvier, M. Methods to monitor the quaternary structure of G protein-coupled receptors. FEBS J 272, 2914-25 (2005).

5. Huang, J., Chen, S., Zhang, J.J. & Huang, X.Y. Crystal structure of oligomeric beta1-adrenergic G protein-coupled receptors in ligand-free basal state. Nat Struct Mol Biol 20, 419-25 (2013).

6. Manglik, A. et al. Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. Nature 485, 321-6 (2012).

7. Wu, B. et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science 330, 1066-71 (2010).

8. Liang, Y. et al. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. J Biol Chem 278, 21655-21662 (2003).

9. Fotiadis, D. et al. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. Nature 421, 127-8 (2003).

10. Armando, S. et al. The chemokine CXC4 and CC2 receptors form homo- and heterooligomers that can engage their signaling G-protein effectors and betaarrestin. FASEB J 28, 4509-23 (2014).

11. Percherancier, Y. et al. Bioluminescence resonance energy transfer reveals ligand-induced conformational changes in CXCR4 homo- and heterodimers. J Biol Chem 280, 9895-903 (2005).

12. Whorton, M.R. et al. A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. Proc Natl Acad Sci U S A 104, 7682-7 (2007).

13. Whorton, M.R. et al. Efficient coupling of transducin to monomeric rhodopsin in a phospholipid bilayer. J Biol Chem 283, 4387-94 (2008).

14. Meral, D. et al. Molecular details of dimerization kinetics reveal negligible populations of transient μ-opioid receptor homodimers at physiological concentrations. Sci Rep 8, 7705 (2018).

15. Mondal, S. et al. Membrane driven spatial organization of GPCRs. Sci Rep 3, 2909 (2013).

16. Gahbauer, S. & Bockmann, R.A. Membrane-Mediated Oligomerization of G Protein Coupled Receptors and Its Implications for GPCR Function. Front Physiol 7, 494 (2016).

17. Kobayashi, H., Ogawa, K., Yao, R., Lichtarge, O. & Bouvier, M. Functional rescue of beta-adrenoceptor dimerization and trafficking by pharmacological chaperones. Traffic 10, 1019-33 (2009).

18. Salahpour, A. et al. Homodimerization of the beta2-adrenergic receptor as a prerequisite for cell surface targeting. J Biol Chem 279, 33390-7 (2004).

19. Han, Y., Moreira, I.S., Urizar, E., Weinstein, H. & Javitch, J.A. Allosteric communication between protomers of dopamine class A GPCR dimers modulates activation. Nat Chem Biol 5, 688-95 (2009).

20. Guo, W., Shi, L., Filizola, M., Weinstein, H. & Javitch, J.A. Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation. Proc Natl Acad Sci U S A 102, 17495-500 (2005).

21. Fung, J.J. et al. Ligand-regulated oligomerization of beta(2)-adrenoceptors in a model lipid bilayer. EMBO J 28, 3315-28 (2009).

22. Yao, X.J. et al. The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex. Proc Natl Acad Sci U S A 106, 9501-6 (2009).

23. Martinez-Munoz, L. et al. Separating Actin-Dependent Chemokine Receptor Nanoclustering from Dimerization Indicates a Role for Clustering in CXCR4 Signaling and Function. Mol Cell 70, 106-119 e10 (2018).

24. Das, R. & Baker, D. Macromolecular modeling with rosetta. Annu Rev Biochem 77, 363-82 (2008).

25. Procko, E. et al. A computationally designed inhibitor of an Epstein-Barr viral Bcl-2 protein induces apoptosis in infected cells. Cell 157, 1644-1656 (2014).

26. Silva, D.A. et al. De novo design of potent and selective mimics of IL-2 and IL-15. Nature 565, 186-191 (2019).

27. Joh, N.H. et al. De novo design of a transmembrane Zn(2)(+)-transporting four-helix bundle. Science 346, 1520-4 (2014).

424 28. Wang, Y. & Barth, P. Evolutionary-guided de novo structure prediction of self-associated
425 transmembrane helical proteins with near-atomic accuracy. Nat Commun 6, 7196 (2015).
426 29. Yin, H. et al. Computational design of peptides that target transmembrane helices. Science
427 315, 1817-22 (2007).
428 30. Qin, L.K., I.; Holden, L.G.; Wang, C.; Zheng, Y.; Zhao, C.; Fenalti, G.; Wu, H.; Han, G.W.;
429 Cherezov, V.; Abagyan, R.; Stevens, R.C.; Handel, T.M. Structural biology. Crystal structure of the
430 chemokine receptor CXCR4 in complex with a viral chemokine. Science 347, 1117-22 (2015).
431 31. Feng, X., Ambia, J., Chen, K.M., Young, M. & Barth, P. Computational design of ligand-
432 binding membrane receptors with high selectivity. Nat Chem Biol 13, 715-723 (2017).
433 32. Angers, S. et al. Detection of beta 2-adrenergic receptor dimerization in living cells using
434 bioluminescence resonance energy transfer (BRET). Proc Natl Acad Sci U S A 97, 3684-9 (2000).
435 33. Namkung, Y. et al. Monitoring G protein-coupled receptor and beta-arrestin trafficking in live
436 cells using enhanced bystander BRET. Nat Commun 7, 12178 (2016).
437 34. Luttrell, L.M. et al. Manifold roles of β-arrestins in GPCR signaling elucidated with siRNA
438 and CRISPR/Cas9. Sci Signal 11(2018).
439 35. Wootten, D. et al. The Extracellular Surface of the GLP-1 Receptor Is a Molecular Trigger for
440 Biased Agonism. Cell 165, 1632-43 (2016).
441 36. Manglik, A. et al. Crystal structure of the μ-opioid receptor bound to a morphinan antagonist.
442 Nature 485, 321-6 (2012).
443 37. Vilardaga, J.P. et al. Conformational cross-talk between alpha2A-adrenergic and mu-opioid
444 receptors controls cell signaling. Nat Chem Biol 4, 126-31 (2008).
445 38. Kufareva, I. et al. A novel approach to quantify G-protein-coupled receptor dimerization
446 equilibrium using bioluminescence resonance energy transfer. Methods Mol Biol 1013, 93-127
447 (2013).
448 39. Manglik, A. et al. Structure-based discovery of opioid analgesics with reduced side effects.
449 Nature 537, 185-190 (2016).

Material and Methods

Modeling CXCR4 inactive state monomer and homodimer structures

The X-ray structure of the antagonist-bound human chemokine receptor CXCR4 homodimer (PDB code: 3ODU) served as a starting template for modeling the CXCR4 inactive state monomer. After removal of detergent and lipid molecules, the two receptor molecules were separated from the dimer structure and the region corresponding to the binding interface was relaxed in implicit lipid membrane environment (The RMSD between the relaxed structure and the starting antagonist bound X-ray structure was 0.1 Å over Cα atoms). The lowest energy relaxed CXCR4 monomer structure was selected as a representative model of the CXCR4 inactive state monomer.

The symmetric flexible docking mode of RosettaMembrane[28] involving inter-monomer rigid-body movements and intra-monomer conformational flexibility was then applied to model CXCR4 homodimer inactive state structures. 10,000 homodimer models were generated starting from the selected CXCR4 inactive state monomer model. The 10% lowest homodimer interface energy CXCR4 homodimer models were selected and then filtered by inter-protomer angles to select quaternary structures that had both optimal homodimer binding energies and proper membrane insertion. Specifically, the relative orientation of the monomers in the X-ray homodimer structure is characterized by an inter-helical angle between helix 5 of 52 degrees which ensures optimal membrane embedding. Hence, all models where such angle was no larger than 85 degrees and no less than -50 degrees were considered compatible with proper embedding. Overall, 80% of the models selected by interface energy were kept after applying this relative orientation filter.

These homodimer models were then clustered by dimer-specific geometric parameters across the dimer binding interface (i.e. θ and d, as described in **Supplementary Fig.S2**) for major dimer orientation analysis. We used the hdbscan-clustering method, which is a density-based clustering method based on hierarchical density estimates[40]. A majority of the models clustered in two large

476 families of distinct dimer conformations (i.e. closed or open) characterized by very different
477 interhelical angles and distances between TMH5 as described in **Supplementary Fig.S2**. The
478 lowest energy structure from each cluster was selected as the representative model of each specific
479 (i.e. closed or open) homodimer conformation.
480
481 **Quaternary structure assembly of CXCR4 active state dimer complexes bound to G-protein**
482 **or β-arrestin**
483
484 The general strategy for modeling G-protein or β-arrestin-bound CXCR4 active state homodimers
485 involved the following steps: First, the CXCR4 monomer was modeled in the active state
486 conformation and then assembled into homodimers. Lastly, the G-protein Gi and β-arrestin-2 were
487 also modeled and assembled onto the CXCR4 active state dimers to generate an optimal signaling
488 complex. The same procedure was applied to model the WT and designed CXCR4 quaternary
489 structures.
490
491 *Modeling CXCR4 active state monomer structures*
492 We applied RosettaMembrane homology modeling method[31,41] to model the agonist-bound
493 conformations of a CXCR4 active state monomer. We used the nanobody and chemokine-bound
494 active state viral GPCR (PDB code: 4XT1, Sequence identity = 30%) as a template because it
495 displayed the highest sequence homology to CXCR4 among active state GPCR structures. 50,000
496 models of CXCR4 monomer were generated and the 10% lowest energy models were clustered
497 based on Cα RMSD. The cluster centers of the top 10 largest clusters were used to build models of
498 active state CXCR4 homodimer.
499
500 *Modeling CXCR4 active state homodimer structures*

Active state CXCR4 homodimer structures were modeled using the same approach than for the inactive state models with the exception that 10 starting active state monomer models were considered. The symmetric flexible docking mode of RosettaMembrane[28] was applied on each monomer model, and, after filtering by interhelical angle, all homodimer models were pooled together prior to the final clustering step. The lowest interface energy decoy from the largest clusters were selected for modeling CXCR4-dimer-β-arrestin-2 or CXCR4-dimer-Gi complexes.

Modeling GPCR-bound β-arrestin-2 conformations

Arrestin binding to GPCRs mainly involves 3 loops which undergo significant conformational changes upon receptor binding. Since β-arrestin-2 was never crystallized in complex with a GPCR, to increase the chance of identifying optimal CXCR4-β-arrestin-2 binding modes, we modeled the receptor-bound conformations of β-arrestin-2 by homology to that of the close homolog arrestin-1 bound to Rhodopsin (Sequence identity = 60%, PDB code: 4ZWJ) using Rosetta homology modeling. 10,000 models were generated, and the lowest 10% energy models were clustered. The lowest energy models of the largest clusters (containing at least 2% of the population) were used to generate CXCR4-dimer-β-arrestin-2 complex structures.

Assembling β-arrestin-2-CXCR4 dimer active state complexes

A total of eight β-arrestin-2 models were selected for optimal docking assembly to each selected active CXCR4 homodimer models. Starting conformations were generated by aligning one subunit of the CXCR4 dimer to Rhodopsin receptor and β-arrestin-2 to visual arrestin in the Rhodopsin-arrestin X-ray structure (PDB code: 4ZWJ). 5,000 models were generated by flexible docking perturbation of the starting structure to optimize the interaction between the different domains of β-arrestin-2 and the intracellular regions of the CXCR4 homodimers. The complexes with the lowest interface energy were selected as representative conformations of β-arrestin-2 bound to one CXCR4 homodimer structure model.

Assembling Gi-CXCR4 dimer active state complexes

The α-subunit of the Gi protein (Gαi)-CXCR4 dimer structure was modeled before the first X-ray structure of a GPCR-Gi complex was solved. The GPCR-bound active state conformation of the C-terminal domain of Gi (including the α5 C-terminal helix) was modeled from the Gs structure bound to the beta2 adrenergic receptor (B2AR) (PDB code: 3SN6, Sequence Identity > 40%). The C-terminal domain model of Gi was grafted onto the N-terminal domain of the GTPγS bound structure of Gi protein α-subunit (PDB code: 1GIA) to model the full length GPCR-bound conformation of Gαi. 10,000 models were generated and the lowest energy 10% models by total energy were clustered. The lowest-energy decoys in the largest clusters were used as representative active state Gαi to assemble CXCR4-dimer-Gαi complex structures.

The starting position of Gαi for docking onto CXCR4 was generated by aligning Gαi and CXCR4 to the β-2 adrenergic receptor and Gs protein α-subunit, respectively in their bound active state structure (PDB code: 3SN6). 5,000 models were generated through perturbation of the starting structures to refine the interaction between the downstream effector and CXCR4 models. The docked structures were filtered by interface energy (lowest 1% effector-interface energy) and clustered. The models with the lowest effector-docking interface energy in the largest clusters were selected as representative conformation for further analysis.

Computational design of CXCR4 dimer conformations with distinct stabilities

Inactive and active state open-dimer and closed-dimer models of the WT receptor served as starting templates for all design calculations performed using the implicit lipid membrane model of RosettaMembrane[28,42,43]. Positions at the interface of the two protomers were systematically scanned *in silico* (~20 positions, 20^{20} possible combinations) to search for mutations that would stabilize the open-dimer or closed-dimer conformation without modifying significantly the stability of

each monomer. This strategy ensured that designed mutations would solely affect the structural and functional properties associated with receptor dimerization. Hence, mutations were selected according to the quantity $\Delta E_{interface} = (E_{interface})_{design} - (E_{interface})_{WT}$ where $E_{interface} = E_{dimer} - 2 * E_{monomer}$ providing $\Delta E_{monomer} = (E_{monomer})_{design} - (E_{monomer})_{WT}$ remained minimal. Any designed mutation that had minimal effects on $\Delta E_{interface}$ (less than 1.0 REU) and/or significantly affected $\Delta E_{monomer}$ (by more than 1.0 REU) was systematically discarded. After each step of sequence selection, the structure of the designed binding interface was refined and optimized using a Monte Carlo Minimization protocol sampling all conformational degrees of freedom.

The distribution between dimer conformations for the final selected designs and the associated functional effects on the binding to G-protein versus β-arrestin were obtained by performing a final round of docking simulations where designed monomers were assembled into GPCR dimers and into complex with G-proteins or β-arrestin as described above for the WT receptor.

Modeling μOR active state dimer structures. Starting from the active state monomeric structure of μOR bound to the G-protein Gi (PDB code: 6DDF), homodimer models of the WT receptor were obtained using the symmetric docking mode of RosettaMembrane described above using the same parameters than for CXCR4. Representative lowest energy homodimer μOR models were selected to assemble Gi and β-arrestin complexes as described above for CXCR4. The bound Gi structure resolved in the 6DDF structure was used for docking onto μOR dimer models. Final models were selected and analyzed using the same unbiased geometric and energetic criteria as for CXCR4. The effect of the W5.34A mutation was obtained by calculating the quantity $\Delta E_{interface}$ after assembling the mutated monomers into GPCR dimers as described in the computational design section.

Calculation of dimerization propensity. The docking simulations performed using the software Rosetta do not reliably calculate free energies of protein associations because they neglect conformational and configurational entropies for example and just provide the enthalpy of a static structure. Nevertheless, differences in dimerization propensities between receptor variants can be estimated from the dimer binding energy calculated for the selected open and closed dimer conformation as follows. In absence of free energies for the monomer and dimer species, we define a reference state, that of the lowest energy primary dimer conformation of the WT receptor, i.e. the open dimer: $(\Delta E_{interface,\ O})_{WT}$. We first calculate the difference in dimer binding energies for each variant and conformation from WT as follows:

$$(\Delta\Delta E_{interface,\ Y})_X = (\Delta E_{interface,\ Y})_X - (\Delta E_{interface,\ O})_{WT}. \quad (1)$$

where X represents WT or any designed receptor variant and Y = O or C and corresponds to the open and closed conformation, respectively.

The Boltzmann factors describing the probability of a variant X to occupy the dimer state in a specific conformation ($(PD_Y)_X$, dimerization propensity) relative to WT can be derived as follows:

$$(PD_Y)_X = \exp(-(0.5((\Delta\Delta E_{interface,\ Y})_X) / RT). \quad (2)$$

where the 0.5 factor roughly converts Rosetta Energy Units to kcal/mol. RT is the thermal scaling factor and equal to 0.593 $kcal.mol^{-1}$.

The sum of the Boltzmann factors for the open and closed conformation are calculated in the inactive and active state (reported in **Supplementary Table S2**) and provides an indication whether a variant has a lower or higher propensity to occupy the dimer state than WT.

Reagents and plasmids

CXCL12 was purchased from Cedarlane. Forskolin, isobutylmethyl xanthine (IBMX), AVP and met-enkephalin were purchased from Sigma. The following plasmids were already described: ~~and~~ HA-CXCR4[44], β-arrestin2-LucII[45], β-arrestin2-YFP[46], $G\alpha_{i1}$-91RLucII[44], $G\alpha_{s}$-117RLucII[47], $G\alpha_{oA}$-91RLucII[48], GFP10-$G_{\gamma 1}$[49], GFP10-$G_{\gamma 2}$[50], GFP10-EPAC-RLucII[51], rGFP-CAAX[33]. The cloning of CXCR4-RLuc and CXCR4-YFP in pcDNA3.1 was previously described[11]. In the present study, CXCR4-RLuc and CXCR4-YFP were amplified and modified by PCR at the N-terminal end to add a myc epitope (EQKLISEEDL) or a HA epitope (YPYDVPDYA), respectively. Myc-CXCR4-RLuc and HA-CXCR4-YFP segments were then subcloned into pIREShyg3 (BsrG1/AflII) and pIRESpuro3 (Nhe1/AflII) respectively. The human μOR and V2R were amplified with a SNAP tag at their N-terminal (NEB) and subcloned in the pcDNA4/TO plasmid (Invitrogen). All the mutants were obtained by site directed mutagenesis using the extension of overlapping gene segments by PCR technique and validated by sequencing.

Cell culture and Transfections

Human Embryonic Kidney 293 T cells (HEK293T cells) were cultured using Dulbecco's Modified Eagle Medium (DMEM with L-glutamine from Wisent) supplemented with 10% vol/vol Fetal Bovine Serum (Wisent). The day before transfection, 600,000 cells were seeded in 6-well plates. Transient transfections were performed using Polyethylenimine 25Kd linear (PEI, Polysciences) as transfection agent, with a 3:1 PEI:DNA ratio.

U87.MG cells stably expressing HA-CXCR4 and HA-CXCR4-W195$^{5.34}$L mutant (U87.CXCR4 and U87.CXCR4-W195$^{5.34}$L, respectively) were established by transfection of pIRES-HA-CXCR4 and pIRES-HA-CXCR4-W195$^{5.34}$L and subsequent cell sorting for equivalent surface expression levels using 1:1000 Alexa Fluor 488-labeled anti-HA antibody (clone 16B12, Biolegend). U87 cells were grown in Dulbecco's modified Eagle medium (Themo Fischer Scientific) supplemented with 15%

vol/vol foetal bovine serum and penicillin/streptomycin (100 Units/ml and 100 µg/ml) (Themo Fischer Scientific). U87.CXCR4 and U87.CXCR4-W195$^{5.34}$L cell lines were maintained under puromycin (0.5 µg/ml) selective pressure.

BRET measurements

Two different BRET configurations were used in this study: $BRET_{480\text{-}YFP}$ and $BRET_{400\text{-}GFP10}$. $BRET_{480\text{-}YFP}$ uses RLuc as energy donor and YFP as the acceptor (excitation peak at 488nm) and coelenterazine-h (coel-h, Nanolight Technology) was used as the substrate (emission peak at 480nm). $BRET_{400\text{-}GFP10}$ uses RLucII as energy donor and GFP10 as the acceptor (excitation peak at 400nm) and coelenterazine-400a (coel-400a, Nanolight Technology) was used as the substrate (emission peak at 400nm). Enhanced bystander BRET (ebBRET) uses RlucII as energy donor, rGFP as the acceptor and is detected using the $BRET_{480\text{-}YFP}$ configuration and Prolume Purple as the substrate (NanoLight Technology). BRET was measured with a Mithras LB940 multimode microplate reader (Berthold Technologies) equipped with a $BRET_{480\text{-}YFP}$ filters set (donor 480 ± 20 nm and acceptor 530 ± 20 nm filters) or a Tristar microplate reader equipped either with a $BRET_{480\text{-}YFP}$ filters set (donor 480 ± 20 nm and acceptor 530 ± 20 nm filters) or a $BRET_{400\text{-}GFP10}$ filters set (donor 400 ± 70 nm and acceptor 515 ± 20 nm filters). All the BRET experiments were performed at room temperature.

CXCR4 dimerization

Cells were transfected with HA-CXCR4-YFP and myc-CXCR4-RLuc, WT or mutant, and seeded in 96-well plates (Culturplate, Perkinelmer) coated with poly-L-ornithine (Sigma-Aldrich) 24h after transfection. The following day, cells were washed with Hank's Balanced Salt Solution (HBSS, Invitrogen) and incubated in HBSS supplemented with 0.1% BSA. Cells were treated with CXCL12 at the indicated times and concentrations. Coel-h (2.5 µM) was added 10 min before reading.

G protein activation

Cells were transfected with the receptor (CXCR4, V2R or μOR) and a three-component BRET-based biosensor: Gαi1-RlucII (CXCR4), Gαs117RlucII (V2R) or Gαo-RlucII (μOR) and Gβ1, and Gγ1-GFP10 (V2R and μOR) or Gγ2-GFP10 (CXCR4). BRET was then monitored as described above using coel-400a as a substrate. The dissociation of the Gα and Gβ/Gγ subunits after activation leads to a decrease in the BRET ratio.

β-arrestin engagement (direct interaction)

Cells were transfected with CXCR4-Rluc and β-arrestin2-YFP. BRET was monitored as described above using Coel-h as a substrate.

β-arrestin engagement (ebBRET)

Cells were transfected with the receptor (HA-CXCR4, SNAP-V2R or SNAP-μOR), β-arrestin2-RLucII and CAAX-rGFP. BRET was monitored as described above using Coel-h. BRET was monitored as described above using Prolume Purple (1.3 μM) as a substrate.

cAMP accumulation

Cells were transfected with HA-CXCR4 and the BRET-based biosensor GFP10-EPAC1-RlucII. BRET was then monitored as described above with the cells first washed with HBSS and then incubated in HBSS + 0.1% BSA containing 500 μM isobutylmethyl xanthine (IBMX), without or with 10 μM forskolin for 15 min, followed by agonist stimulation.

ERK phosphorylation assay monitored by HTRF

U87, U87.CXCR4 and U87.CXCR4-W195$^{5.34}$L cells were seeded in 96-well plates (1 x 10^4 cell/well). 72 hours later, culture medium was replaced with FBS-free, phenol-red free DMEM. After 4-hour

starvation, CXCL12 was added to cells at a final concentration of 10 nM and incubated for the indicated times. ERK phosphorylation was evaluated using a Homogenous Time-Resolved FRET (HTRF)-based Phospho-ERK (Thr202/Tyr204) cellular kit (Cisbio). Cells were lysed for 30 minutes with the lysis buffer provided and incubated for 2 hours with pERK1/2-specific antibodies conjugated with Eu^{3+}-cryptate donor and d2 acceptor at recommended dilutions. HTRF was measured with Tecan GENios Pro plate reader equipped with 612 ± 10 (donor) and 670 ± 25 (acceptor) filters. HTRF ratio was calculated as follows:

$$Ratio = \frac{A_{670}}{D_{612}} \times 10000$$

A_{670} = emission at 670 nm (RFU)

D_{612} = emission at 612 nm (RFU)

Elisa

To control for the cell surface expression of HA-CXCR4, HA-CXCR4-YFP and myc-CXCR4-Rluc, and their respective mutant receptors, ELISA were performed in parallel of BRET experiments, using an antibody directed at the extracellular epitope (HA or Myc). 24h after transfection, cells were seeded in 24-well plates coated with poly-L-ornithine. The day of the experiment, media was removed and a solution of PBS with 3.7% paraformaldehyde was added for 5 min. Cells were then washed 3 times with Phosphate Buffered Saline (PBS). Blocking solution (PBS + 1% BSA) was added for 45 min then replaced by PBS + 1% BSA containing HA antibody 1:1000 (12CA5, Santacruz) or Myc antibody 1:1000 (D84C12, Cell Signaling) for 45 min. After antibody addition, cells were washed three times with PBS and incubated 45 min with PBS+1% BSA containing an anti-mouse HRP antibody (1:2000). After labelling, cells were washed three times with PBS and incubated with SigmaFastOPD (SigmaAldrich) at room temperature. Reaction was stopped using

3N HCl, supernatant transferred in a 96-well plate, and reading was performed using a Spectramax multimode microplate reader (molecular devices) at 492nm.

Flow cytometry

Endogenous CXCR4 expression on the surface of HEK and U87 cells was monitored by flow cytometry using CXCR4-specific phycoerythrin-conjugated mAb 12G5 or the corresponding isotype control (1:20, R&D Systems) in a BD FACS LSR Fortessa cytometer (BD Biosciences). U87 were chosen as cellular background for the absence of endogenous CXCR4 and ACKR3, as previously demonstrated[52,53]. U87 cells stably expressing the HA-tagged CXCR4 or variants thereof were obtained following puromycin selection and subsequent single-cell sorting using BD FACSAria II cell sorter (BD Biosciences). Equivalent surface expression level was verified using an Alexa Fluor 488-conjugated anti-HA-tag mAb (clone 16B12, 1:1000 Biolegend). Flow cytometry data were analyzed using FlowJo V10 software.

Data and statistical analysis

All data were analysed using GraphPad Prism (GraphPad Software, Inc). Statistical significance between the groups was assessed with a one-way ANOVA followed by Tukey's post hoc test.

References (Material and Methods)

40. McInnes, L.H., J.; Astels, S. hdbscan: Hierarchical density based clustering. Journal of Open Source Software 2(2017).

41. Chen, K.Y., Sun, J., Salvo, J.S., Baker, D. & Barth, P. High-resolution modeling of transmembrane helical protein structures from distant homologues. PLoS Comput Biol 10, e1003636 (2014).

42. Barth, P., Schonbrun, J. & Baker, D. Toward high-resolution prediction and design of transmembrane helical protein structures. Proc Natl Acad Sci U S A 104, 15682-7 (2007).

722 43. Chen, K.Y., Zhou, F., Fryszczyn, B.G. & Barth, P. Naturally evolved G protein-coupled
723 receptors adopt metastable conformations. Proc Natl Acad Sci U S A 109, 13284-9 (2012).
724 44. Quoyer, J. et al. Pepducin targeting the C-X-C chemokine receptor type 4 acts as a biased
725 agonist favoring activation of the inhibitory G protein. Proc Natl Acad Sci U S A 110, E5088-97
726 (2013).
727 45. Paradis, J.S. et al. Receptor sequestration in response to beta-arrestin-2 phosphorylation by
728 ERK1/2 governs steady-state levels of GPCR cell-surface expression. Proc Natl Acad Sci U S A
729 112, E5160-8 (2015).
730 46. Khoury, E., Nikolajev, L., Simaan, M., Namkung, Y. & Laporte, S.A. Differential regulation of
731 endosomal GPCR/beta-arrestin complexes and trafficking by MAPK. J Biol Chem 289, 23302-17
732 (2014).
733 47. Thomsen, A.R.B. et al. GPCR-G Protein-β-Arrestin Super-Complex Mediates Sustained G
734 Protein Signaling. Cell 166, 907-919 (2016).
735 48. Busnelli, M. et al. Functional selective oxytocin-derived agonists discriminate between
736 individual G protein family subtypes. J Biol Chem 287, 3617-29 (2012).
737 49. Galés, C. et al. Probing the activation-promoted structural rearrangements in preassembled
738 receptor-G protein complexes. Nat Struct Mol Biol 13, 778-86 (2006).
739 50. Gales, C. et al. Real-time monitoring of receptor and G-protein interactions in living cells.
740 Nat Methods 2, 177-84 (2005).
741 51. Zimmerman, B. et al. Differential beta-arrestin-dependent conformational signaling and
742 cellular responses revealed by angiotensin analogs. Sci Signal 5, ra33 (2012).
743 52. Szpakowska, M. et al. Human herpesvirus 8-encoded chemokine vCCL2/vMIP-II is an
744 agonist of the atypical chemokine receptor ACKR3/CXCR7. Biochem Pharmacol 114, 14-21 (2016).
745 53. Szpakowska, M. et al. Mutational analysis of the extracellular disulphide bridges of the
746 atypical chemokine receptor ACKR3/CXCR7 uncovers multiple binding and activation modes for its
747 chemokine and endogenous non-chemokine agonists. Biochem Pharmacol 153, 299-309 (2018).

Acknowledgements

This work was supported by a grant from the National Institute of Health (1R01GM097207), a grant from the Swiss National Science Foundation (31003A_182263), a supercomputer allocation from XSEDE (MCB120101) to P.B., and a grant from the Canadian Institute for health Research (CIHR) (Foundation grant #148431) to M.B.. P.B. was also supported by funds from EPFL and the Ludwig Institute for Cancer Research. M.S. and A.C. were supported by the Luxembourg Institute of Health (LIH) MESR (grants 20160116 and 20170113) and the Luxembourg National Research Fund (INTER/FWO "Nanokine" - grant 15/10358798). J.P. had studentships from the 'Groupe de Recherche Universitaire sur le Médicament' and 'la Faculté des Études Supérieures et postdoctorales de l'Université de Montréal'. B.M. had a fellowship from the 'Fondation pour la Recherche Médicale (France). M.B. Holds a Canada Research Chair in Signal Transduction and Molecular Pharmacology. The authors thank Dr. Monique Lagacé for her critical reading of the manuscript.

Author contributions

P.B., A.C. and M.B. designed the study; X.F., R.J. and P.B. performed the modeling and design calculations; J.S.P., B.M., M.S., M.H., N.D.B. and F.M.H. performed the experiments under the supervision of M.B, A.C. and M.J.S.; All authors analyzed the data; P.B., J.S.P., X.F. and M.B. wrote the manuscript.

Data availability

The authors declare that all data supporting the findings in this study are presented within the article and its Supplementary Information files. All the biosensors can be obtained and used without limitations for non-commercial purpose with a standard academic materials transfer

agreement (MTA) on request. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability.

The details (i.e. input files and command lines) of the calculations performed in this study as well as the source codes and executables of the modeling and design methods are available from the corresponding authors upon request and will be released free of charge for academic users in the software Rosetta.

Competing Interests statement

The authors declare no competing interest

Figure legends

Figure 1. Computational modeling and design of GPCR associations with reprogrammed structures and functions. a. Framework for the modeling and design of specific receptor quaternary active state conformations eliciting various degree of functional selectivity. The WT receptor modeled in the active state is assembled into dimers and then into ternary complex with G-proteins (green) or β-arrestin (orange) to identify the distribution of quaternary conformations and their ability to recruit intracellular signaling proteins. The dimer binding interface is redesigned to stabilize and/or destabilize specific quaternary conformations. This design strategy enhances the quaternary conformational selectivity of the receptor and reprograms the functional bias of the receptor oligomer (**Supplementary Fig.S1**, **Methods**). **b.** Quaternary structural changes act as a functional switch as the closed-dimer conformation interferes with the binding of a GPCR monomer to β-arrestin. **c.** Mutations designed to selectively stabilize the CXCR4 open-dimer conformation without affecting CXCR4 monomer stability were identified in the extracellular and TMH regions. **d.** Mutations designed to selectively stabilize the CXCR4 closed-dimer conformation without affecting CXCR4 monomer stability were identified in the extracellular region. Key atomic contacts are represented as red dotted lines. **e, f.** Schematic conformational energy landscapes of CXCR4 dimerization in the inactive (**e**) and active (**f**) states. The conformational energies reported in **Supplementary Table S1** were used to plot the energy curves.

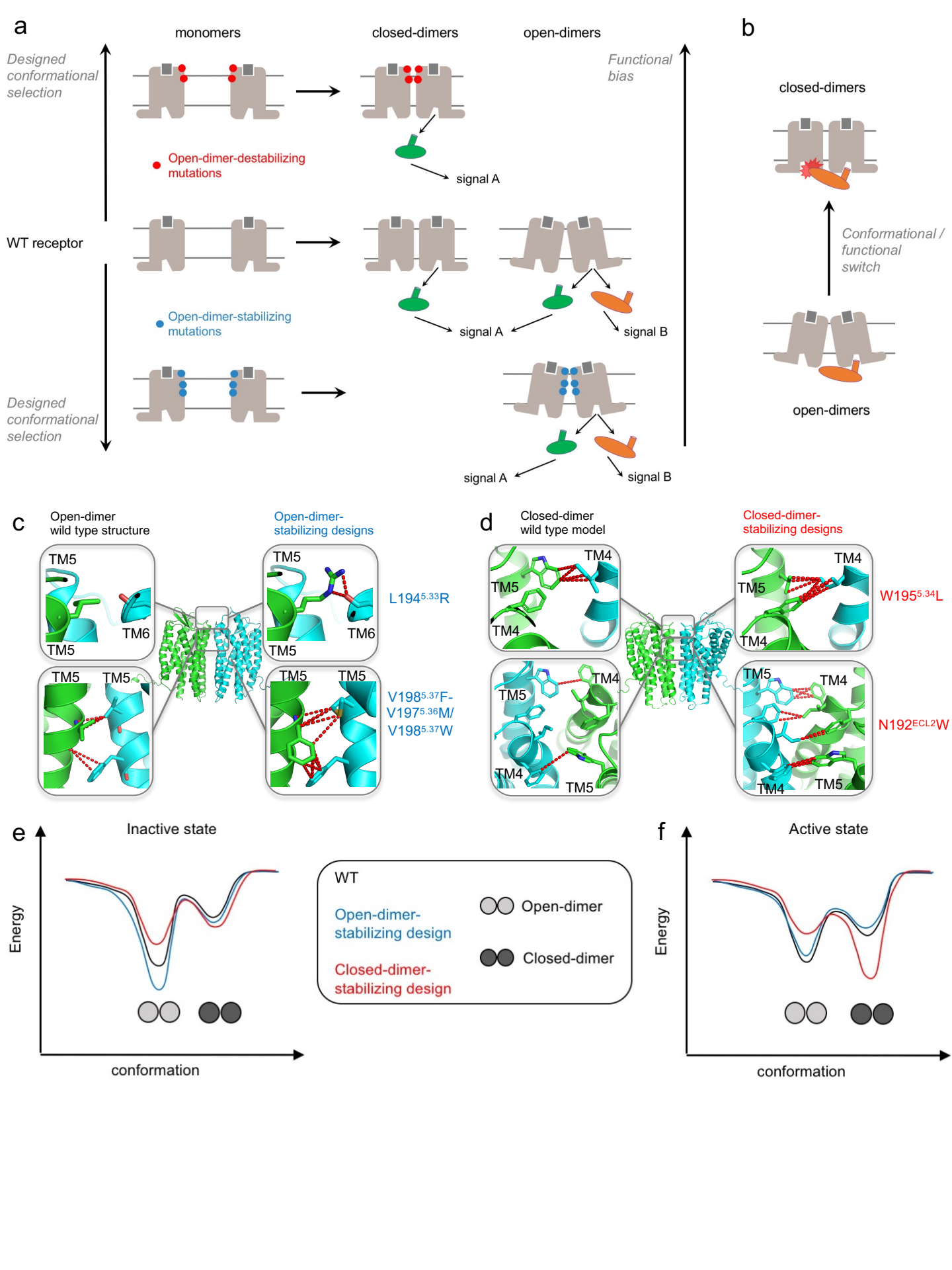
Figure 2. CXCR4 association and Gi activation. a. (Left) Schematic representation of the CXCR4 dimerization BRET-based assay. (Right) CXCR4 association measured by BRET before (black) and after agonist stimulation (grey) in HEK293T cells transfected with CXCR4-RLuc and its counterpart CXCR4-YFP, WT or mutant as indicated. $BRET_{480\text{-}YFP}$ was measured after the addition of coel-h (10 min) and CXCL12 (15 min). **b.** (Left) Schematic representation of the BRET-based ligand-induced Gi activation assay. (Right) CXCL12-promoted Gi activation measured by BRET in

HEK293T cells transfected with HA-CXCR4, WT or mutant as indicated, Gαi1-RLucII, Gβ1 and Gγ2-GFP10. $BRET_{400\text{-}GFP10}$ was measured after the addition of coel-400a (10min) and CXCL12 (3 min) **c.** (Left) Schematic representation of the BRET-based EPAC sensor to measure cAMP production. (Right) CXCL12-promoted EPAC inhibition was measured by BRET in HEK293T cells transfected with HA-CXCR4, WT or mutant as indicated, and RLuc-EPAC-YFP. $BRET_{480\text{-}YFP}$, reporting the conformation rearrangement of the EPAC sensor from an open to a close conformation, was measured after addition of coel-h (10 min) and CXCL12 (5 min). CXCR4 mutations predicted to stabilize the open-dimer or the closed-dimer conformation are annotated with a blue or red dimer symbol, respectively. Data are expressed as net BRET (calculated by subtracting background luminescence) or ΔBRET (agonist-promoted BRET). Data shown represent the mean ± SEM of at least three independent experiments. ***$p < 0.001$; n.s., not significant. † is used to compare basal values between the mutants and * to compare basal to CXCL12 treated conditions.

Figure 3. β-arrestin recruitment and ERK activation. a. (Left) Schematic representation of the BRET-based ligand-induced β-arrestin 2 (βarr2) translocation assay. (Right) CXCL12-promoted βarr2 recruitment to CXCR4 measured by BRET in HEK293T cells transfected with CXCR4-RLuc, WT or mutant as indicated, and βarr2-YFP. $BRET_{480\text{-}YFP}$ between CXCR4-RLuc and βarr2-YFP was measured after the addition of coel-h (10 min) and CXCL12 (15 min). Data are represented as ΔBRET **b.** (Left) Schematic representation of ERK activation by CXCR4. (Right) ERK phosphorylation in U87 stably expressing equivalent levels of WT and $W195^{5.34}L$ CXCR4 induced by stimulation with 10 nM CXCL12 for the indicated times was monitored by HTRF. CXCR4 mutations predicted to stabilize the open-dimer or the closed-dimer conformation are annotated with a blue or red dimer symbol, respectively. Data shown represent the mean ± SEM of at least three independent experiments.

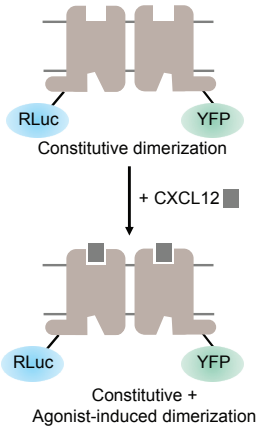
Figure 4. Distinct quaternary structures selectively control G-protein and β-arrestin recruitment. a. Surface representation of the CXCR4 inactive state monomeric structure highlighting the distinct oligomerization interfaces controlling either β-arrestin recruitment (extracellular side and TM core of TMH5, blue) or Gi activation and nanocluster formation[23] (intracellular side of TMH6, orange). **b.** The hotspot binding sites controlling CXCR4 oligomerization through TMH5 (designed residues in red) are poorly conserved in the chemokine receptor family, except for the β-arrestin signaling switch W5.34. Aromatic residues are highly enriched at position 5.34 of other dimerizing GPCR families. **c.** Conserved position and conformation of W5.34 in human chemokine receptor X-ray structures. Superposition of W5.34 conformations is shown in the center. **d.** Conservation of the β-arrestin signaling switch in the vasopressin/oxytocin receptors and opioid receptors.

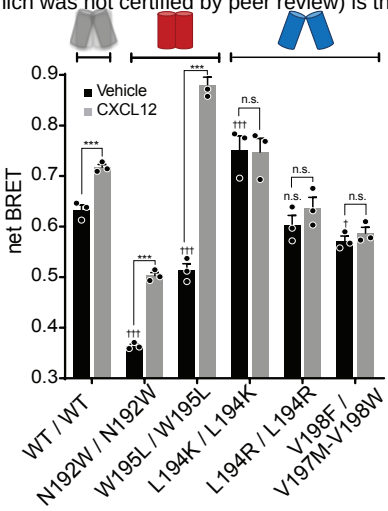
Figure 5. W5.34 is a common biased signaling switch in dimerizing peptide-binding GPCRs. a, b. Ligand-promoted Gs (V2R) or Go (μOR) and βarr2 recruitment to the membrane in the presence of V2R-WT or V2R-W200$^{5.33}$A (**a**), and μOR-WT or μOR-W230$^{5.34}$A (**b**). Gs and Go activation were detected by monitoring the dissociation between Gα and Gβγ by BRET whereas β-arrestin recruitment to the plasma membrane was assessed in cells transfected with rGFP-CAAX, WT or mutant receptors, as indicated, and βarr2-RlucII. Data are represented as ΔBRET and were normalized to the maximal response of the WT receptor. *$p < 0.05$; ***$p < 0.001$.


a
monomers
closed-dimers
open-dimers
Designed conformational selection
Functional bias
Open-dimer-destabilizing mutations
signal A
WT receptor
Open-dimer-stabilizing mutations
signal A
signal B
Designed conformational selection
signal A
signal B
b
closed-dimers
Conformational / functional switch
open-dimers
c
Open-dimer wild type structure
Open-dimer-stabilizing designs
TM5
TM6
L194$^{5.33}$R
V198$^{5.37}$F-
V197$^{5.36}$M/
V198$^{5.37}$W
d
Closed-dimer wild type model
Closed-dimer-stabilizing designs
TM4
TM5
W195$^{5.34}$L
N192^{ECL2}W
e
Inactive state
Energy
conformation
WT
Open-dimer-stabilizing design
Closed-dimer-stabilizing design
Open-dimer
Closed-dimer
f
Active state
Energy
conformation

a

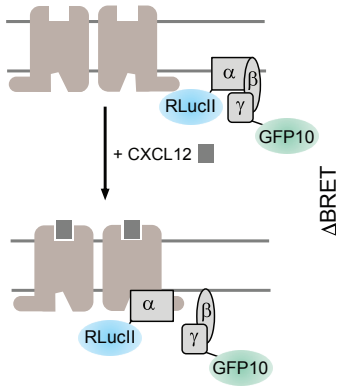
CXCR4 DIMERIZATION

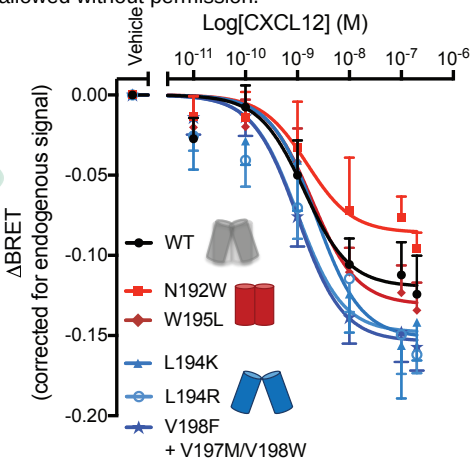




b

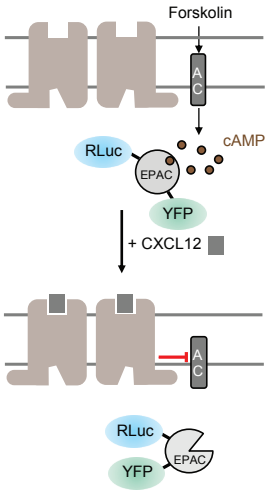
Gi ACTIVATION

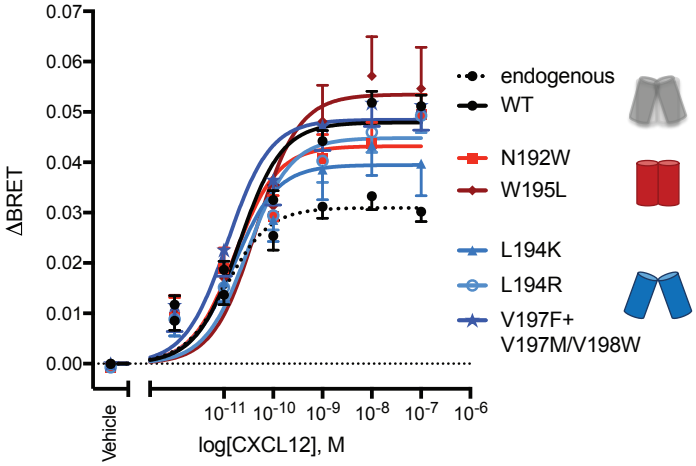


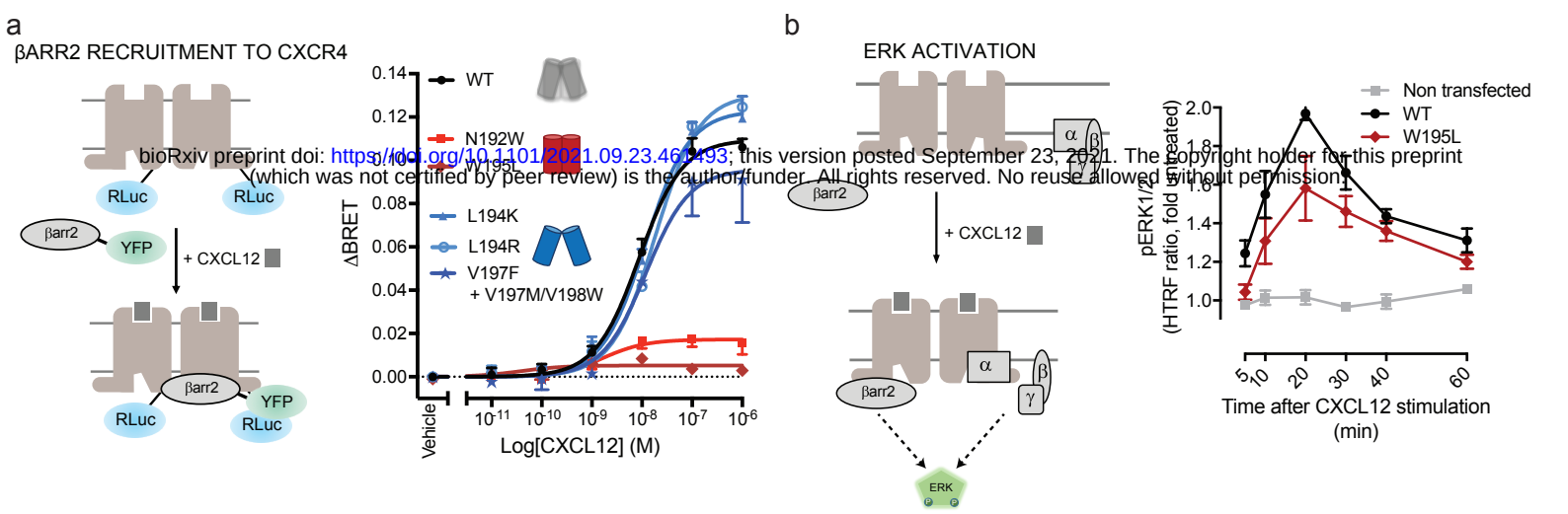


c

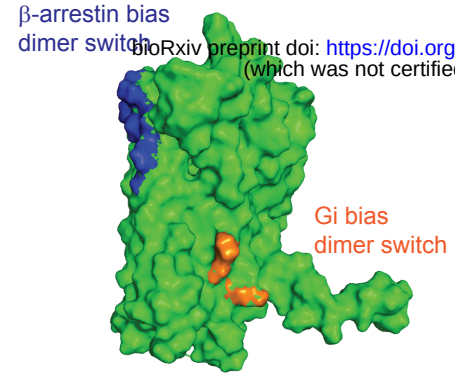
cAMP PRODUCTION



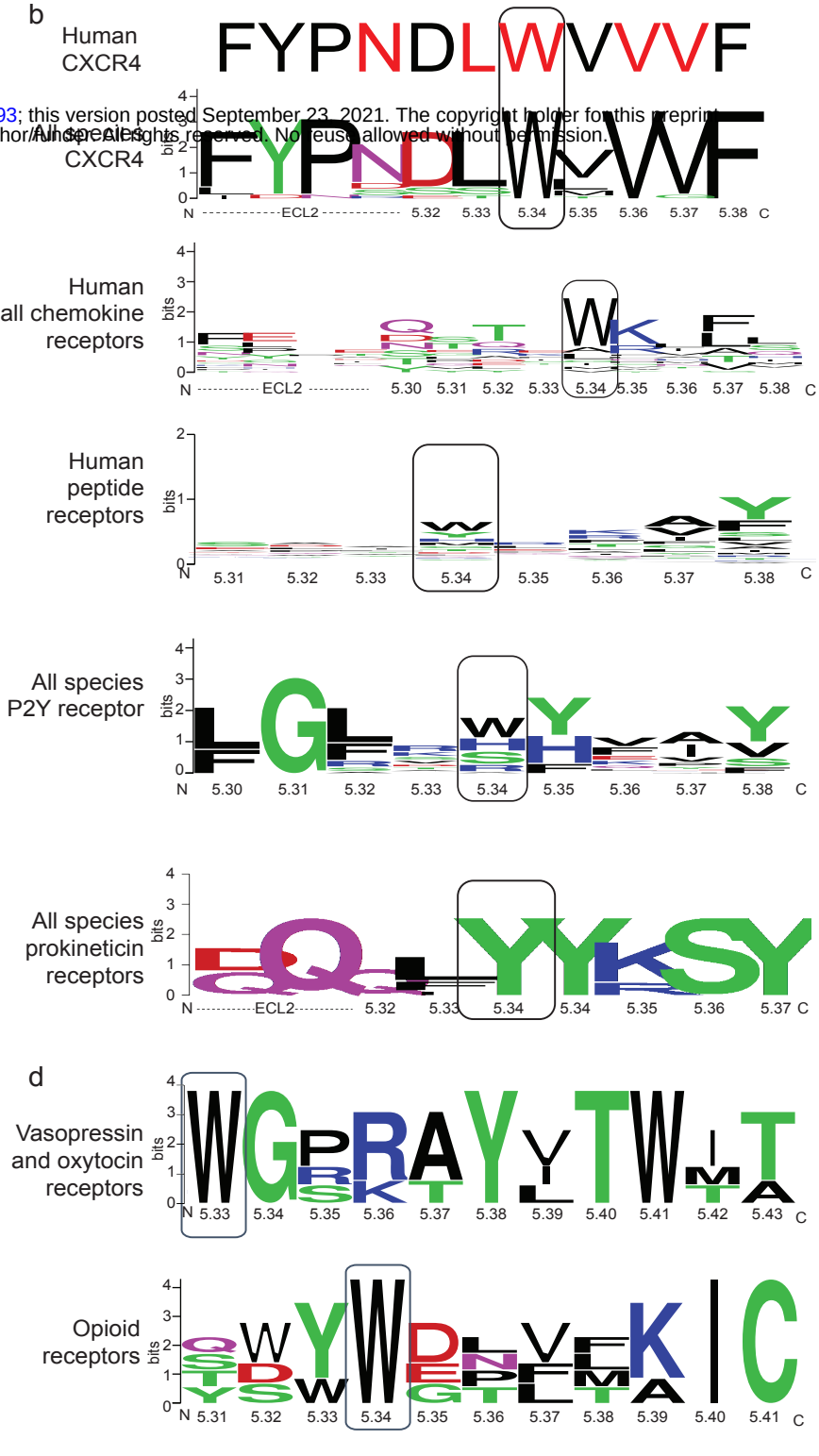




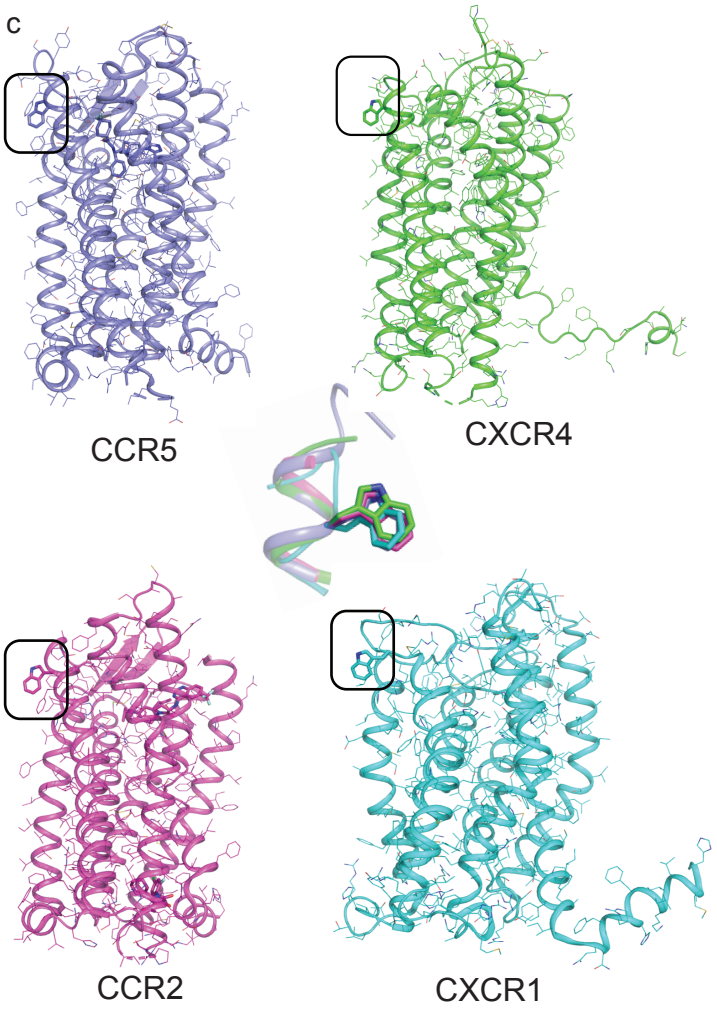
a



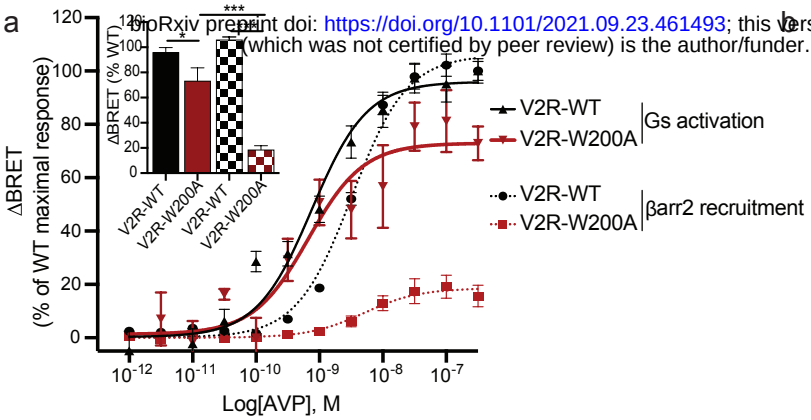
b

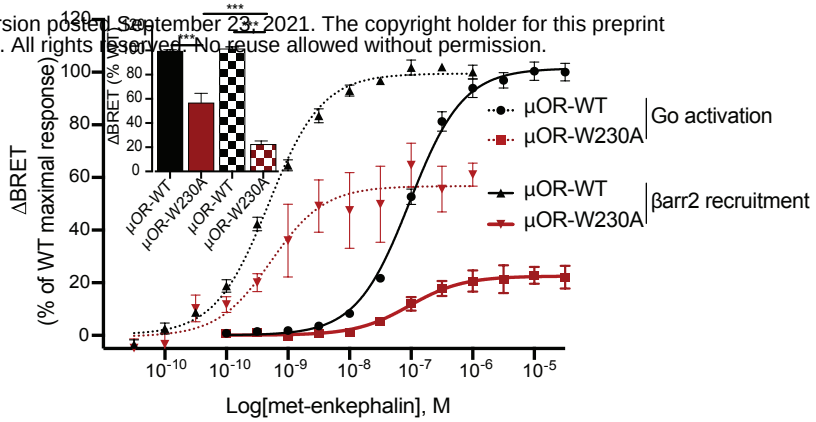


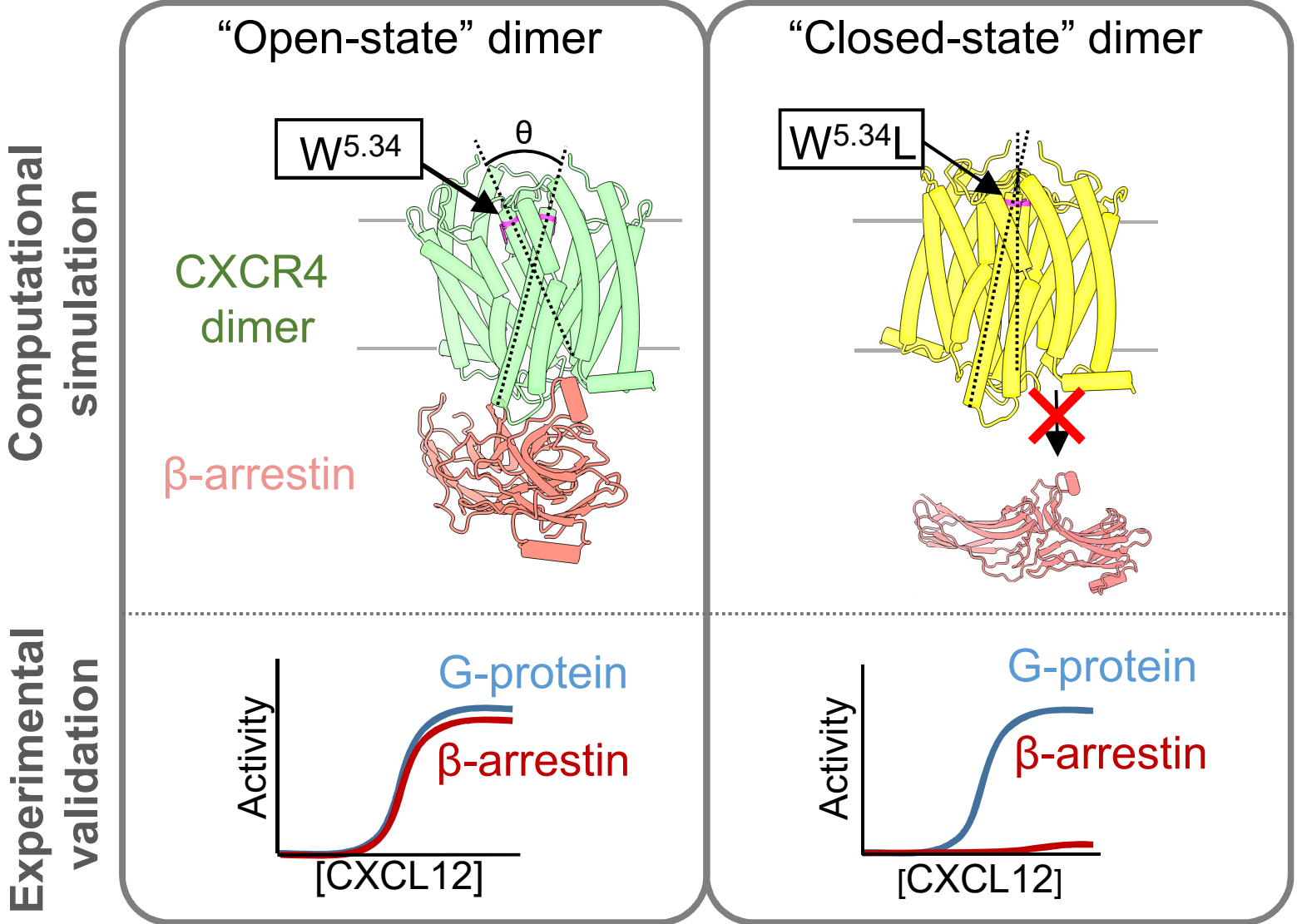
c



d






Computational simulation
Experimental validation
“Open-state” dimer
“Closed-state” dimer
W5.34
θ
W5.34L
CXCR4 dimer
β-arrestin
G-protein
β-arrestin
Activity
[CXCL12]
G-protein
β-arrestin
Activity
[CXCL12]