Research Article



Designing Structure-specific and Switchable Allosteric Effectors for GPCRs Based on the Causality and Energetics of Inherent Signaling ☆

Bingxue Dong [1,†], **Wei-Ven Tee** [1,†], **and Igor N. Berezovsky** [1,2,*]

*1 - **Bioinformatics Institute (BII),** Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01, Matrix, 138671, Singapore*

*2 - **Department of Biological Sciences (DBS),** National University of Singapore (NUS), 8 Medical Drive, 117579, Singapore*

Correspondence to Igor N. Berezovsky: Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01, Matrix, 138671, Singapore. *igorb@bii.a-star.edu.sg (I.N. Berezovsky)*

https://doi.org/10.1016/j.jmb.2025.169293

Edited by Ruth Nussinov

Abstract

The pervasiveness and versatility of G protein-coupled receptors (GPCRs) in virtually all physiological processes is based on the receptors' capability to activate intracellular signal transduction pathways in response to diverse extracellular stimuli. While the importance of GPCRs makes them the largest group of drug targets (comprising about one-third of marketed medicines), more than half of GPCRs lack pre-clinical drug candidates because of high structural conservatism of their orthosteric sites. Recognizing that the mechanisms of GPCR function and regulation are chiefly allosteric in nature, we explore their allosteric control and the potential for developing allosteric drugs with high specificity. First, we obtained a picture of allosteric signaling in 280 non-olfactory GPCRs in human and explored archetypal structure-based patterns and sequence-determined variations of allosteric communication. We showed how the causality of allosteric effects due to ligand binding can be quantified, using the β_2-adrenergic receptor (ADRB2) and glucagon-like peptide 1 receptor (GLP1R) as case studies. Implementing our *directed design protocol* for developing allosteric drug candidates, switchable agonist–antagonist pairs were obtained for GPL1R. Moreover, we showed the predictive power of our approach for identification of latent allosteric sites and the capability for *de novo* design of agonistic and antagonistic effectors. We also performed efficacy-based design of ligands, demonstrating on the homologous GLP1R and GIPR that the strength of allosteric signaling induced by an effector can determine specificity to a target receptor. Our computational framework not only provides a foundation for addressing the problem of "difficult" GPCR targets with the allosteric approach, but also allows rational design of effectors with controllable mode switching and high specificity in general. The comprehensive data on allosteric signaling in classes A, B1, B2, C, F, and T GPCRs at single-residue resolution is available in the AlloMAPS database (https://allomaps.bii.a-star.edu.sg/browse/gpcraf).

© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

G protein-coupled receptors (GPCRs) is a large and diverse family of transmembrane proteins that play a crucial role in detecting and transducing

☆ This article is part of a special issue entitled: 'Allostery in disease (2025)' published in Journal of Molecular Biology.

0022-2836/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

extracellular signals [1] important in regulation of various physiological processes, including sensory, neurotransmission, and endocrine functions [2]. In humans, GPCRs are classified into five main classes based on sequence homology and functional similarity [3]: classes A (rhodopsin-like), B1 and B2 (secretin and adhesion), C (glutamate), and F (Frizzled). Structurally, most of the class A GPCRs share a conserved architecture comprising seven transmembrane helices (TM1–TM7) connected by three intracellular and three extracellular loops, forming the transmembrane domain (TMD). The class A GPCRs represents the largest and most extensively studied class, comprising over 286 non-olfactory receptors [3]. The class A receptors are prime drug targets in pharmacology of multiple pathologies, for instance cardiovascular diseases, psychiatric disorders, metabolic syndromes, and cancers [2]. Examples of established targets include the β_2-adrenergic receptor (ADRB2) involved in cardiovascular and respiratory pathophysiology [4] and chemokine receptors known to play critical roles in inflammatory, autoimmune disorders, tumor microenvironment formation, and infectious diseases [5]. The class B1 GPCRs are receptors of peptide hormones working in numerous physiological processes. For example, the glucagon receptors regulate appetite and energy homeostasis, representing prominent drug targets for various clinical disorders, such as type 2 diabetes, obesity, and depression [6,7].

Despite some clinical successes, a majority of therapeutically important GPCRs remain undrugged. The main reason why various GPCRs are considered as difficult or even "undruggable" targets can be attributed to the requirement on drug selectivity upon targeting the conserved binding sites for endogenous ligands [8,9]. The endogenous binding sites are also termed as orthosteric sites. Many initially promising agonists and antagonists obtained for the orthosteric sites were found to cause severe off-target effects in related receptors and toxicity in subsequent clinical trials [10]. Moreover, the B1 peptide receptors pose additional challenges in drug discovery as their natural ligands bind to large pockets spanning TMD and the extracellular domain [11]. While injectable peptide drugs have demonstrated clinical success [12] in targeting class B1 receptors, such as glucagon-like peptide-1 receptor (GLP1R) - a highly promising target in diabetes management [13], the development of small-molecule agonists remains limited. Exactly for these reasons, allostery has emerged as a promising strategy to alleviate the challenges in targeting GPCRs [9,14,15].

Allostery is one of the fundamental and ubiquitous mechanisms of functional regulation in proteins, where perturbations at one site (*e.g.* binding of a ligand, also known as an allosteric effector) modulate the dynamics and activity at a distant site in the protein [16–18]. A key attribute of the allosteric action is an indirect regulation of the function upon non-competitive ligand binding outside of the functional sites. Based on a rigorous analysis of a large number of allosteric and orthosteric sites and ligands [15], distinctive characteristics of the former have been determined, facilitating the identification of allosteric sites and design of highly specific allosteric effectors that provide long-range regulation in a targeted member of structurally-conserved families of proteins.

How can the "allosteric perspective" be beneficial in the development of GPCR-targeting drugs? The structural and dynamical properties of GPCRs enable a ligand-binding event in the extracellular part of the receptor to induce dynamic conformational changes at the distant cytoplasmic portion where the Gα subunit (termed here as GA) or β-arrestin (Arr) bind [16,20–22]. The signal transduction via allosteric signaling permits functional regulation of GPCRs in different modes and extents by diverse ligands, allowing for greater versatility than molecular events that require direct steric interactions, such as catalysis and its inhibition. Complementing allosteric signaling from the endogenous sites (*i.e.* orthosteric sites), the binding to other (allosteric) locations can remotely regulate the binding of GA and Arr, or modulate signals initiated by ligands bound to endogenous sites. Generally, the repertoire of allosteric effectors for GPCRs includes agonists and antagonists that regulate the function via allosteric signaling to GA and Arr sites, as well as positive and negative allosteric modulators (PAMs/NAMs) that tune/modulate effects of the agonists and antagonists. In addition to alleviating potential off-target toxicity by targeting non-endogenous and unique binding sites in specific receptors from conserved families/subtypes, prospective allosteric ligands and modulators offer fine-tuned regulation of target site(s), presenting a promising approach to tackle the complex challenges of personalized medicine [23–25]. Therefore, despite the limited number of FDA-approved allosteric drugs so far, the ubiquity of allosteric signaling in protein structures underscores its significant potential in modern drug design [18,21,26].

To systematically examine the complex GPCR functionality and its multi-faceted regulation, we use here our understanding of allostery and, particularly, the experience in developing and applying the Structure-Based Statistical Mechanical Model of Allostery (SBSMMA) – a physical model that accounts for the causality and energetics of allosteric communication in proteins at per-residue resolution [27,28]. The predictive power and design potential of SBSMMA have been demonstrated in a number of works, including prediction of allosteric sites and design of highly specific and tunable drug candidates that precisely modulate a protein of interest in a controllable manner [15,29,30]. Examples of case studies include the highly conserved matrix metalloproteinase

family [31], examining the role of disorder as a trigger of allosteric signaling [32], analysis of mutational effects from SARS-CoV-2 variants of concern [33], and exploring features of allosteric communication in protein folds and functional superfamilies [34], to name a few. SBSMMA has been implemented in the AlloSigMA server [35,36], and used to build the AlloMAPS database [37] recently complemented by the data obtained for predicted structures from AlphaFold database [38]. More recently, we proposed and demonstrated a generic SBSMMA-based *directed design protocol* for the rational design of allosteric drug candidates in proteins with desired allosteric response and affinity [23,30,31].

In this work, we first developed a database of allosteric communication in 280 non-olfactory human GPCRs from different families and classes (the complete data is available in the interactive AlloMAPS database at https://allomaps.bii.a-star.edu.sg/browse/gpcraf). Then, we performed a rigorous analysis of the conservatism and diversity of allosteric signaling in 23 chemokine receptors, most of which remain largely uncharacterized, showing how differences in the signaling may be utilized to regulate specific receptors within the family. Next, we focused on the extensively studied ADRB2 and GLP1R from classes A and B, respectively, quantifying the allosteric effects of various known ligands, including agonists, antagonists, PAMs, NAMs, and allosteric antagonists. Finally, based on the achieved understanding of intrinsic allosteric mechanisms in various GPCRs, we present here a generic framework for the rational design of allosteric drug candidates with desired efficacy and functional regulation. Key steps in the initial development of allosteric effectors include: (i) identification of latent allosteric sites, followed by (ii) a characterization of effector candidates in terms of the predicted binding free energy and allosteric response (both factors are instrumental for evaluating the drug efficacy) induced upon effector binding, and (iii) determining the specificity of a drug candidate's action to the target protein. Here, we show that allosteric effectors with controllable switching between agonistic and antagonistic modes, and simultaneous adjustments of corresponding effectors and binding sites (site-effector pairs) for precise regulation of a specific receptor, can be designed and performed in a systematic way. In the former, we demonstrate how to engineer switchable allosteric effectors with similar structures for GLP1R starting from known binding sites of distinct agonist and antagonist, as well as in scenarios where a priori information is not available. In the latter, an efficacy-driven approach was used for obtaining receptor structure-specific regulation, distinguishing the closely related glucagon receptors GLP1R and GIPR (gastric inhibitory polypeptide receptor).

Results

Conservatism and diversity of allosteric signaling in the chemokine receptors family

The first objective of this work is a rigorous analysis of allosteric communication in GPCRs, aiming to detect variations of signaling in related receptors at a per-residue level and characterize the archetypal signaling patterns of GPCR classes and families. To this end, we used two quantitative representations obtained from exhaustive perturbations at every residue position – Allosteric Signaling Maps (ASMs [28]) and Allosteric Probing Maps (APMs [36]), both providing a picture of allosteric signaling based on the SBSMMA formalism. The high-throughput ASMs and APMs are computed by quantifying the energetics associated with allosteric signals emanated to every residue position i upon a perturbation P (*e.g.* a mutation) in a structure, termed as the allosteric modulation ($\Delta h_i^{(P)}$, kcal/mol). The ASM exhaustively quantifies the allosteric signaling from every residue position in the protein and APM – from the sequential binding of a three-residue probe to every consecutive three-residue segment of the protein (see Materials and Methods and [28,38]). Positive allosteric modulation values reflect increased dynamics of the affected residue that might lead to potential conformational changes, whereas negative modulation values indicate overstabilization of the residue as a result of allosteric response.

We obtained the comprehensive ASMs and APMs describing allosteric signaling in 280 GPCRs belonging to classes A, B1, B2, C, F, and T (the complete data is presented in the AlloMAPS database). Focusing here on the chemokine receptors family, Figure 1A illustrates a structure of CCR1 and Figure 1B shows ASMs for CCR1 and CCR5 (Figure 1B). The ASMs reveal common patterns characterized by pronounced signaling (large positive Δh) between extracellular parts of transmembrane helices (TM) and intracellular loops, indicating strong communication between these distant regions. For instance, a perturbation at the ECL2 (extracellular loop 2) linking TM4 and TM5 (Figure 1A) originates a large positive modulation at multiple intracellular locations in CCRs 1 and 5, including ICL1 (intracellular loop 1) between TM1 and TM2 and ICL3 between TM5 and TM6 (Figure 1B). The dynamics and/or structure of ICL2 and TM4 are strongly modulated by various distant positions in CCR5, while allosteric signals emanated to the same locations in CCR1 show a picture of weaker modulation values. Despite some variations in long-range signaling among chemokine receptors, the archetypal signaling map of the family (Figure 1C), obtained by averaging the ASMs computed from the structures of 23 chemokine receptors (see Table S1 in Supplementary File 1 for information on the receptors), shows consistent patterns of

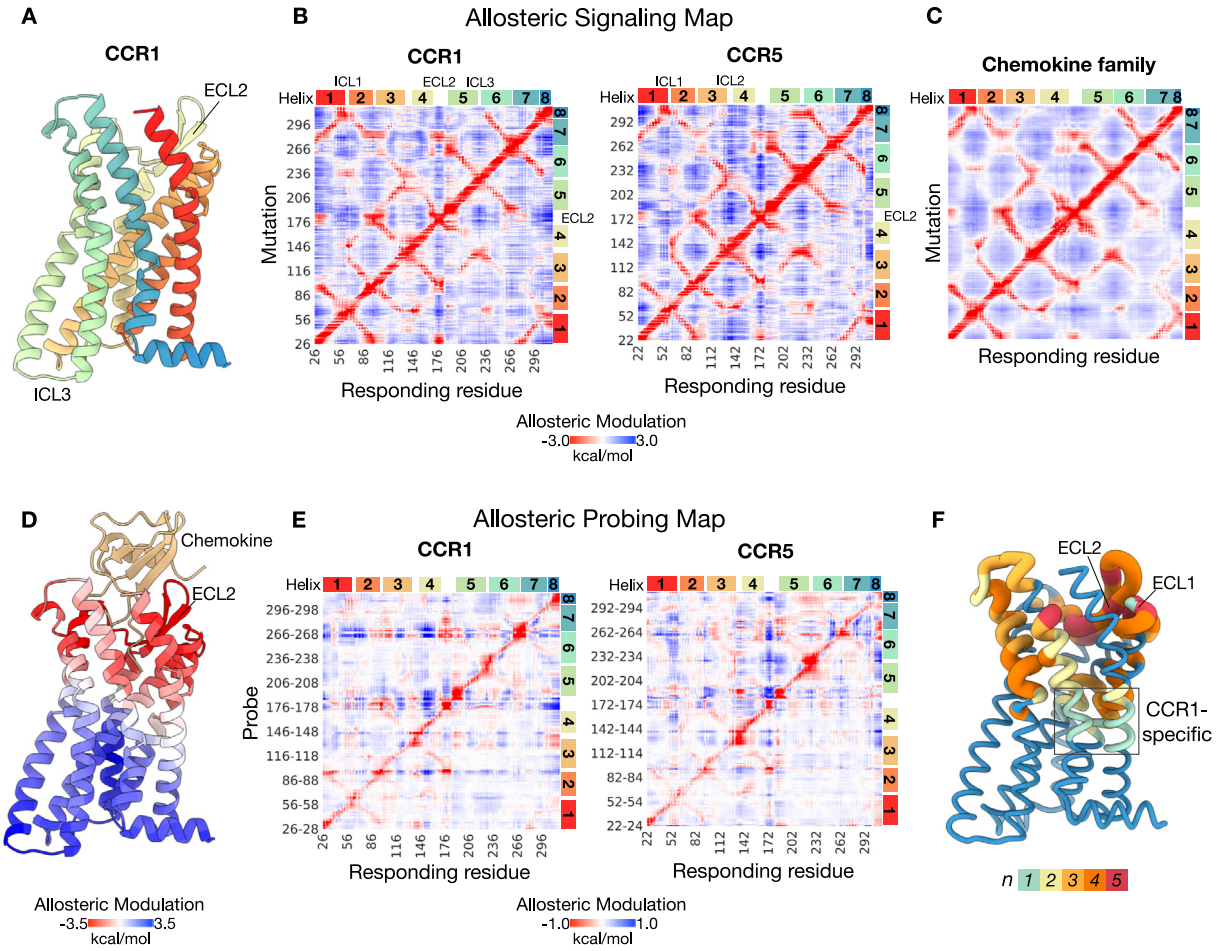

Figure 1. Conservation and diversity of allosteric communication in class A chemokine receptors. A. A CCR1 structure showing the archetypal seven-transmembrane helical architecture (TM1–TM7) connected by three extracellular (ECL1–ECL3) and three intracellular loops (ICL1–ICL3), colored in red-blue spectrum. **B.** ASMs of representative CCR1 and CCR5, illustrating variations and conservatism in allosteric signaling across different members of the chemokine family. **C.** The averaged ASM of class A chemokine receptors. **D.** Allosteric modulation caused by binding of a CCR3 ligand (gold) to CCR1 receptor quantified using SBSMMA. **E.** APMs of representative CCR1 and CCR5. **F.** Conservation and specificity of allosteric signaling among closely related receptors (CCR1, CCR2, CCR5, CX3CR1, and CXCR2). The distribution of the top 20% of allosteric probe residues affecting binding of Gα is color-coded by the number of times a probe is identified to strongly modulate the GA sites in respective receptors: red (observed in all five receptors), orange (four), light orange (three), yellow (two), and light cyan (CCR1-specific). A structure of CCR1 is used here for reference.

strong signaling between extracellular loops, which interact with chemokines, and intracellular locations including ICL1, 2 and 3 where the Gα subunit (GA) and β-arrestin (Arr) bind.

We next explored how the allosteric effect of chemokine binding (Figure 1D shows CCR1 as an example) can be analyzed using the APMs at per-residue resolution. The APMs of CCRs 1 and 5 (Figure 1E) show that probes sending strong positive signals are typically clustered at ECL2 and the extracellular end of TM4, a region coincides with the interaction interface with chemokines [39]. To further investigate this pattern, we analyzed the APMs of several closely related chemokine receptors, including CCR1, CCR2, CCR5, CX3CR1 and CXCR2. Specifically, we identified residue triplets that originate the largest positive modulation (top 20%) to the GA binding sites of the five receptors upon probe binding. We examined the distribution of strongly modulating probed triplets at corresponding positions in the structures based on a structure-based sequence alignment (provided in Supplementary File 2), delineating those that are unique to individual receptors and those inducing strong allosteric response in multiple receptors (Figure 1F, for clarity, a structure of CCR1 is used as a reference). For example, allosteric probing hotspots found in all five receptors (colored in red) are mostly located in ECL1 and ECL2, whereas those specific to CCR1 are found in the transmembrane portion of the adjacent TM1 and TM7 (see Figure 1 legend for details).

Overall, the result shows an existence of conserved signaling patterns in the chemokine receptors family and allosteric hotspots for screening or design of effectors (APM) and protein engineering (ASM). At the same time, each receptor exhibits distinctions in signaling signatures, underscoring the potential for designing structure-specific allosteric effectors. As a result of this study, we developed a database containing exhaustive signaling and probing maps obtained for a large set of 280 GPCRs in human studied here (https://allomaps.bii.a-star.edu.sg/browse/gpcraf). These comprehensive data could be utilized as starting points for exploring allostery in GPCRs, detecting latent allosteric sites, and designing allosteric effectors with desired specificity and efficacy.

Allosteric effects of ligands in class A and class B1 receptors

Among experimentally characterized classes A and B1 GPCRs, the human β_2-adrenergic receptor (ADRB2) and glucagon-like peptide-1 receptor (GLP1R) are notable due to extensive structural data. To date, at least 48 crystal structures of ADRB2 and 46 structures of GLP1R have been deposited in the RCSB PDB database, revealing a picture of diverse interactions with a range of ligands exhibiting distinct functional properties. The functional properties of these ligands encompass a broad spectrum [40] including inverse agonists, agonists, partial agonists, antagonists, as well as various types of modulators including positive and negative allosteric modulators (PAMs and NAMs), and allosteric antagonists. The complex interplay between agonists/antagonists and different modulators further complicates the understanding of GPCR regulation. Moreover, ligands binding to partially overlapping sites can induce distinct allosteric responses at distant intracellular locations and, thus, varied functional outcomes. As a result, a subtle change in the molecular structure and intermolecular interactions can lead to a switch between agonistic and antagonistic modes. In the analysis performed here, first, the binding sites of diverse ligands in ADRB2 and GLP1R were identified from the crystal structures. The ligands and binding sites are named according to the ligand codes from PDB database (see Supplementary File 1, Table S2 for complete information). Then, we explored and quantified the per-residue allosteric modulation $\Delta h_i^{(lig)}$ emanated by these ligands in corresponding receptors using SBSMMA. For ADRB2, we considered the binding sites of agonist ERC, antagonists JRZ, CAU and TIM, a NAM called M3J, and an allosteric antagonist 8VS. For GLP1R, the endogenous agonist glucagon-like peptide 1 (GLP-1) and synthetic agonist UK4, and PE5, QW7 and 97Y, which represent antagonist, PAM, and NAM, respectively, were investigated.

In ADRB2, binding of the agonist ERC stabilizes (negative modulation values) proximal residues, such as those surrounding the ECL1 (residues 90–106), to a greater extent than that by the bound antagonists TIM and JRZ (Figure 2, top row). At the same time, ERC agonist induces large positive $\Delta h_{i\in GA}^{(ERC)}$ values at residues in the distant GA site (shaded regions, Figure 2, center), suggesting conformational changes particularly at residues 131–146 spanning the cytoplasmic ends of helix 3 and ICL2. Despite binding to similar locations as ERC, the antagonists TIM and JRZ cause the opposite effect – overstabilization – in ICL2 and TM3 residues forming the GA site, suggesting their roles in precluding structural and dynamical changes required for interacting with the G protein. On the other hand, we observed negative $\Delta h_{i\in GA}^{(8VS)}$ values caused by the binding of allosteric antagonist 8VS at the intracellular surface of the receptor, including adjacent ICL2 residues in GA site. Noteworthily, 8VS exerts positive modulation values at residues surrounding the distant orthosteric site (Figure 2). In addition to its role in sterically blocking interactions with Gα, 8VS allosterically disrupts the structure of the orthosteric site, consistent with experimental findings on its effects in reducing the affinity for agonists [41]. We also found that negative modulator M3J alone is unable to elicit allosteric response, it is likely that the modulator works together with other ligands to block ADRB2 function.

The GLP1R agonists, UK4 and GLP-1, originate a large positive allosteric modulation at the intracellular parts of TMs 5–7, as well as ICL2 and ICL3 forming the GA site, indicating conformational changes that might facilitate binding of the G protein (Figure 2, bottom row). As crystal structures of GLP1R with bound antagonists are currently unavailable, we mapped the binding residues of the antagonist PE5 from the closely related GLP receptor onto GLP1R using structure-based sequence alignment (Supplementary File 1, Table S2; Supplementary File 2). Simulations with SBSMMA confirmed the antagonist effects, showing the absence of allosteric responses in the structure, consistent with its role in preventing agonist binding and receptor activation. Similar to the observation on M3J in ADRB2, binding of modulators QW7 (PAM) and, separately, 97Y (NAM) were largely unable to elicit allosteric responses in distant regions of GLP1R by themselves (Figure 2). Overall, one can conclude that there is a common trend in both receptors, in which binding of agonists destabilizes the GA site, likely facilitates conformational changes associated with receptor activation, while antagonists and modulators typically lead to very weak allosteric responses. It is also important to note that the lack of significant allosteric response from binding of a modulator alone suggests that they might work mostly in presence of other ligands, prompting us to investigate the combinatorial effects of GPCR ligands.

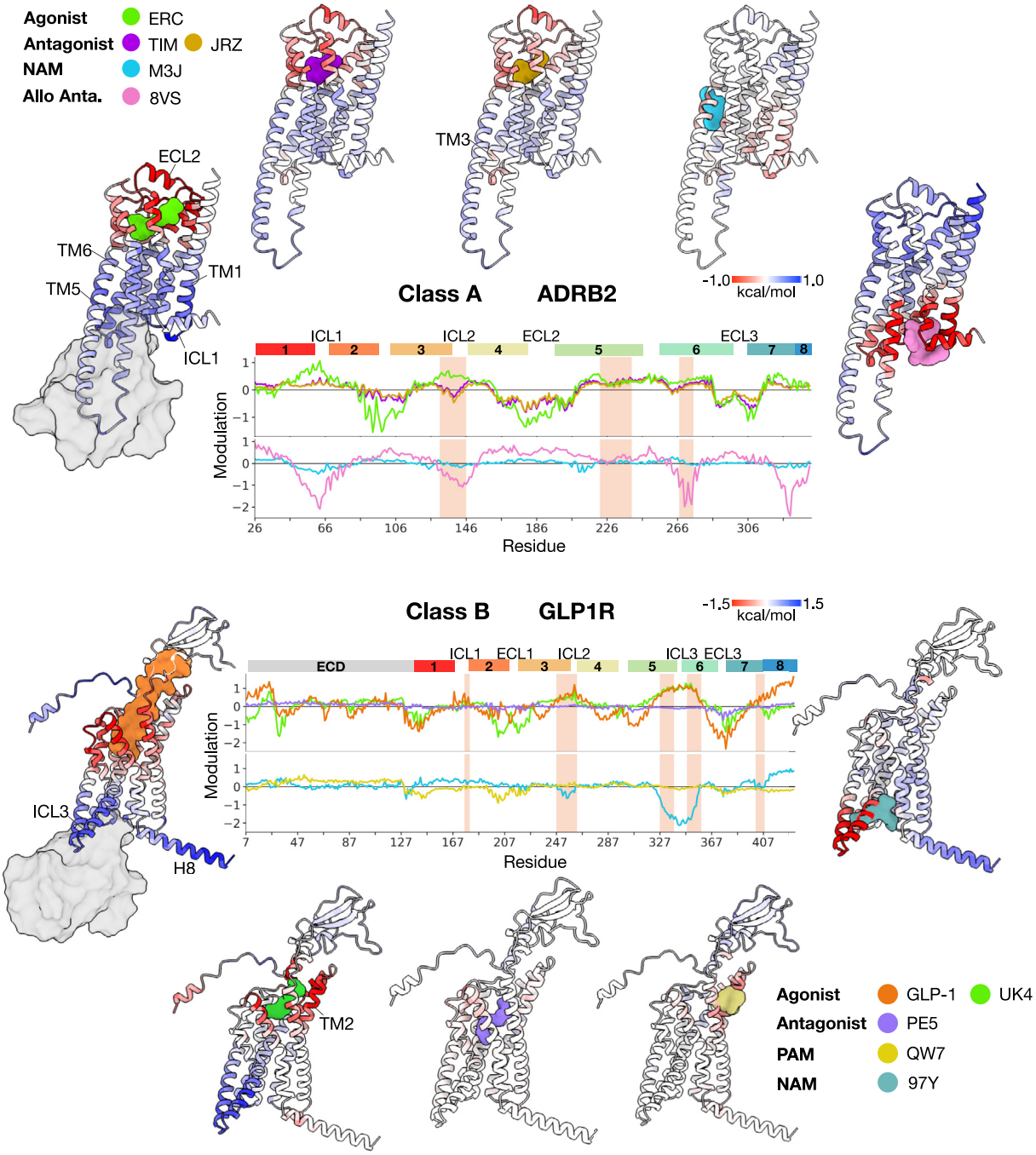

Figure 2. Allosteric modulation induced by various ligands in the class A ADRB2 receptor and class B1 GLP1R receptor. The per-residue modulation values (kcal/mol) exerted at all residues upon binding to various ligands are shown. Residues forming the GA site are indicated by orange shaded areas. For ADRB2, we obtained the binding sites using the following crystal structures: ERC (3PDS), JRZ (3NY8), CAU (5X7D), TIM (6PS6), M3J (6OBA) and 8VS (5X7D). For GLP1R: GLP-1 (7DUQ), UK4 (7LCI), PE5 (mapped from GLP receptor, 5EE7), QW7 (6VCB) and 97Y (5VEX).

Complex interplay of allosteric communication between sites in ADRB2 and GLP1R

While the above analysis illustrates signaling emanated from binding of individual ligands to every residue in a structure, the complex picture of allosteric communication between various locations calls for a straightforward quantitative representation of communication between protein sites. Here, we quantified allosteric signaling exerted to a site by averaging the per-residue modulation values of all residues *i* forming a site of interest $\Delta h_{site}^{(lig)} = \langle \Delta h_i^{(lig)} \rangle_{site}$ upon distant binding of a ligand, which enables us to compare the effects of allosteric modulation at different sites. Next, we explored combinatorial effects of multiple

ligands in originating a signal to functional sites $\Delta h_{site}^{(lig1,\ lig2,\ ...)}$, allowing us to study regulatory mechanisms of known modulators, as well as predicting functional regulation by combinations of ligands which have yet to be characterized in experiments. Figure 3 depicts a schematic presentation of communication between selected sites (center) and its adjustments by modulators (outer circle), focusing on the signaling to GA and ICL3 in ADRB2, and GA and Arr sites in GLP1R. The complete communication data is available in Supplementary File 1, Tables S3 and S4.

In ADRB2, we observed differential modulation of GA site by the agonist ERC and antagonists TIM, JRZ and CAU, all of which interact with the class A receptor at partially overlapping sites (Figure 3, left). The agonist causes a positive per-residue modulation $\Delta h_{GA}^{(ERC)}$ = 0.4 kcal/mol, whereas the antagonists lead to much weaker response at 0.2, 0.1 and 0.1 kcal/mol, respectively, suggesting their roles in sterically blocking competing agonist binders. Interestingly, perturbing the intracellular GA site elicits a strong positive modulation to the extracellular orthosteric site where agonists and antagonists bind, for example, $\Delta h_{ERC}^{(GA)}$ = 1.2 kcal/mol. We have previously proposed and validated that reversing allosteric communication by modeling the binding of a substrate in a functional site enables one to predict latent allosteric sites – a technique coined by us as reverse perturbation [42]. The positive modulation exerted on the orthosteric site upon perturbing the binding interface of Gα serves here as a clear demonstration of the predictive power of the reverse perturbation method. While the positive modulator KBY alone results in a negligible modulation at ICL3 (Supplementary File 1, Table S3), it enhances the magnitude of signaling from the agonist ERC when both are concurrently bound ($\Delta h_{ICL3}^{(ERC)}$ = 0.4 and $\Delta h_{ICL3}^{(ERC,KBY)}$ = 0.8 kcal/mol; Figure 3, top left). The combination of KBY and

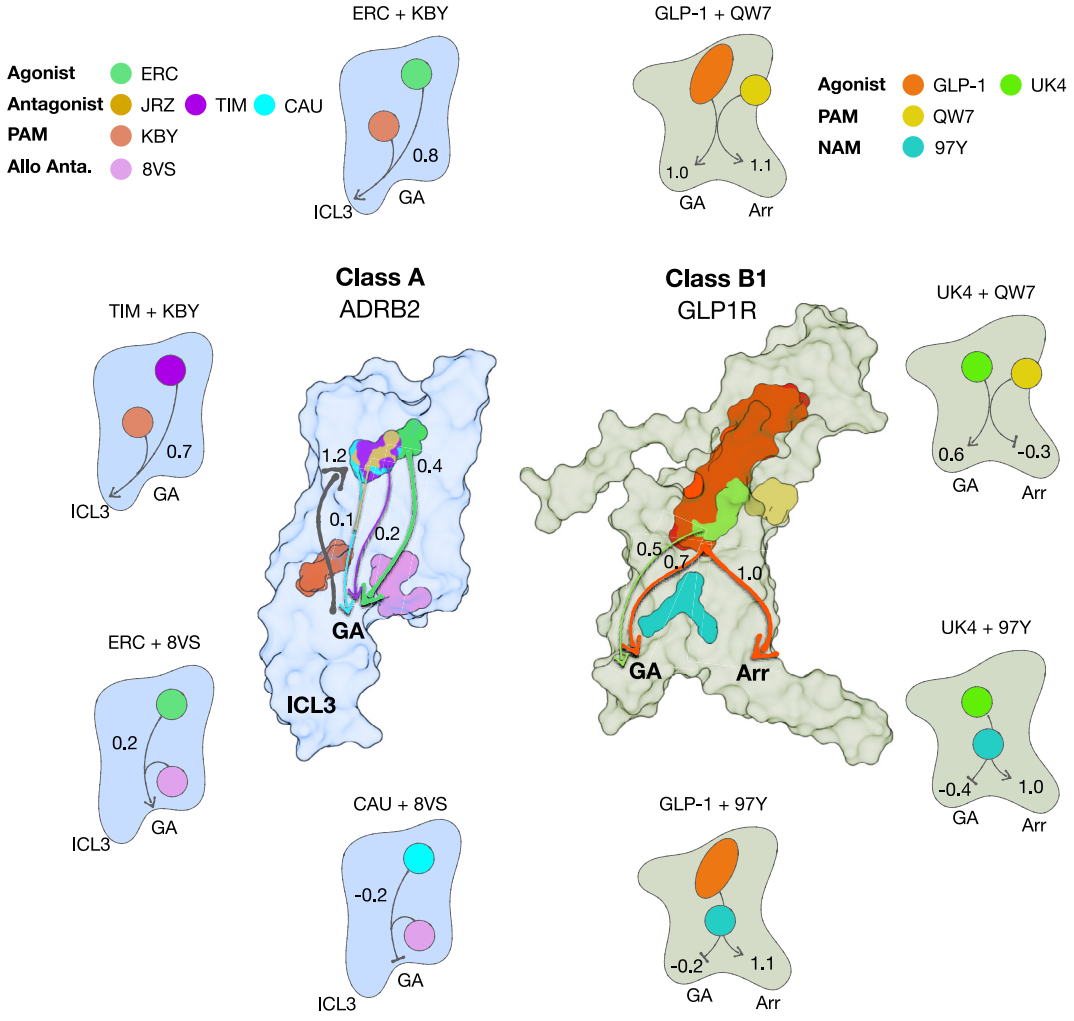

Figure 3. Allosteric communication between functionally important sites and its alteration by modulators. The figure illustrates the mode and strength of allosteric signaling Δh_{site} between ligand binding and functional sites including the GA and Arr sites. The outer parts show examples of combined effects of agonists/antagonists and PAMs/NAMs. Complete data can be found in Tables S3 and S4 in Supplementary File 1.

antagonist TIM leads to larger positive modulation compared to KBY alone ($\Delta h_{ICL3}^{(TIM)}$ = 0.2 and $\Delta h_{ICL3}^{(TIM,KBY)}$ = 0.7 kcal/mol). The large positive modulation values indicate conformational changes that may be required for ICL3-Gα interaction, and in turn, activates the Gα-dependent signaling pathway. Besides potentially reducing the affinity for agonists described above (Figure 2), the allosteric antagonist 8VS could attenuate the signaling to GA site from bound ERC at the orthosteric site ($\Delta h_{GA}^{(ERC)}$ = 0.4 and $\Delta h_{GA}^{(ERC,8VS)}$ = 0.2 kcal/mol). In the case of antagonist CAU, the mode of modulation exerted at GA switches from weakly positive at 0.1 kcal/mol with CAU alone, to −0.2 kcal/mol when both CAU and 8VS are bound.

In GLP1R, we found that binding of the endogenous peptide GLP-1 causes a stronger per-residue modulation at GA ($\Delta h_{GA}^{(GLP1)}$ = 0.7 kcal/mol, Figure 3, right) compared to that by the synthetic agonist UK4 (0.5 kcal/mol). While GLP-1 also sends a signal to Arr at 1.0 kcal/mol, UK4 does not modulate the site (Supplementary File 1, Table S4). Consistent with experiment findings showing positive modulator QW7 increased the efficacy of GLP-1 [43], binding of QW7 further enhances GLP-1 signaling to GA ($\Delta h_{GA}^{(GLP1,QW7)}$ = 1.0 kcal/mol). The stronger modulation is also observed for the UK4 and QW7 pair ($\Delta h_{GA}^{(UK4,QW7)}$ = 0.6 kcal/mol). Moreover, UK4 agonist in the presence of QW7 induces biased allosteric signaling to the distinct Gα and β-arrestin binding sites, likely causes conformational changes in the former but stabilizes the latter site ($\Delta h_{GA}^{(UK4,QW7)}$ = 0.6 and $\Delta h_{Arr}^{(UK4,QW7)}$ = −0.3 kcal/mol). Given that β-arrestin plays a critical role in receptor desensitization and internalization, as well as alternative Gα-independent pathways, this signaling bias might result in selective modulation of cellular signal transduction pathways. On the other hand, simultaneous binding of the negative allosteric modulator 97Y with either GLP-1 and UK4 (both agonists), stabilizes GA but leads to increased dynamics in Arr. The biased modulation by the agonists, as a result of 97Y modulator, may lead to differential regulation of Gα and β-arrestin mediated pathways (Figure 3, bottom right).

From learning the mechanisms of allosteric communication to designing switchable and structure-specific allosteric drug candidates

The analysis above shows divergent allosteric response caused by known ligands that bind to partially overlapping sites. It is also known that slight changes of an agonist's structure and interactions with a receptor can often perturb the protein dynamics in alternative ways, inducing different allosteric response in the Gα interface across the transmembrane domain [19]. Based on the capability of SBSMMA to account for the causality of allosteric communication caused by known agonists, antagonists and modulators (Figures 2 and 3), we introduce here a computational framework for prediction of latent allosteric sites and rational design of new allosteric effectors/ligands with therapeutic potential and desired efficacy. The GLP1R is used here as a model system, and three important aspects in designing prospective allosteric drugs are demonstrated here. First, in cases where allosteric sites and ligands for a target receptor are not known, delineating intrinsic allosteric communication in the target enables identification of latent sites and new effector ligands. We demonstrate below two case studies: in the first one, the locations of characterized allosteric ligands provided a basis for obtaining new ones (Figure 4); in another, the Allosteric Probing Map was used to explore latent allosteric sites in the structure. Second, we utilized a phenomenon known as mode switching, in which compounds with similar structures can exhibit varied regulatory effects as agonists, antagonists, PAMs or NAMs [8] upon slight modifications of their structures and binding locations (Figure 5). We show that agonists and antagonists interacting with different residues can be, indeed, derived from the same fragment-like compounds for GLP1R. Third, the proposed framework allows to identify unique allosteric sites and develop effectors that specifically modulate a functional site of a target receptor (GLP1R, Figure 6), while sparing a closely related one (GIPR).

Obtaining new switchable allosteric effectors using information on known ligands. The binding locations of the potent agonist UK4, WB2 with weaker agonist activity, and the antagonist PE5 form a basis for designing a series of effectors with mode switching capability starting from fragment-like compounds (with molecular weight below 300 Da) from DSI-poised fragment library [44]. Briefly, major steps of the procedure (explained in detail with the flowcharts in Figure S1, Supplementary File 2) include: (i) exhaustive mapping of fragment binding sites in different locations to yield a repertoire of corresponding binding site-fragment pairs; (ii) evaluating the efficacy of each site-fragment pair in terms of allosteric modulation at GA site and estimated binding free energy; (iii) iterative optimization of the efficacies of shortlisted site-fragment pairs towards those of strong agonists or antagonists, by modifications of the molecular structures and binding sites using the *directed design protocol* proposed previously [30].

First, we explored the binding locations of UK4 and WB2 (strong and weak agonists) and antagonist PE5, by performing virtual screening of 860 diverse fragments from DSI-poised fragment library [44]. Within each location, we obtained multiple binding poses of fragments with their binding residues (defined as residues within 4 Å from any ligand atom), which we referred to as site-fragment pairs. Each corresponding pair was characterized by the estimated binding free energy

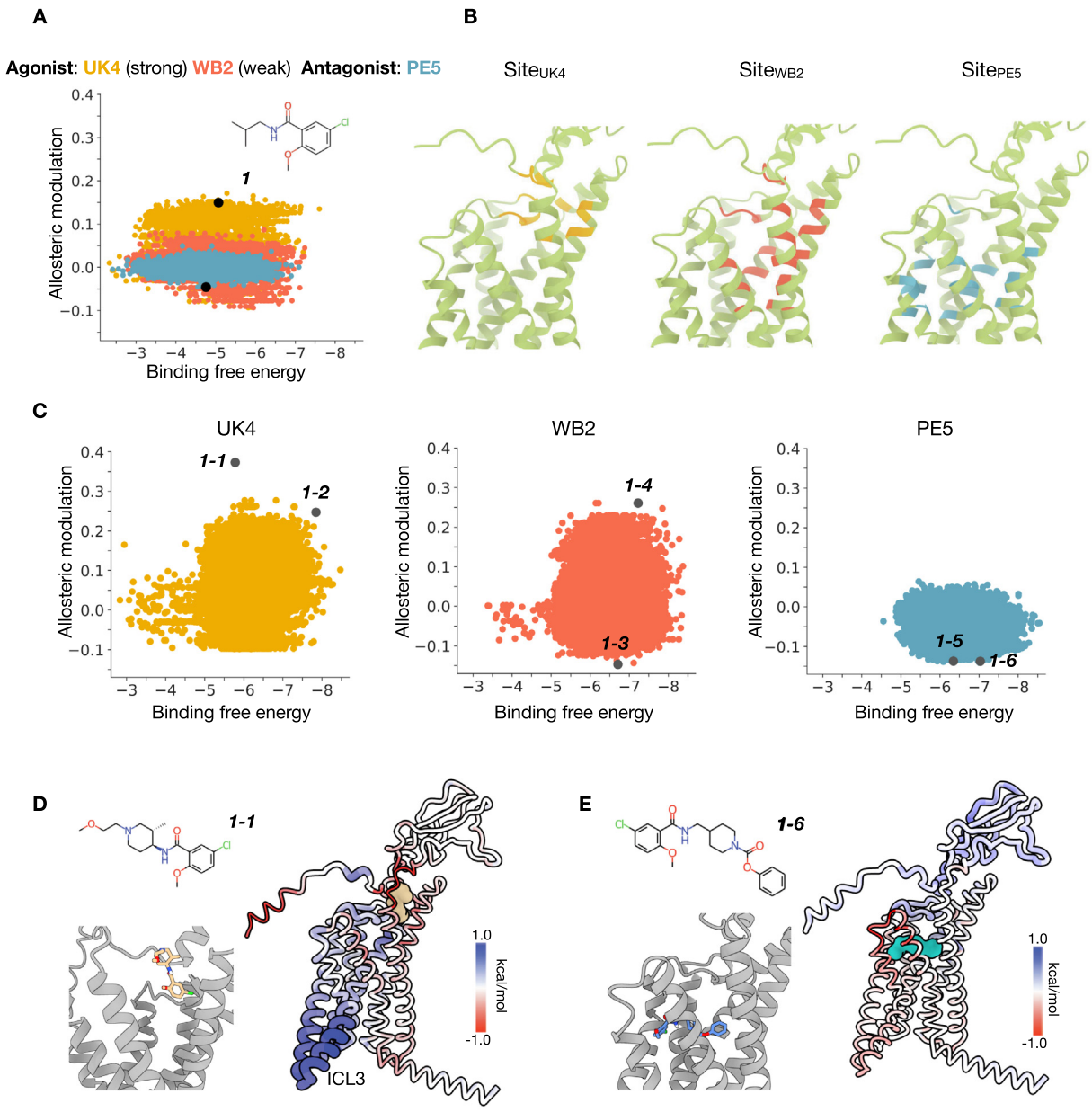

Figure 4. Design of switchable allosteric effectors based on known ligand binding sites in GLP1R. A. Analysis of fragment-like compounds towards locations corresponding to the strong agonist UK4, weaker agonist WB2, and antagonist PE5, respectively. An example of fragments with mode switching potential, fragment ***1*** is highlighted. **B.** Residues forming Site$_{UK4}$, Site$_{PE5}$ and Site$_{WB2}$ for subsequent docking of derived compounds. **C.** Analysis of selected derivatives binding to Site$_{UK4}$, Site$_{PE5}$ and Site$_{WB2}$. **D** and **E.** Binding poses and modulation caused by ***1-1*** and ***1-6***.

Δg_{bind} from AutoDock Vina and allosteric modulation Δh_{GA} evaluated by SBSMMA (both in kcal/mol), allowing one to predict the efficacies of various fragments. The result shows varied distributions of Δg_{bind} and Δh_{GA} values of site-fragment pairs within partially overlapping locations correspond to UK4, WB2 and PE5 (Figure 4A). In particular, fragments binding to sites proximal to that of strong agonist UK4 generally induce larger positive modulations compared to the same set of fragments in WB2 (weak agonist) and PE5 (antagonist) locations. In most of the cases, fragments binding to the PE5 location result in very weak signals to GA site. At the same time, we found that fragment *1* induces divergent allosteric response from the UK4 and PE5 locations (Figure 4A). Therefore, it was subsequently selected as a "seed" for deriving switchable effector ligands exhibiting distinct regulatory effects.

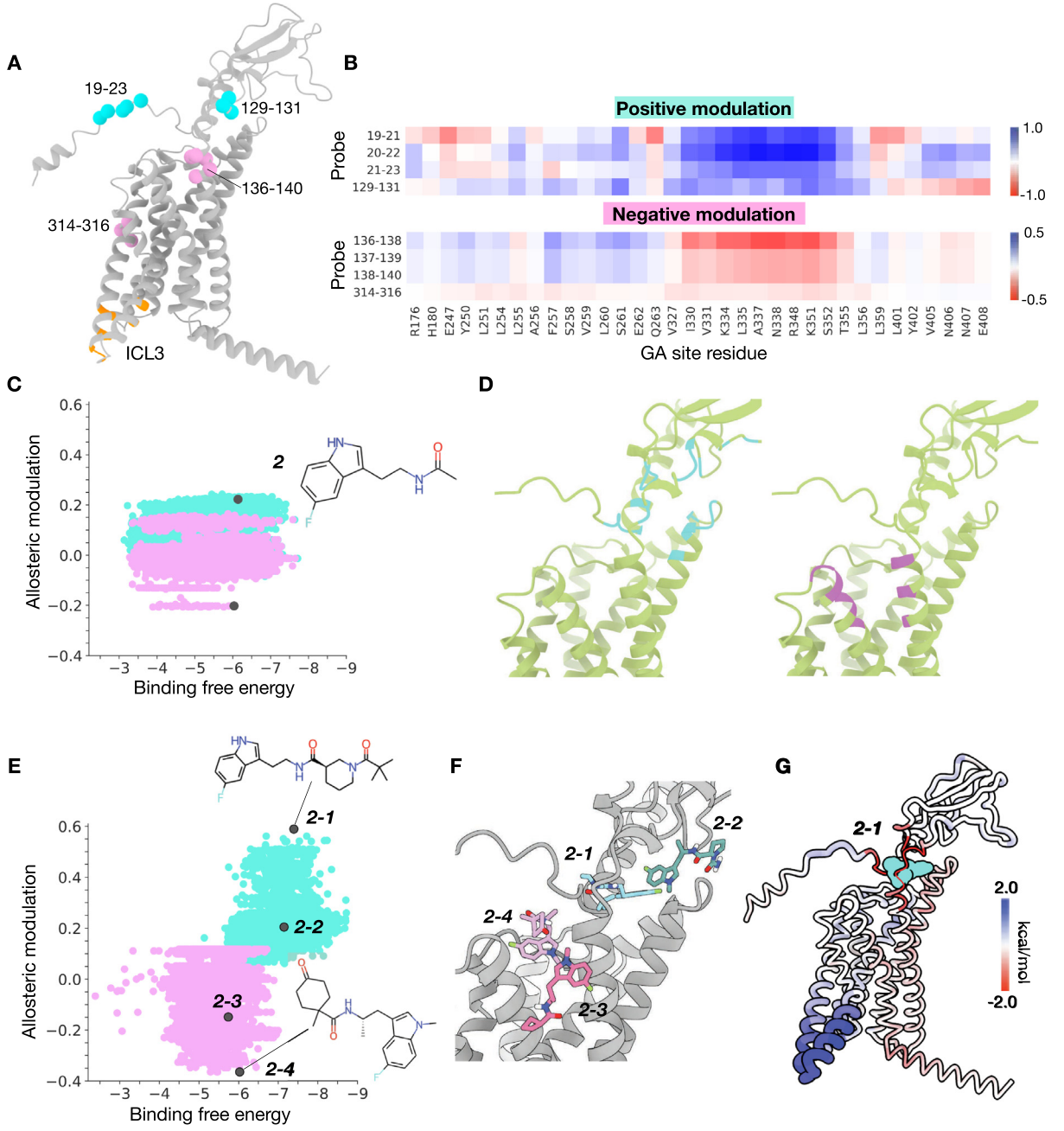

Figure 5. ***De novo*** **design of switchable allosteric effectors for GLP1R. A.** Selected allosteric probes inducing strong positive or negative modulation at the GA site. **B.** Modulation effects of strongly modulating probes on individual GA residues. **C.** Analysis of fragments targeting locations based on strongly modulating probes, cyan – residues 129–131 and magenta – residues 136–140. An example of fragments with mode switching potential, fragment ***2*** is highlighted. **D.** Latent allosteric site residues for screening of derivatives**. E**. Analysis of selected derivatives binding to the latent sites. Compounds ***2-1*** to ***2-4*** with distinct efficacies are indicated. **F.** Binding poses of derivatives. **G.** Allosteric modulation of compound ***2-1***, a potent agonist.

The exhaustive fragment mapping revealed key residues capable of causing large positive modulation values in GA when perturbed, likely facilitating conformational changes required for G protein binding, as well as those precluding the changes via stabilization of GA site (negative modulation). Figure 4B shows locations denoted as $Site_{UK4}$ (left) and $Site_{PE5}$ (right), which are composed of residues involved in 50 site-fragment pairs obtained in the UK4 location with the fragment poses providing

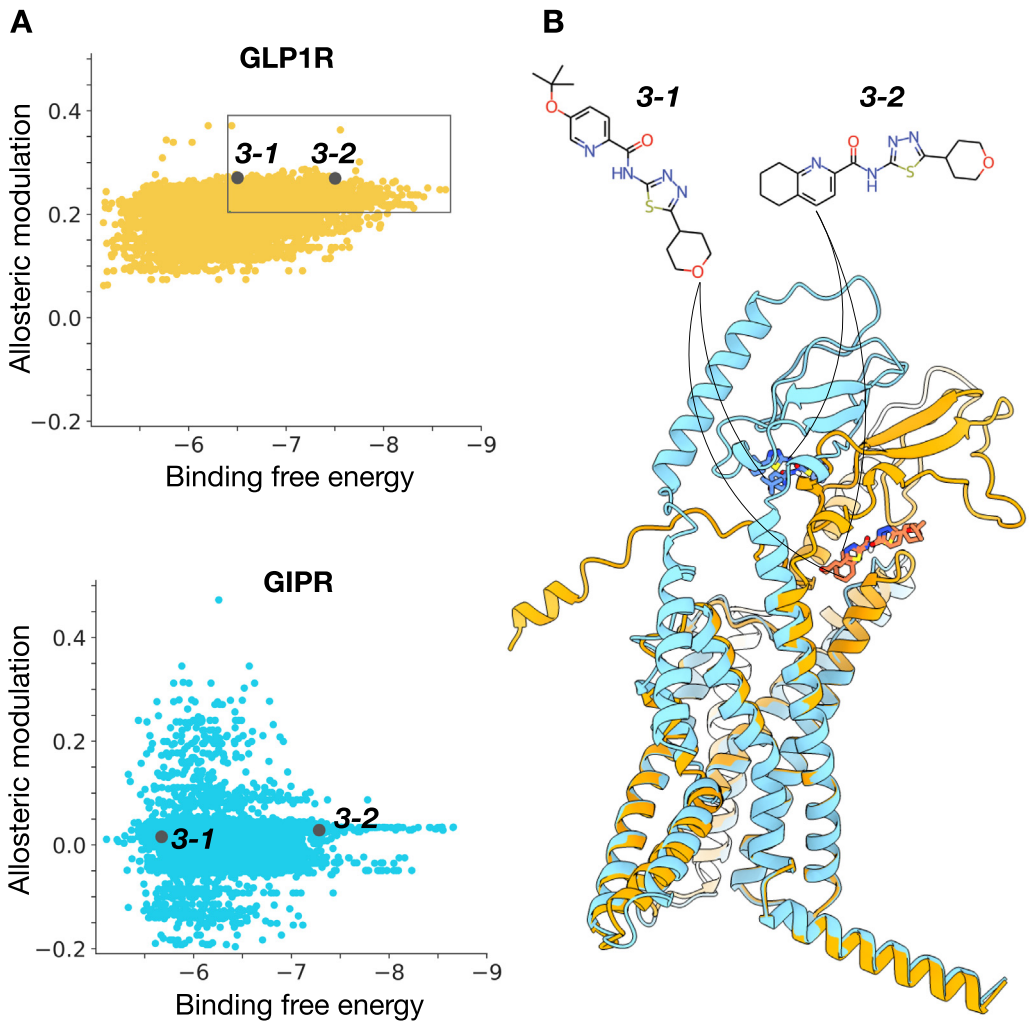

Figure 6. Efficacy-based design of effectors with high specificity in GLP1R. A. Binding free energies and modulation values of derived compounds upon interactions with GLP1R and GIPR. For the counter-screen on GIPR (below), promising compounds from GLP1R (*above,* highlighted in rectangle), including ***3-1*** and ***3-2***, were used. **B.** Binding of ***3-1*** and ***3-2*** on aligned structures of GLP1R and GIPR.

the highest positive Δh_{GA} values, and separately, 50 site-fragment pairs in the PE5 location with the fragment poses providing the largest negative modulation. $Site_{WB2}$ contains residues involved in all site-fragment pairs, regardless of the modulation values (middle).

Subsequently, the *directed design* approach was used to identify prospective agonists and antagonists using fragment *1* as a seed and focusing on $Site_{UK4}$, $Site_{WB2}$ and $Site_{PE5}$ (see Figure S1, Supplementary File 1 for detail). Derivatives of *1* were obtained from ZINC database [45] via substructure-based matching, yielding 8,609 compounds after discarding those that are likely promiscuous and/or with poor pharmacokinetics using PAINS [46] and Brenk filters [47], and prioritizing those with a potential to be developed as oral drugs – estimated by the Quantitative Estimate of Drug-likeness, QED [48]. Virtual screening of the derived ligands followed by a quantification of modulation using SBSMMA (Figure 4C) shows that it is possible to design and fine-tune diverse allosteric site-ligand pairs with different modes and strengths of the efficacy. For instance, putative agonists (*e.g. 1-1*, *1-2* and *1-4*) and antagonists (*1-3*, *1-5* and *1-6*) were obtained from a common fragment *1*, demonstrating the design of allosteric effectors with controllable mode switching based on a single precursor. The agonist *1-1* binds to an allosteric site on GLP1R and causes positive modulation at GA site, especially at the intracellular loop 3 (Figure 4D), which is comparable to that by the strong agonist UK4 ($\Delta h_{GA}^{(UK4)}$ = 0.5 kcal/mol, Figures 2 and 3). As opposed to the negligible modulation by the antagonist PE5 (Figure 2), derivatives *1-3*, *1-5* and *1-6* bind to partially overlapping sites and

allosterically stabilize the dynamics of ICL3, likely working as antagonists (*e.g. 1-6*, Figure 4E). The SMILES strings of the compounds are available in Supplementary File 1, Table S5.

Prediction of latent allosteric sites and de novo design of switchable effectors. A key advantage of the generic framework proposed here is the capability to discover allosteric ligands for GPCRs with unknown endogenous ligands and/or physiological function, also coined as orphan GPCRs [49]. Moreover, rational design of allosteric ligands for activating or inhibiting the receptor function and biological pathways can shed light on the GPCR function when coupled with experimental assays. In this section, the APM containing comprehensive information on signaling was used as a starting point in finding allosteric binding sites and *de novo* design of effectors targeting GLP1R. We used here our targeted analysis method [29], which selects allosteric probes from the APM providing the highest modulation on a targeted site – in this case the Δh_{GA} modulation on the Gα binding interface. The targeted analysis allowed us to determine residues important for originating allosteric signaling to GA. Importantly, residues involved in binding of known ligands were excluded from the targeted analysis, allowing us to explore new locations and structures of allosteric sites. Figures 5A and B show residues identified by the method and the per-residue modulation values exerted at GA site. Notably, allosteric signals originated by strongly modulating probes from alternative locations converge at residues that are part of ICL3 (orange, from I330 to T335). Virtual screening of DSI-poised fragment library was performed based on locations centered at residues 129–131 in the ECD (positive Δh_{GA}) and 136–140 in TM1 (negative Δh_{GA}). Fragment *2* was found to elicit divergent modes of modulation upon binding to these locations (Figure 5C), which was subsequently used for deriving larger compounds using the directed design approach described above. Residues bound by fragment poses that provide the largest positive (top 150, cyan) and negative (top 150, magenta) modulations (Supplementary File 3) were used to define the locations (Figure 5D) for screening of 2,634 compounds derived from fragment *2*. Figures 5E and F illustrate that effector ligands sharing a common sub-structure can lead to varied allosteric response and regulatory effects, *i.e.* mode switching. In particular, binding of putative agonist *2-1* to the hinge region between ECD and TMD induces a larger modulation than that by UK4 (Figures 5F and G). Switching to the antagonistic mode, the structurally similar derivative *2-4* results in an opposite allosteric effect. While *2-1* and *2-4* interact with several residues that are involved in binding of the endogenous peptide GLP-1, we also identified various putative agonists and antagonists binding to allosteric sites that are non-overlapping with the binding sites of known ligands, such as derivatives *2-2* and *2-3*. To summarize, we show here a generic framework for *de novo* design of allosteric drug candidates. It includes the whole procedure from finding the allosteric sites for targeting a functional site, to developing an array of allosteric ligands capable of regulating the receptor function in varied ways. Specifically, we started from the exhaustive probing perturbation of GLP1R, then, performing targeted analysis based on the APM, we identified residues of the receptor that are allosterically coupled to GA site and not involved in the binding of known ligands, allowing us to obtain new allosteric site-ligand pairs. The SMILES strings of the putative effector ligands can be found in Supplementary File 1, Table S5.

Tuning the specificity of agonist–antagonist pairs for targeted allosteric sites. Importantly, the computational framework used here allows an optimization of the structures of derived ligands and the composition of binding sites in an iterative process in order to achieve a desired efficacy. In other words, depending on specifics and requirements of the task, the candidate ligands exemplified above could undergo further iterations of modification within the framework towards certain characteristics. For example, some of the candidate ligands used as representatives for switchable effectors in the previous sections, such as putative agonist *1-1* and antagonists *1-6, 2-3* and *2-4*, exhibit slightly weaker predicted binding free energies for the designated agonistic and antagonistic locations, respectively, compared to other locations (additional data in Figures S2 and S3, Supplementary File 1), while agonists *2-1* and *2-2* bind preferentially to the designated locations. Therefore, the structures of *1-1*, *1-6*, *2-3* and *2-4* may be further modified using the directed design protocol – substructure-based search followed by an evaluation of the binding free energies and resulted allosteric modulation – to achieve a fine-tuning of specificity for the targeted location. Alternatively, other candidates derived from fragments *1* and *2* may also be used, such as the putative agonist co_2616 and antagonist co_4391 (both obtained from fragment *1*), and co_746 (from fragment *2*) which show stronger estimated binding free energies to the designated locations compared to "off-target" locations (see details in Figures S2, S3 and Table S5 in Supplementary File 1). Overall, the specificity of allosteric drug candidates for targeted regulatory sites and the functional modulation can be considered and iteratively fine-tuned in the directed design

procedure on the basis of the affinity and allosteric signaling, respectively.

Efficacy-based design of highly specific allosteric drug candidates for a selected member of the glucagon receptor family

Due to the high structural similarity among GPCRs, particularly within the same receptor family, the consideration of drug specificity to a target receptor is essential for reducing the likelihood of the off-target toxicity. In this section, we show how specificity towards a target receptor can be achieved by using the ligand efficacy as an optimization parameter. GIPR and GLP1R from the glucagon receptor family share remarkable structural resemblance, with C_α root mean square deviation (RMSD) value at 0.9 Å and approximately 40% sequence similarity [50]. Notably, the endogenous GLP1R ligand, GLP-1, also partially activates GIPR. Although both GIP (gastric inhibitory polypeptide, an endogenous ligand of GIPR) and GLP-1 regulate food intake and satiety while stimulating insulin secretion in pancreatic β-cells, their effects on glucagon secretion from pancreatic α-cells differ. Specifically, GIP exhibits a glucagonotropic effect during hypoglycemia, whereas GLP-1 exerts a glucagonostatic effect during hyperglycemia [13,51]. Additionally, GIPR activation facilitates fat storage by stimulating lipoprotein lipase activity and enhancing both fatty acid and glucose uptake [51], and GLP1R activation mitigates fat accumulation and promotes lipid breakdown [13].

Traditional approaches in drug design typically focus on binding affinity-driven specificity, where drugs bind their target receptors (or target binding sites as discussed in previous section) with higher affinity compared to non-targets. However, the specificity of an allosteric drug's effects can diminish at high drug concentrations, leading to off-target effects. Therefore, an alternative efficacy-based strategy, wherein a ligand preferentially activates or inhibits its target receptor while exhibiting minimal effects on off-target receptors even at saturating concentrations, has been developed [52]. Here, we implemented the idea of efficacy-based effector design, using an advantage of the two-component allosteric mode of action – binding and signaling. Our goal here is to achieve specificity of the effector based on the second component – allosteric signaling and respective response in the target. Working with the closely related GLP1R and GIPR, we show that regardless of comparable binding free energies in both receptors, one can still obtain drug specificity on the basis of distinct allosteric response – efficient in one protein, while absent in another.

We focus on designing effectors with high specificity in GLP1R. From targeted analysis on the APMs of GLP1R and GIPR, residue positions correspond to the strongest positive modulation probes (top 20% in Δh_{GA}) were identified in each of the receptors, allowing us to define the allosteric locations in GLP1R and GIPR. Focusing on strongly modulating residues that are not involved in the binding of known ligands, we obtained fragment *3* (targeting GLP1R) as an example, derivatives of which were characterized based on their efficacies, in terms of allosteric modulation and predicted affinity, in GLP1R (Figure 6A, above). Importantly, promising derivatives with Δh_{GA} and Δg_{bind} above the respective averages were shortlisted (highlighted in rectangle) and subsequently examined in a counter-screen on GIPR (Figure 6A, below; see the flowchart in Figure S4, Supplementary File 1 for additional detail). The derived compounds *3-1* and *3-2* (SMILES available in Supplementary File 1, Table S5), which likely cause conformational changes in the Gα binding site in GLP1R, while not affecting that in GIPR, are shown here as examples (Figure 6B).

Discussion

As the largest family of human membrane proteins and with crucial roles in virtually all physiological processes, G protein-coupled receptors (GPCRs) is one of the major drug targets associated with over 30% of all FDA-approved drugs [53]. At the same time, only 108 out of nearly 800 GPCRs encoded in the human genome are currently targeted by approved drugs [10]. A substantial number of GPCRs remain classified as *undruggable*, primarily due to challenges in targeting their endogenous/orthosteric sites. In particular, the high structural and sequence conservation within GPCR families complicates drug development, making it difficult to specifically target a receptor of interest among closely related ones in the same family/subtype, often leading to undesirable side effects [8,54]. To address this challenge, development of allosteric effectors/modulators that bind outside of conserved orthosteric sites has emerged as a promising strategy for GPCRs [2,9,15]. Allosteric effectors offer several advantages, including higher specificity to a target of interest compared to orthosteric ligands, which may reduce the likelihood of off-target effects. In GPCRs, positive and negative allosteric modulators acting together with agonists/antagonists and allosteric ligands provide an even more precise and subtle control over the receptors' function [9].

Using our Structure-Based Statistical Mechanical Model of Allostery (SBSMMA, [27,28]) which quantitatively accounts for the causality and energetics of allosteric effects induced by different perturbations, we show here how to explore the causal mechanisms of allostery in GPCRs, providing quantitative insights into the functional modulation of these receptors. We specifically focused here on a rigorous analysis of the human ADRB2 (class A) and GLP1R (class B1). Particularly, using the ASMs, we observed patterns of conserved allos-

teric communication in the chemokine family of class A GPCRs (Figure 1C), while showing the role of sequence variations in the diversification of signaling in two representatives – CCR1 and CCR5. The APM analysis, in turn, revealed allosteric locations unique to receptors of interest which can be instrumental in the design of structure-specific drug candidates (Figure 1D-F). The strength of allosteric signaling originated by endogenous and synthetic ligands and modulators in ADRB2 and GLP1R was quantified at single-residue resolution. Using SBSMMA, we delineated the allosteric effects of agonists, antagonists, and PAMs/NAMs, quantifying the allosteric signaling biased towards Gα or β-arrestin binding sites (Figures 2, 3 and Tables S3 and S4 in Supplementary File 1). We also showed that starting with a consideration of long-range communication from individual agonists and antagonists to the binding sites of Gα and β-arrestin, as well as the effects caused by modulators, it is possible to model complex regulation of the receptor function by combinations of ligands and modulators, corroborated by experimental findings (Figure 3).

Motivated by the omnipresence of allosteric connectivity in GPCRs and critical demands on development of allosteric drugs for various important but "difficult" targets, we present a generic framework for rational design of allosteric effectors for GPCRs. The phenomenon of mode switching in a series of similar compounds with divergent effects [8] and the engineering of an antagonist based on the structure of an agonist [55] prompted us to explore a potential for tuning and optimization of both allosteric effects and binding affinity – two components of the allosteric drug action – starting from fragment seeds. To this end, we used our *directed design protocol* [30], simultaneously designing and adjusting allosteric site-ligand pairs in order to achieve required affinity and modulation strength. As a proof of principle, we pursued here the controllable design of switchable agonist–antagonist pairs for GLP1R. First, we used preliminary information on the locations of known ligands as starting points for developing new effectors (Figure 4). Then, we demonstrated that comprehensive description of intrinsic allosteric communication provided by the APMs enables *de novo* design of an array of putative agonists and antagonists (Figure 5). Importantly, some of the newly designed effectors exhibit higher predicted binding free energies and stronger allosteric modulation compared to those of known ligands, demonstrating a great potential of the approach in general, and in obtaining structurally similar effectors with switchable modes of action. Lastly, an efficacy-based approach in identifying new effectors with highly specific regulation of a target receptor (exemplified with GLP1R), while causing minimal allosteric effects when bound to another closely related receptor (GIPR) was introduced. This case study was motivated by the overlapping roles of GIPR antagonists and GLP1R agonists in the treatment of diabetes and obesity [56] and by the observation that a combination of a GIPR antagonist with a GLP1R agonist leads to synergistic effects in reducing body weight with minimal or no adverse effects [57,58]. Importantly, the latter observation suggests that a compound acting as an agonist for GLP1R and, simultaneously, as an antagonist for GIPR could offer therapeutic advantages. Overall, controllable design of "molecular switches" that share similar (or the same) structures but varied allosteric effects in different proteins, starting from single precursor molecules, can be a promising strategy for achieving distinct regulation of multiple proteins of interest, which should be pursued and further developed.

To conclude, we show that most of the effectors considered in the examples here, including various agonists and antagonists, allosterically modulate the dynamics of intracellular binding sites of Gα and/or β-arrestin. A quintessence of GPCR function is the inherent allosteric mechanisms coupling extracellular ligand binding to conformational changes in the intracellular parts of the receptor. We focus on exploring the underlying allosteric mechanisms and using them for development of allosteric drug candidates. We, indeed, found a multitude of latent allosteric regulatory sites, revealing a great versatility for the design of allosteric drugs with unique characteristics. Based on the *directed design protocol* [23,30], we elaborated here a computational framework for obtaining structure-specific and switchable allosteric drug candidates, opening a perspective for precise and tunable targeting of GPCRs with unmet clinical needs. While the computational design of allosteric drug candidates proposed here will undoubtedly benefit from complementary experimental data and verification, new ideas and advances in the methodology demonstrated in this work should be considered as a starting point in the quest for rational design of GPCR drugs.

Materials and Methods

Structure-based statistical mechanical model of allostery (SBSMMA)

In this work, we use our structure-based statistical mechanical model of allostery (SBSMMA, [27,28]), which accounts for causality and quantifies the energetics of allosteric communication. The model considers a source of allosteric signal, such as ligand binding, mutation and different combinations thereof, as a perturbation. There are three major steps in the model briefly described here (see [16,27–29] for details). First, protein dynamics is modeled based on the elastic network model of the proteins and uses a C*a* harmonic model with the energy function $E^{(P)}(\boldsymbol{r}) = E^{(0)}(\boldsymbol{r}) + \alpha_P V^{(P)}(\boldsymbol{r})$, where the first term is an energy function of the

unperturbed state $E^{(0)}(\boldsymbol{r}) = \frac{1}{2}\sum_{i,j} k_{ij}\left(d_{ij} - d_{ij}^{0}\right)^2$ and second term describes a perturbation that causes allosteric signaling [27,28]. Second, the energetics of configurational ensembles of the original (free) and perturbed structures at the single residue level is characterized via calculation of the microscopic allosteric potential $U_i(\sigma) = \frac{1}{2}\sum_{\mu}\varepsilon_{\mu,i}\sigma_{\mu}^2$, where $\varepsilon_{\mu,i} = \sum_j \left|\boldsymbol{e}_{\mu,i} - \boldsymbol{e}_{\mu,j}\right|^2$ are intensity parameters defined via normal modes $\boldsymbol{e}_{\mu}$ characterizing the configurational ensemble of a native/perturbed state obtained from the Hessian matrix $K = \partial^2 E/\partial \boldsymbol{r}_i \partial \boldsymbol{r}_j$. Third, using the microscopic allosteric potential that evaluates the allosteric signal delivered on a residue, ensembles of configurations of neighboring residues are considered via integrating over all of their possible displacements, allowing one to calculate the per-residue partition function as $z_i = \Pi_{\mu}\left(\pi 2 k_B T/\varepsilon_{\mu,i}\right)^{1/2}$ and, consequently, the free energy $g_i = -k_B T \ln z_i$ associated with an allosteric signal transmitted to a residue. Then, considering two (original/perturbed) protein states, the allosteric signal delivered to a residue as a result of a perturbation (P) is represented via the free energy difference $\Delta g_i^{(P)} = \frac{1}{2} k_B T \sum_{\mu} \ln \frac{\varepsilon_{\mu,i}^{(P)}}{\varepsilon_{\mu,i}^{(0)}}$. The free energy difference $\Delta g_i^{(P)}$ (kcal/mol) shows results of allosteric signaling: positive $\Delta g_i^{(P)} > 0$ suggests a potential for conformational changes of residue *i*, and stabilization of the residue in the case of negative $\Delta g_i^{(P)} < 0$. Allosteric signaling exerted to a protein site is measured by averaging the per-residue $\Delta g_i^{(P)}$ of all residues *i* forming a site of interest $\Delta g_{site}^{(P)} = \langle \Delta g_i^{(P)} \rangle_{site}$.

There are several important aspects in using outputs of SBSMMA for the analysis of allosteric communication and signaling. The deviation of the free energy difference $\Delta g_i^{(P)}$ from its mean value over all residues of a protein chain containing residue *i*, called *allosteric modulation* $\Delta h_i^{(P)}$, is used for estimating a purely allosteric effect: $\Delta h_i^{(P)} = \Delta g_i^{(P)} - \langle \Delta g_i^{(P)} \rangle_{Chain}$.

The allosteric effect of mutations is generalized by considering perturbations at a residue position $\Delta h_i^{(m\downarrow\uparrow)}$, termed as the allosteric modulation range $\Delta h_i^{(m\downarrow\uparrow)} = \Delta h_i^{(m\uparrow)} - \Delta h_i^{(m\downarrow)}$, which evaluates an allosteric signal exerted at residue *i* caused by a mutation from the smallest (Ala/Gly-like) to the bulkiest residues (for example, Phe or Trp) at residue position *m*. A complete picture of the allosteric response upon residue-by-residue mutations is used for derivation of the Allosteric Signaling Maps (ASMs), which provides an exhaustive description of allosteric signaling in a protein. The ASMs are used here for the analysis of conserved and specific patterns of allosteric communication in GPCRs. The Allosteric Probing Maps (APMs) are derived via calculation of the allosteric modulation on residue *i* originated by binding of a small probe to a three-residue segment of the protein modeled sequentially from residue 1 to residue N-2 (for a protein chain of N residues). The allosteric modulation caused by a probe is evaluated as a background-free effect: $\Delta h_i^{(Probe)} = \Delta g_i^{(Probe)} - \langle \Delta g_i^{(Probe)} \rangle_{Chain}$. The exhaustive probing simulated in APMs is used here for finding locations in GPCRs that send allosteric signals of required mode (negative or positive) and strength to regulated sites upon binding of a probe. For example, "top 20% probes" consists of residues in probed segments that are among those originating top 20% strongest $\Delta h_{site}^{(Probe)}$ (both positive and negative) in a targeted site.

Data collection and analysis of protein structures and sequences

The AlphaFold GPCR structures. In this study, we used AlphaFold models of GPCRs [59]. The protein structures predicted by AlphaFold were filtered based on a per-residue confidence metric, specifically the predicted Local Distance Difference Test (pLDDT) score [60]. Residues with pLDDT scores below 70 were classified as low-confidence residues (LCRs). Continuous stretches of low-confidence residues in the N-terminal and C-terminal loops were excluded from consideration. Additionally, we required that structures contain no continuous stretches of LCRs longer than 30 residues, and the total proportion of LCRs not exceeding 25%. This threshold was established due to the presence of three intracellular loops and three extracellular loops characteristic of GPCR structures. The accession codes of all AlphaFold structures used here are available in Table S6 (Supplementary File 1) as well as on AlloMAPS.

Structure-based sequence alignment. Structure-based sequence alignment of the AlphaFold-predicted structures within the same receptor family was performed using 3D-Coffee (version 13.46.1. b8b01e06). This process incorporated the SAP (TM-align) algorithm [61] and the slow_pair alignment method [62]. Based on the resulting sequence alignment (Supplementary File 2), the average ASM was calculated for the chemokine receptors family.

Binding residue identification. Binding residues were identified from X-ray crystallographic structures by applying a distance cutoff of 4 Å from any heavy atom of a bound ligand or a protein (e.g., GA or Arr). Residues within this distance were classified as binding residues. The accession codes of all crystallographic structures used here are available in Supplementary Figure S1, Table S2.

Generic design framework of allosteric effectors

We provided flowcharts and detailed description on the design framework in Figures S1 and S4 (Supplementary File 1). Briefly here, the ZINC database [45] was used for deriving compounds from selected DSI-poised library fragments [44]

using a substructure-based search method in all case studies. The PAINS [46] and Brenk [47] filters were applied to exclude potentially toxic compounds, and compounds with promising oral drug-like properties were prioritized based on QED values (QED > 0.8). All fragments and derived compounds were evaluated in terms of the predicted binding free energy from molecular docking and allosteric modulation (using SBSMMA) considering the 10 poses with the highest affinities. Particularly, in the first case starting from known ligand locations (Figure 4), 8,609 compounds were derived from selected fragment *1*. In the *de novo* design of switchable ligands (Figure 5), residues in the top 20% probes targeting GA site identified from the APM were used to define initial locations for exhaustive fragment mapping. Subsequently, 2,634 compounds were derived from the switchable fragment *2*. Similar procedure was used for the efficacy-based design (Figure 6), we obtained 3830 compounds (Supplementary File 4) with high specificity in GLP1R starting from fragment *3*.

Molecular docking. Molecular docking was performed using AutoDock Vina 1.2.0 [63]. A docking exhaustiveness parameter of 32 was applied to ensure adequate searching. Docking was conducted using predefined ligands and docking boxes on receptors. Predicted structures of ADRB2, GLP1R and GIPR from AlphaFold were used (IDs: P07550, P43220 and P48546). For docking box placement, the geometric center of predefined binding site residues was used as the box center (see Tables S7–S11, Supplementary File 1 for details).

CRediT authorship contribution statement

Bingxue Dong: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Wei-Ven Tee:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Igor N. Berezovsky:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the OF-IRG grant MOH-001402-00 to INB.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jmb.2025.169293.

Received 18 March 2025;
Accepted 9 June 2025;
Available online 11 June 2025

Keywords:
G protein-coupled receptors;
drug design;
allostery in GPCRs;
allosteric drugs;
allosteric mutations

† Both authors are first co-authors.

References

[1]. Rosenbaum, D.M., Rasmussen, S.G., Kobilka, B.K., (2009). The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363.
[2]. Zhang, M., Chen, T., Lu, X., Lan, X., Chen, Z., Lu, S., (2024). G protein-coupled receptors (GPCRs): advances in structures, mechanisms, and drug discovery. *Signal Transduct. Target. Ther.* **9**, 88.
[3]. Yang, D., Zhou, Q., Labroska, V., Qin, S., Darbalaei, S., Wu, Y., et al., (2021). G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduct. Target. Ther.* **6**, 7.
[4]. Johnson, M., (2006). Molecular mechanisms of beta(2)-adrenergic receptor function, response, and regulation. *J. Allergy Clin. Immunol.* **117**, 18–24. quiz 5.
[5]. Proudfoot, A.E., (2002). Chemokine receptors: multifaceted therapeutic targets. *Nature Rev. Immunol.* **2**, 106–115.
[6]. Liang, Y.-L., Belousoff, M.J., Zhao, P., Koole, C., Fletcher, M.M., Truong, T.T., et al., (2020). Toward a structural understanding of class B GPCR peptide binding and activation. *Mol. Cell* **77**, 656–668.
[7]. Jamaluddin, A., Gorvin, C.M., (2023). RISING STARS: Targeting G protein-coupled receptors to regulate energy homeostasis. *J. Mol. Endocrinol.* **70**
[8]. Wenthur, C.J., Gentry, P.R., Mathews, T.P., Lindsley, C. W., (2014). Drugs for allosteric sites on receptors. *Annu. Rev. Pharmacol. Toxicol.* **54**, 165–184.
[9]. Conn, P.J., Lindsley, C.W., Meiler, J., Niswender, C.M., (2014). Opportunities and challenges in the discovery of allosteric modulators of GPCRs for treating CNS disorders. *Nature Rev. Drug Discov.* **13**, 692–708.
[10]. Xie, X., Yu, T., Li, X., Zhang, N., Foster, L.J., Peng, C., et al., (2023). Recent advances in targeting the "undruggable" proteins: from drug discovery to clinical trials. *Signal Transduct. Target. Ther.* **8**, 335.
[11]. Cary, B.P., Zhang, X., Cao, J., Johnson, R.M., Piper, S.J., Gerrard, E.J., et al., (2023). New insights into the structure and function of class B1 GPCRs. *Endocr. Rev.* **44**, 492–517.
[12]. Wang, L., Wang, N., Zhang, W., Cheng, X., Yan, Z., Shao, G., et al., (2022). Therapeutic peptides: current

applications and future directions. *Signal Transduct. Target. Ther.* **7**, 48.
[13]. Zheng, Z., Zong, Y., Ma, Y., Tian, Y., Pang, Y., Zhang, C., et al., (2024). Glucagon like peptide 1 receptor: mechanisms and advances in therapy. *Signal Transduct. Target. Ther.* **9**, 234.
[14]. García-Cárceles, J., (2018). Novel strategies targeting G-protein-coupled receptors: an overview. *Cell Biol.*, 1–9. https://doi.org/10.1002/9780470015902.a0028406.
[15]. Tan, Z.W., Tee, W.V., Berezovsky, I.N., (2022). Learning about allosteric drugs and ways to design them. *J. Mol. Biol.* **434**, 167692.
[16]. Guarnera, E., Berezovsky, I.N., (2019). On the perturbation nature of allostery: sites, mutations, and signal modulation. *Curr. Opin. Struct. Biol.* **56**, 18–27.
[17]. Guarnera, E., Berezovsky, I.N., (2016). Allosteric sites: remote control in regulation of protein activity. *Curr. Opin. Struct. Biol.* **37**, 1–8.
[18]. Gunasekaran, K., Ma, B., Nussinov, R., (2004). Is allostery an intrinsic property of all dynamic proteins?. *Proteins* **57**, 433–443.
[19]. Latorraca, N.R., Venkatakrishnan, A.J., Dror, R.O., (2017). GPCR dynamics: structures in motion. *Chem. Rev.* **117**, 139–155.
[20]. Heydenreich, F.M., Marti-Solano, M., Sandhu, M., Kobilka, B.K., Bouvier, M., Babu, M.M., (2023). Molecular determinants of ligand efficacy and potency in GPCR signaling. *Science* **382**, eadh1859.
[21]. Changeux, J.P., Christopoulos, A., (2016). Allosteric modulation as a unifying mechanism for receptor function and regulation. *Cell* **166**, 1084–1102.
[22]. Thal, D.M., Glukhova, A., Sexton, P.M., Christopoulos, A., (2018). Structural insights into G-protein-coupled receptor allostery. *Nature* **559**, 45–53.
[23]. Tee, W.V., Berezovsky, I.N., (2024). Allosteric drugs: new principles and design approaches. *Curr. Opin. Struct. Biol.* **84**, 102758.
[24]. Hauser, A.S., Chavali, S., Masuho, I., Jahn, L.J., Martemyanov, K.A., Gloriam, D.E., et al., (2018). Pharmacogenomics of GPCR drug targets. *Cell* **172**, 41–54.e19.
[25]. Chen, K.M., Keri, D., Barth, P., (2020). Computational design of G Protein-coupled Receptor allosteric signal transductions. *Nature Chem. Biol.* **16**, 77–86.
[26]. Guarnera, E., Berezovsky, I.N., (2020). Allosteric drugs and mutations: chances, challenges, and necessity. *Curr. Opin. Struct. Biol.* **62**, 149–157.
[27]. Guarnera, E., Berezovsky, I.N., (2016). Structure-based statistical mechanical model accounts for the causality and energetics of allosteric communication. *PLoS Comput. Biol.* **12**, e1004678.
[28]. Guarnera, E., Berezovsky, I.N., (2019). Toward comprehensive allosteric control over protein activity. *Structure* **27**, 866–878.e1.
[29]. Tee, W.V., Tan, Z.W., Lee, K., Guarnera, E., Berezovsky, I.N., (2021). Exploring the allosteric territory of protein function. *J. Phys. Chem. B* **125**, 3763–3780.
[30]. Tee, W.V., Lim, S.J., Berezovsky, I.N., (2024). Toward the design of allosteric effectors: gaining comprehensive control of drug properties and actions. *J. Med. Chem.* **67**, 17191–17206.
[31]. Tee, W.V., Raechell, G.E., Berezovsky, I.N., (2025). Sequence and structure at play in designing allosteric drugs and alleviating the drug resistance. *J. Mol. Biol.*169206.
[32]. Tee, W.V., Guarnera, E., Berezovsky, I.N., (2020). Disorder driven allosteric control of protein activity. *Curr. Res. Struct. Biol.* **2**, 191–203.
[33]. Tan, Z.W., Tee, W.V., Samsudin, F., Guarnera, E., Bond, P.J., Berezovsky, I.N., (2022). Allosteric perspective on the mutability and druggability of the SARS-CoV-2 Spike protein. *Structure* **30**, 590–607.e4.
[34]. Tee, W.V., Tan, Z.W., Guarnera, E., Berezovsky, I.N., (2022). Conservation and diversity in allosteric fingerprints of proteins for evolutionary-inspired engineering and design. *J. Mol. Biol.*167577.
[35]. Guarnera, E., Tan, Z.W., Zheng, Z., Berezovsky, I.N., (2017). AlloSigMA: allosteric signaling and mutation analysis server. *Bioinformatics* **33**, 3996–3998.
[36]. Tan, Z.W., Guarnera, E., Tee, W.V., Berezovsky, I.N., (2020). AlloSigMA 2: paving the way to designing allosteric effectors and to exploring allosteric effects of mutations. *Nucleic Acids Res.* **48**, W116–W124.
[37]. Tan, Z.W., Tee, W.V., Guarnera, E., Booth, L., Berezovsky, I.N., (2019). AlloMAPS: allosteric mutation analysis and polymorphism of signaling database. *Nucleic Acids Res.* **47**, D265–D270.
[38]. Tan, Z.W., Tee, W.V., Guarnera, E., Berezovsky, I.N., (2023). AlloMAPS 2: allosteric fingerprints of the AlphaFold and Pfam-trRosetta predicted structures for engineering and design. *Nucleic Acids Res.* **51**, D345–D351.
[39]. Arimont, M., Sun, S.L., Leurs, R., Smit, M., de Esch, I.J. P., de Graaf, C., (2017). Structural analysis of chemokine receptor-ligand interactions. *J. Med. Chem.* **60**, 4735–4779.
[40]. Lorente, J.S., Sokolov, A.V., Ferguson, G., Schioth, H.B., Hauser, A.S., Gloriam, D.E., (2025). GPCR drug discovery: new agents, targets and indications. *Nature Rev. Drug Discov.* **24**, 458–479.
[41]. Ahn, S., Kahsai, A.W., Pani, B., Wang, Q.T., Zhao, S., Wall, A.L., et al., (2017). Allosteric “beta-blocker” isolated from a DNA-encoded small molecule library. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1708–1713.
[42]. Tee, W.V., Guarnera, E., Berezovsky, I.N., (2018). Reversing allosteric communication: from detecting allosteric sites to inducing and tuning targeted allosteric response. *PLoS Comput. Biol.* **14**, e1006228.
[43]. Bueno, A.B., Sun, B., Willard, F.S., Feng, D., Ho, J.D., Wainscott, D.B., et al., (2020). Structural insights into probe-dependent positive allosterism of the GLP-1 receptor. *Nature Chem. Biol.* **16**, 1105–1110.
[44]. Cox, O.B., Krojer, T., Collins, P., Monteiro, O., Talon, R., Bradley, A., et al., (2016). A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330.
[45]. Irwin, J.J., Tang, K.G., Young, J., Dandarchuluun, C., Wong, B.R., Khurelbaatar, M., et al., (2020). ZINC20—a free ultralarge-scale chemical database for ligand discovery. *J. Chem. Inf. Model.* **60**, 6065–6073.
[46]. Baell, J.B., Holloway, G.A., (2010). New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **53**, 2719–2740.

[47]. Brenk, R., Schipani, A., James, D., Krasowski, A., Gilbert, I.H., Frearson, J., et al., (2008). Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem* **3**, 435–444.
[48]. Bickerton, G.R., Paolini, G.V., Besnard, J., Muresan, S., Hopkins, A.L., (2012). Quantifying the chemical beauty of drugs. *Nature Chem.* **4**, 90–98.
[49]. Tang, X.L., Wang, Y., Li, D.L., Luo, J., Liu, M.Y., (2012). Orphan G protein-coupled receptors (GPCRs): biological functions and potential drug targets. *Acta Pharmacol. Sin.* **33**, 363–371.
[50]. Li, W., Zhou, Q., Cong, Z., Yuan, Q., Li, W., Zhao, F., et al., (2024). Structural insights into the triple agonism at GLP-1R, GIPR and GCGR manifested by retatrutide. *Cell Discov.* **10**, 77.
[51]. Fukuda, M., (2021). The role of GIP receptor in the CNS for the pathogenesis of obesity. *Diabetes* **70**, 1929–1937.
[52]. Brubaker, P.L., Drucker, D.J., (2002). Structure-function of the glucagon receptor family of G protein-coupled receptors: the glucagon, GIP, GLP-1, and GLP-2 receptors. *Recept. Channels* **8**, 179–188.
[53]. Sriram, K., Insel, P.A., (2018). G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs?. *Mol. Pharmacol.* **93**, 251–258.
[54]. Powers, A.S., Pham, V., Burger, W.A., Thompson, G., Laloudakis, Y., Barnes, N.W., et al., (2023). Structural basis of efficacy-driven ligand selectivity at GPCRs. *Nature Chem. Biol.* **19**, 805–814.
[55]. Shiriaeva, A., Park, D., Kim, G., Lee, Y., Hou, X., Jarhad, D.B., et al., (2022). GPCR agonist-to-antagonist conversion: enabling the design of nucleoside functional switches for the A2A adenosine receptor. *J. Med. Chem.* **65**, 11648–11657.
[56]. Sabnis, R.W., (2024). Novel GIPR antagonists for treating obesity and type 2 diabetes mellitus. *ACS Med. Chem. Lett.* **15**, 2091–2092.
[57]. Jensen, M.H., Sanni, S.J., Riber, D., Holst, J.J., Rosenkilde, M.M., Sparre-Ulrich, A.H., (2024). AT-7687, a novel GIPR peptide antagonist, combined with a GLP-1 agonist, leads to enhanced weight loss and metabolic improvements in cynomolgus monkeys. *Mol. Metab.* **88**, 102006.
[58]. Véniant, M.M., Lu, S.-C., Atangan, L., Komorowski, R., Stanislaus, S., Cheng, Y., et al., (2024). A GIPR antagonist conjugated to GLP-1 analogues promotes weight loss with improved metabolic parameters in preclinical and phase 1 settings. *Nature Metabol.* **6**, 290–303.
[59]. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al., (2022). AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444.
[60]. Mariani, V., Biasini, M., Barbato, A., Schwede, T., (2013). lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722–2728.
[61]. Taylor, W.R., Orengo, C.A., (1989). Protein structure alignment. *J. Mol. Biol.* **208**, 1–22.
[62]. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., et al., (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* **34**, W604–W608.
[63]. Eberhardt, J., Santos-Martins, D., Tillack, A.F., Forli, S., (2021). AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* **61**, 3891–3898.