Mechanistic Insights into Specific G Protein Interactions with Adenosine Receptors Revealed by Accelerated Molecular Simulations

Jinan Wang and Yinglong Miao*

Center for Computational Biology and Department of Molecular Biosciences, University of
Kansas, Lawrence, KS 66047, USA

* Corresponding email: miao@ku.edu

Abstract

Coupling between G-protein-coupled receptors (GPCRs) and the G proteins is a key step in cellular signaling. Despite extensive experimental and computational studies, the mechanism of specific GPCR-G protein coupling remains poorly understood. This has greatly hindered effective drug design of GPCRs that are primary targets of ~1/3 of currently marketed drugs. Here, we have employed all-atom molecular simulations using a robust Gaussian accelerated molecular dynamics (GaMD) method to decipher the mechanism of the GPCR-G protein interactions. Adenosine receptors (ARs) were used as model systems based on very recently determined cryo-EM structures of the A₁AR and A_{2A}AR coupled with the G_i and G_s proteins, respectively. Changing the G_i protein to the G_s led to increased fluctuations in the A₁AR and agonist adenosine (ADO), while agonist 5'-N-ethylcarboxamidoadenosine (NECA) binding in the A_{2A}AR could be still stabilized upon changing the G_s protein to the G_i. Free energy calculations identified one stable low-energy conformation for each of the ADO-A₁AR-G_i and NECA-A_{2A}AR-G_s complexes as in the cryo-EM structures, similarly for the NECA-A_{2A}AR-G_i complex. In contrast, the ADO agonist and G_s protein sampled multiple conformations in the ADO-A₁AR-G_s system. GaMD simulations thus indicated that the ADO-bound A₁AR preferred to couple with the G_i protein to the G_s, while the A_{2A}AR could couple with both the G_s and G_i proteins, being highly consistent with experimental findings of the ARs. More importantly, detailed analysis of the atomic simulations showed that the specific AR-G protein coupling resulted from remarkably complementary residue interactions at the protein interface, involving mainly the receptor transmembrane 6 helix and the G_α α5 helix and α4-β6 loop. In summary, the GaMD simulations have provided unprecedented insights into the dynamic mechanism of specific GPCR-G protein interactions at an atomistic level, which is expected to facilitate future drug design efforts of the GPCRs.

Introduction

G-protein-coupled receptors (GPCRs) are key cellular signaling proteins and represent primary targets of ~1/3 of currently marketed drugs.¹ Particularly, four subtypes (A₁, A_{2A}, A_{2B}, and A₃) of GPCRs mediate the effects of adenosine, an endogenous nucleoside modulator that plays a critical role in cytoprotective function. Adenosine receptors (ARs) have emerged as important therapeutic targets for treating many human diseases such as cardiac ischemia, neuropathic pain and cancer.² During function, the A₁AR and A₃AR preferentially couple to the G_{i/o} proteins, while the A_{2A}AR and A_{2B}AR preferentially couple to the G_s proteins. Nevertheless, increasing evidence suggests that GPCRs including the ARs can couple to multiple G proteins.³⁻⁶

Few complex structures of GPCRs coupled with the G protein or its mimic have been determined using X-ray crystallography or cryo-EM so far.⁷ ARs are the sole subfamily of GPCRs that have structures in complex with different G proteins, *i.e.*, the adenosine (ADO)-bound A₁AR coupled with the G_i protein⁸ and the 5'-N-ethylcarboxamido-adenosine (NECA)-bound A_{2A}AR coupled with an engineered G_s protein.⁹ Both structures were obtained via cutting-edge cryo-EM and published very recently in 2018. The GPCR-G protein complex structures provide valuable information about active conformations of the GPCRs and G proteins. However, they are rather static images. The dynamic mechanism of specific GPCR-G protein interactions remains unclear.

Experimental techniques including mutagenesis, nuclear magnetic resonance, hydrogen-deuterium exchange mass spectrometry, double electron-electron resonance spectroscopy and structural biology have been utilized to investigate GPCR-G protein interactions.¹⁰⁻¹³ While the C-terminal α_5 helix in the G $_{\alpha}$ subunit has been suggested as the primary driver for specific receptor recognition, the G $_{\alpha}$ α_N helix and receptor intracellular loop (ICL) 2 and transmembrane (TM) 6 helix further contribute to the GPCR-G protein coupling specificity. In addition, dynamic regions

in the complex and agonist binding can be crucial for the coupling through allosteric modulation.^{11,12}

A bioinformatics approach has been applied to determine a selectivity barcode (patterns of amino acids) of GPCR–G protein coupling.¹⁴ While universally conserved residues in the barcode allow GPCRs to bind and activate G protein in a similar manner, different receptors recognize the unique positions of the G-protein barcode through distinct residues. Molecular dynamics (MD) simulations have identified several important regions for coupling of the G protein with activated GPCRs, including the receptor TM6 and G_{α} $\alpha 5$ helices.^{15–17} MD simulations have shown that conformational dynamics of the GPCR-G protein complex depends on the bound ligands.^{18,19} Moreover, MD simulations have suggested that the binding of the active GPCR is necessary for nucleotide release from the G protein.^{20–23} However, due to limited timescales, conventional MD (cMD) simulations often suffer from insufficient sampling, precluding proper free energy calculations to characterize GPCR-G protein interactions quantitatively.

To overcome the limitations of cMD, enhanced sampling methods have been applied to investigate GPCR–G protein interactions. Umbrella sampling has been used to calculate free energy profiles of the TM6 outward movement during receptor coupling to the G proteins.¹⁷ Metadynamics simulations have been performed to investigate the dynamic effects of different GPCR ligands and intracellular binding partners²⁴ and examine differences of GPCRs coupled by the G protein versus its mimetic nanobody.²⁵ Nevertheless, these enhanced simulation methods require predefined collective variables and may apply constraints on the conformational space of the proteins. In this regard, a novel and robust Gaussian accelerated MD (GaMD) method has been developed to allow for unconstrained enhanced sampling and free energy calculations of large biomolecules^{26–28}. GaMD has been applied to successfully simulate protein folding^{26,27}, protein-ligand binding and unbinding^{26,27,29}, GPCR activation²⁹, large-scale conformational transitions of

the CRISPR-Cas9 gene-editing system³⁰, T cell receptor signaling protein³¹ and human dystonia related protein³², and so on. Notably, GaMD has been recently applied to capture spontaneous binding of the G-protein mimetic nanobody to a muscarinic GPCR.³³

In this study, we have employed all-atom enhanced sampling simulations using the robust GaMD method on the latest cryo-EM structures of the ADO-A₁AR-G_i and NECA-A_{2A}AR-G_s protein complexes, as well as complexes with the G proteins switched (**Table 1** and **Figure S1**). A computational model was prepared for the receptor-G protein complexes in explicit lipids and solvent (**Figure S2**). The GaMD simulations allowed us to characterize structural flexibility and low-energy conformations of the AR-G protein complexes, which provided important insights into the mechanism of specific GPCR–G protein interactions.

Results

Variations of structural flexibility in different adenosine receptor-G protein complexes. In GaMD simulations of the ADO-A₁AR-G_i and NECA-A_{2A}AR-G_s complexes, the receptors underwent small fluctuations except the extracellular loop 2 (ECL2) and TM6 intracellular end (**Figure S3**). Overall, the G proteins exhibited higher flexibility than the receptors, especially in the α 5 helix, α 4- β 6 loop and α 4- β 5 loop in the G $_{\alpha}$ subunit and terminal regions of the G $_{\beta\gamma}$ subunits. Both the A₁AR and A_{2A}AR showed flexibility change upon switching of the G proteins. For the A₁AR, changing the G_i protein to the G_s led to increased fluctuations in the ADO agonist and the receptor ECL2, TM6 intracellular end and helix 8 (H8) (**Figure 1A**). These motifs were suggested to be important in previous studies for activation of the A₁AR and receptor coupling with the G protein.^{8,34} For the A_{2A}AR, changing the G_s protein to the G_i, however, appeared to stabilize the receptor with slightly decreased fluctuations in the latter, other than the ECL2 and intracellular loop 2 (ICL2) regions (**Figure 1B**).

Next, we examined flexibility change of the G proteins upon coupling to the different receptors. In the G_s protein, the C-terminus of the G_α $\alpha 5$ helix exhibited higher fluctuations when the $A_{2A}AR$ was changed to the A_1AR (**Figure 1C**). In the G_i protein, while the G_α $\alpha 5$ helix became stabilized with lower fluctuations, the $\alpha 4$ - $\beta 5$ loop and switch III exhibited higher flexibility when the A_1AR was changed to the $A_{2A}AR$ (**Figure 1D**). These regions were shown earlier to play a key role in activation and receptor recognition of the G protein.^{21,35}

Distinct binding modes of the G proteins and agonists in adenosine receptors. Free energy profiles were calculated from GaMD simulations to identify low-energy conformations of the GPCR-G protein complexes. RMSD of the agonist relative to the cryo-EM structures and the distance between the receptor NPxxY motif in the TM7 intracellular end and the C-terminus of the G_α $\alpha 5$ helix were first used as the reaction coordinates. In the ADO- A_1AR - G_i and NECA- $A_{2A}AR$ - G_s protein complexes, both the G proteins and agonists maintained their cryo-EM conformations (**Figures 2A and 2B, Table 1**). In the NECA- $A_{2A}AR$ - G_i complex, the NECA agonist maintained the cryo-EM conformation as in the NECA- $A_{2A}AR$ - G_s complex, but the G_i protein sampled a different state with the receptor:NPxxY- G_α $\alpha 5$ distance decreased to ~ 11.2 Å (**Figure 2C**). The $G_{\alpha i}$ $\alpha 5$ helix moved ~ 2 Å towards the TM7 NPxxY motif of the $A_{2A}AR$ relative to the $G_{\alpha s}$ $\alpha 5$ helix in the NECA- $A_{2A}AR$ - G_s structure (**Table 1**). Nevertheless, the NECA- $A_{2A}AR$ - G_i complex adopted a stable low-energy conformation in the free energy profile (**Figure 2C**).

In the ADO- A_1AR - G_s system, the ADO agonist sampled two low-energy conformational states, denoted “L1” and “L2”, for which agonist RMSD relative to the cryo-EM conformation was ~ 3.0 Å and ~ 7.5 Å, respectively (**Figure 2D**). The “L1” conformation of ADO was similar to the cryo-EM structure with slight sliding of the purine ring by ~ 2 Å at the orthosteric site (**Figure 3A**). In the “L2” conformation, ADO formed interactions with residues Tyr^{1.35} and Tyr^{7.36} in the “sub-

pocket 2” of the A₁AR described earlier³⁶ (**Figure S4**). Residue superscripts denote Ballesteros and Weinstein (BW) numbering of GPCRs.³⁷ The G_s protein sampled two low-energy conformational states, which were similar to cryo-EM conformations of the G_i protein in the ADO-A₁AR-G_i complex and the G_s protein in the NECA-A_{2A}AR-G_s complex. The receptor:NPxxY-G_α:α5 distance was ~11.8 Å and ~13.5 Å in the G_i- and G_s-bound A₁AR, respectively (**Figure 2D** and **Table 1**).

High flexibility and conformational changes of ECL2 in the A₁AR. In the A₁AR, different conformational states of the ECL2 helix were identified from GaMD simulations, including the “open”, “semi-open” and “closed” (**Figure 3B**). In comparison, the ECL2 in the A_{2A}AR sampled only one low-energy conformational state (**Figure S5**). In the ADO-A₁AR-G_i complex, the ECL2 sampled a distinct “semi-open” conformation with ~5.5 Å RMSD in the helix region compared with the starting “open” cryo-EM conformation, for which the ECL2 helix tilted towards the receptor TM bundle by ~4 Å (**Figure 3B**). An energy barrier of ~1.1 kcal/mol was found between the “open” and “semi-open” states of ECL2 in the ADO-A₁AR-G_i complex (**Figure S5A**). In the ADO-A₁AR-G_s system, the receptor ECL2 transitioned between the “open” and “closed” states with ~1.2 kcal/mol energy barrier (**Figure 4A**). In the “closed” state, the ECL2 helix tilted towards the receptor TM bundle by ~9 Å relative to the “open” cryo-EM conformation (**Figure 3B**). Meanwhile, the TM2 extracellular domain could move outwards by ~10 Å (**Figure S6B**). Such movement was observed in binding of a covalent antagonist DU172 to the A₁AR that involved an induced fit mechanism.^{36,38} During conformational transition from the “open” state to the “closed” in ECL2 of the A₁AR, residue Trp156^{ECL2} switched its hydrogen bonding partner from Gly163^{ECL2} to Val166^{ECL2} (**Figure S6**). This finding was consistent with previous mutagenesis experiments,

suggesting that residues Trp156^{ECL2} and Val166^{ECL2} were important in the activation and allosteric modulation of the A₁AR.^{34,39}

Comparatively weak coupling between the A₁AR and G_s protein. Overall, the ADO-A₁AR-G_i, NECA-A_{2A}AR-G_s and NECA-A_{2A}AR-G_i complexes appeared to be stable during the GaMD simulations (**Table 1**). Each of them sampled only one low-energy conformation in the free energy profiles (**Figures 2, S7, S9 and S10**). In comparison, coupling of the G_s protein to the ADO-bound A₁AR was significantly weaker. The ADO-A₁AR-G_s system deviated from the simulation starting structure, visiting multiple low-energy conformational states. In particular, the TM6 intracellular end of the A₁AR sampled two distinct low-energy states, referred to as “Active” and “Over-active”, for which the Arg^{3.50}-Glu^{6.30} distance was ~17.5 Å and ~22.5 Å, respectively (**Figure 4B**). The “Active” state exhibited ~0.8 kcal/mol lower free energy than the “Over-active” state. The receptor TM6 intracellular end moved ~5 Å away from the TM bundle in the “Over-active” state compared with the “Active” state (**Figures 3C and 4B**). When the A₁AR visited the “Active” and “Over-active” states, the G_α α5 helix adopted an orientation angle of ~123° and ~130° (**Figure 4C**) and the distance between the G_α and G_β subunits increased by ~1.2 Å and ~3.2 Å, respectively (**Figure 4D**).

The above results suggested that the G_s protein could not stabilize the ADO-bound A₁AR. On the other hand, the G_{αs} and G_{βs} subunits tended to dissociate from each other when the G_s protein coupled to the A₁AR. In contrast, the G_{αi} and G_{βi} subunits formed closer interaction when the G_i protein coupled to the ADO-bound A₁AR, for which the G_{αi}-G_{βi} distance decreased by ~1.3 Å in the energy minimum conformation of the ADO-A₁AR-G_i complex (**Figure S10A**). Therefore, the A₁AR induced closer interaction of the G_α and G_β subunits in the G_i protein, but dissociation of

the $G_{\alpha s}$ and $G_{\beta s}$ subunits from each other (**Table 1**). In summary, coupling of ADO-bound A_1AR to the G_s protein was weaker than to the G_i protein.

Mechanism of specific adenosine receptor-G protein interactions. Sequence alignments of the ARs and G proteins (**Figure S11**) and detailed comparison of residue interactions at the protein interface (**Figure 5**) enabled us to identify the origin of specific AR-G protein interactions. In the A_1AR-G_i and $A_{2A}AR-G_s$ complexes, analysis of low-energy conformations identified from the GaMD simulations highlighted specific residue interactions (**Figures 5A and 5B**) that were similar to those obtained previously by comparing the A_1AR-G_i and β_2AR-G_s experimental structures.⁸ The last five residues of the $G_{\alpha} \alpha 5$ helix formed significantly stronger receptor interactions in the A_1AR-G_i complex than in the $A_{2A}AR-G_s$ complex. Residue Asp351 or GH5.22 with common G_{α} numbering (CGN)⁴⁰ in the G_i protein formed salt-bridge interactions with Arg^{3.53} and Lys^{8.49} in the A_1AR . In the G_s protein, residue Glu382 (GH5.24) formed similar salt-bridge interactions with Arg^{7.56} and Arg^{8.51} in the $A_{2A}AR$. However, these five residues formed closer van der Waals interactions with the TM3, TM5, TM6 and H8 of the A_1AR . Second, residues GH5.8 – GH5.21 of the $G_{\alpha} \alpha 5$ helix formed more interactions with the receptor ICL2 and TM5 helix in the $A_{2A}AR-G_s$ complex instead than in the A_1AR-G_i complex. In addition, the $\alpha 4$ – $\beta 6$ loop of the G protein formed distinct interactions with the receptor TM5 helix in the $A_{2A}AR-G_s$ complex compared with the A_1AR-G_i complex. Residues His347 (Gh4s6.13) and Tyr348 (Gh4s6.20) of the G_s protein formed non-polar interactions with Gln^{5.71} and Met^{5.72} in the $A_{2A}AR$ (**Figure 5B**). In contrast, residue Asp316 (Gh4s6.9) of the G_i protein formed polar interactions with residues Lys^{6.25} and Tyr^{6.26} in the A_1AR (**Figure 5A**).

In comparison, the $A_{2A}AR-G_i$ complex appeared to form more residue interactions at the protein interface (**Figure 5C**) than both the A_1AR-G_i and $A_{2A}AR-G_s$ complexes. Notably, the last five

residues of $G_{\alpha i}$ $\alpha 5$ helix formed extensive polar and non-polar interactions with the TM3, TM6, TM7, ICL2 and H8 of the $A_{2A}AR$. Residue Phe355 (GH5.26) formed close interactions with Lys^{6.29}, His^{6.32} and Ser^{6.36} of the $A_{2A}AR$. Residues GH5.8-GH5.21 of the $G_{\alpha i}$ $\alpha 5$ helix formed additional interactions with H8 of the $A_{2A}AR$, apart from those with the receptor TM5 and ICL2 as observed in the A_1AR - G_i and $A_{2A}AR$ - G_s complexes. Furthermore, the $\beta 2$ - $\beta 3$ loop of the G_i protein formed new interactions with the ICL2 of the $A_{2A}AR$. Residue Asp194 (Gs2s3.2) formed a salt-bridge with Arg111^{ICL2} in the $A_{2A}AR$. These interactions greatly contributed to strong coupling of the $A_{2A}AR$ and G_i protein, which showed stable low-energy conformations in the free energy profiles (e.g., **Figure 2**) and small fluctuations (**Figure 1**).

When the G_s protein coupled to the A_1AR , residue interactions at the protein interface were decreased overall (**Figure 5D**). In the “Active” state, the TM6 helix of the A_1AR formed significantly fewer interactions with the G_s protein (**Figure 5D**) than with the G_i protein (**Figure 5A**). The αN - $\beta 1$, $\beta 2$ sheet and $\beta 2$ - $\beta 3$ loop of the G_s protein involving residues R38 (Ghns1.2), A39 (Ghns1.3), H41 (GS1.2), D215 (Gs2s3.1) and V217 (GS3.1) formed new interactions with the TM2 helix and ICL2 of the A_1AR (**Figure 5D**). However, both clusters of residues in the G_{α} $\alpha 5$ helix (GH5.8-GH5.21 and GH5.22-GH5.26) greatly reduced receptor interactions in the A_1AR - G_s system compared with the A_1AR - G_i and $A_{2A}AR$ - G_s complexes. Similar results were observed in the “Over-active” conformation of the A_1AR - G_s system (**Figure S12**). Therefore, reduced residue interactions were found at the protein interface between the A_1AR and G_s protein, leading to their weaker coupling compared with the other three AR-G protein complexes (**Figure 5**).

In summary, the ADO-bound A_1AR preferred to bind the G_i protein to the G_s , while the $A_{2A}AR$ could bind both the G_s and G_i proteins (**Figure 6**). For the A_1AR , when the G_i protein was changed to the G_s , the receptor ECL2 and TM6 intracellular end underwent higher fluctuations and sampled multiple conformational states, similarly for the agonist and G protein. The G_s protein could not

stabilize ADO binding in the A₁AR, and vice versa. Coupling of the A₁AR to the G_s protein was significantly weaker than to the G_i protein. The G_{αs} and G_{βs} subunits tended to dissociate from each other (**Figure 6A**). In contrast, both the G_s and G_i proteins could stabilize agonist NECA binding in the A_{2A}AR. The G_i protein became more compact with decreased distance between the G_α and G_β subunits, contributing to strong coupling of the G_i protein to the NECA-bound A_{2A}AR (**Figure 6B**).

Discussions

ARs are so far the only subfamily of GPCRs that have available experimental structures in complex with different G proteins. The AR-G protein complex structures were obtained via cutting-edge cryo-EM and published very recently in 2018. Here, all-atom GaMD enhanced simulations have unprecedentedly allowed us to characterize the structural flexibility and low-energy conformations of the different AR-G protein complexes. Further detailed analysis of the AR-G protein complex simulations has highlighted a network of remarkable complementary residue interactions at the protein interface, which are important for specific G protein coupling to the A₁AR and A_{2A}AR.

The ADO-bound A₁AR preferred to bind the G_i protein to the G_s. This was consistent with previous experimental studies that the A₁AR coupled to different G proteins with the following rank order of “preference”: G_i>G_s>G_q, and the G protein coupling depends on the bound agonist potency.^{4,5} Particularly, the NECA-bound A₁AR was shown to preferentially couple to the G_i protein compared with the G_s protein.⁴ Considering similar binding affinities of ADO and NECA in the A₁AR,^{41,42} the ADO-bound A₁AR likely prefers to couple to the G_i protein as well. When the A₁AR coupled with the G_s protein, the ADO agonist exhibited high fluctuations and sampled two different binding poses (“L1” and “L2”). In the “L2” binding pose, ADO formed interactions with residues Tyr^{1.35} and Tyr^{7.36} in the sub-pocket 2 of the A₁AR as described earlier.³⁶ This was

similar to the 5UIG X-ray structure of the A_{2A}AR,⁴³ in which the 8D1 antagonist interacted with the same residues of the A_{2A}AR (**Figure S4**). The ECL2 of the A₁AR was highly flexible (**Figures 1 and S3**) and sampled “open”, “semi-open” and “closed” conformations (**Figure S5**). Both experimental and computational studies suggested that flexibility of the ECL2 was important for activation and allosteric modulation of the A₁AR.^{39,44,45} Therefore, highly flexibility of the ECL2 contributed to activation of the A₁AR and receptor coupling to the G protein.

The A_{2A}AR could couple to both the G_s and G_i proteins. This correlated with a recent experimental study that the A_{2B}AR coupled with both the G_s and G_i proteins in human cells.³ The A_{2B}AR was able to activate different downstream signaling pathways via different G proteins (including the G_s and G_i) in the same cell type (e.g., HEK293 kidney and T24 bladder cancer cells) and couple to the same pathway via different G proteins in different cell types.³ Considering high similarity of A_{2A}AR and A_{2B}AR (72%), especially at the G protein coupling interface (**Figure S13**), we assume that the A_{2A}AR would also couple to both the G_s and G_i proteins.

With low-energy conformations of AR-G protein complexes obtained from the GaMD simulations, further analysis revealed that complementary residue interactions were key for specific GPCR-G protein coupling. When coupling to different G proteins, one receptor could change its conformation and flexibility (notably in the TM6 intracellular domain), similarly for one G protein as coupled to different receptors (**Figures 1 and 6**). Provided highly complementary residue interactions at the interface, the A_{2A}AR could strongly couple to the G_i protein in addition to the G_s. However, coupling of the A₁AR to the G_s protein became weaker than to the G_i, due to significantly reduced residue interactions (**Figure 5**). The complementary residue interactions were identified to involve the receptor TM6, TM5, H8 and ICL2, as well as the Gα α5 helix, α4-β6 loop and αN-β1 loop. These regions have been highlighted to be important determinants for

specific GPCR-G protein coupling in extensive experimental and computational studies as reviewed earlier.^{10-12,46}

In summary, the GaMD simulations with unconstrained enhanced sampling and free energy calculations have provided important insights into the mechanism of specific G protein coupling to the A₁AR and A_{2A}AR. Nevertheless, effects of binding different extracellular ligands (e.g., agonists of varied potencies and allosteric modulators) on the GPCR-G protein interactions are subject to future studies. Furthermore, challenges remain to accurately predict the thermodynamic and kinetic properties of the G protein binding to the GPCRs in order to fully understand the dynamics of GPCR-G protein interactions. It is important to characterize both the association and dissociation pathways of the G protein coupling to GPCRs. Developments in computing power and enhanced simulation methodologies will be needed to address these problems in the future.

Materials and Methods

Gaussian accelerated molecular dynamics (GaMD). GaMD enhances the conformational sampling of biomolecules by adding a harmonic boost potential to reduce the system energy barriers.²⁶ When the system potential $V(\vec{r})$ is lower than a reference energy E , the modified potential $V^*(\vec{r})$ of the system is calculated as:

$$V^*(\vec{r}) = V(\vec{r}) + \Delta V(\vec{r})$$

$$\Delta V(\vec{r}) = \begin{cases} \frac{1}{2}k(E - V(\vec{r}))^2, & V(\vec{r}) < E \\ 0, & V(\vec{r}) \geq E, \end{cases} \quad (1)$$

Where k is the harmonic force constant. The two adjustable parameters E and k are automatically determined on three enhanced sampling principles. First, for any two arbitrary potential values $v_1(\vec{r})$ and $v_2(\vec{r})$ found on the original energy surface, if $V_1(\vec{r}) < V_2(\vec{r})$, ΔV should be a monotonic function that does not change the relative order of the biased potential values; i.e.,

$V_1^*(\vec{r}) < V_2^*(\vec{r})$. Second, if $V_1(\vec{r}) < V_2(\vec{r})$, the potential difference observed on the smoothened energy surface should be smaller than that of the original; i.e., $V_2^*(\vec{r}) - V_1^*(\vec{r}) < V_2(\vec{r}) - V_1(\vec{r})$. By combining the first two criteria and plugging in the formula of $V^*(\vec{r})$ and ΔV , we obtain

$$V_{max} \leq E \leq V_{min} + \frac{1}{k}, \quad (2)$$

Where V_{min} and V_{max} are the system minimum and maximum potential energies. To ensure that **Eq. 2** is valid, k has to satisfy: $k \leq 1/(V_{max} - V_{min})$. Let us define: $k = k_0 \cdot 1/(V_{max} - V_{min})$, then $0 < k_0 \leq 1$. Third, the standard deviation (SD) of ΔV needs to be small enough (i.e. narrow distribution) to ensure accurate reweighting using cumulant expansion to the second order: $\sigma_{\Delta V} = k(E - V_{avg})\sigma_V \leq \sigma_0$, where V_{avg} and σ_V are the average and SD of ΔV with σ_0 as a user-specified upper limit (e.g., $10k_B T$) for accurate reweighting. When E is set to the lower bound $E = V_{max}$ according to **Eq. 2**, k_0 can be calculated as

$$k_0 = \min(1.0, k'_0) = \min\left(1.0, \frac{\sigma_0}{\sigma_V} \cdot \frac{V_{max} - V_{min}}{V_{max} - V_{avg}}\right), \quad (3)$$

Alternatively, when the threshold energy E is set to its upper bound $E = V_{min} + 1/k$, k_0 is set to:

$$k_0 = k''_0 \equiv \left(1 - \frac{\sigma_0}{\sigma_V}\right) \cdot \frac{V_{max} - V_{min}}{V_{avg} - V_{min}}, \quad (4)$$

If k''_0 is calculated between 0 and 1. Otherwise, k_0 is calculated using **Eq. 3**.

Energetic Reweighting of GaMD Simulations. For energetic reweighting of GaMD simulations to calculate potential of mean force (PMF), the probability distribution along a reaction coordinate is written as $p^*(A)$. Given the boost potential $\Delta V(r)$ of each frame, $p^*(A)$ can be reweighted to recover the canonical ensemble distribution $p(A)$, as:

$$p(A_j) = p^*(A_j) \frac{\langle e^{\beta \Delta V(r)} \rangle_j}{\sum_{i=1}^M \langle p^*(A_i) e^{\beta \Delta V(r)} \rangle_i}, \quad j = 1, \dots, M, \quad (5)$$

where M is the number of bins, $\beta = k_B T$ and $\langle e^{\beta \Delta V(r)} \rangle_j$ is the ensemble-averaged Boltzmann factor of $\Delta V(r)$ for simulation frames found in the j^{th} bin. The ensemble-averaged reweighting factor can be approximated using cumulant expansion:

$$\langle e^{\beta \Delta V(r)} \rangle = \exp \left\{ \sum_{k=1}^{\infty} \frac{\beta^k}{k!} C_k \right\}, \quad (6)$$

where the first two cumulants are given by:

$$\begin{aligned} C_1 &= \langle \Delta V \rangle, \\ C_2 &= \langle \Delta V^2 \rangle - \langle \Delta V \rangle^2 = \sigma_v^2. \end{aligned} \quad (7)$$

The boost potential obtained from GaMD simulations usually follows near-Gaussian distribution.²⁸ Cumulant expansion to the second order thus provides a good approximation for computing the reweighting factor.^{26,47} The reweighted free energy $F(A) = -k_B T \ln p(A)$ is calculated as:

$$F(A) = F^*(A) - \sum_{k=1}^2 \frac{\beta^k}{k!} C_k + F_c, \quad (8)$$

where $F^*(A) = -k_B T \ln p^*(A)$ is the modified free energy obtained from GaMD simulation and F_c is a constant.

System Setup. Cryo-EM structures of the ADO-A₁AR-G_i (PDB: 6D9H)⁸ and NECA-A_{2A}AR-G_s (PDB: 6GDG)⁹ were used for setting up simulation systems. Nanobody Nb35 in the cryo-EM structure of the NECA-A_{2A}AR-G_s was deleted for simulation. In the 6GDG cryo-EM structure, residues that were missing in extracellular loop (ECL) 2 and the C terminus of the A_{2A}AR were added using atomic coordinates obtained from the X-ray structure of the A_{2A}AR bound by the mini-G_s protein (PDB: 5G53)⁴⁸ after aligning the receptor transmembrane (TM) domain. Initial models of the ADO-A₁AR-G_s and NECA-A_{2A}AR-G_i protein complexes were obtained by switching the G proteins in the NECA-A_{2A}AR-G_s and ADO-A₁AR-G_i complexes after aligning the receptor TM domain (**Figure S1**). There was no clash between the ARs and G proteins.

According to previous findings, intracellular loop (ICL) 3 is highly flexible^{21,49} and removal of ICL3 does not appear to affect GPCR function. The ICL3 was thus omitted as in the cryo-EM structures for the simulations. In addition, helical domains of the G_i and G_s proteins missing in the cryo-EM structures were not included in the simulation models. This was based on earlier simulation of the β_2 AR-G_s complex, which showed that the helical domain fluctuated substantially.²¹ All chain termini were capped with neutral groups (acetyl and methylamide). All the disulphide bonds in the receptors and G proteins (i.e., Cys80^{3,25}-Cys169^{ECL2} and Cys260^{6,61}-Cys263^{ECL3} in the A₁AR, Cys74^{3,22}-Cys146^{ECL2}, Cys77^{3,25}-Cys166^{ECL2}, Cys71^{ECL1}-Cys159^{ECL2} and Cys259^{6,61}-Cys262^{ECL3} in the A_{2A}AR, and Cys121-Cys149 in the G β subunit of the G_s protein) that were resolved in the cryo-EM structures were maintained in the simulations. Using the *psfgen* plugin in VMD,⁵⁰ protein residues were set to the standard CHARMM protonation states at neutral pH. For each of the complex systems, the receptor was inserted into a palmitoyl-oleoyl-phosphatidyl-choline (POPC) bilayer with all overlapping lipid molecules removed using the membrane plugin in VMD. The system charges were then neutralized at 0.15M NaCl using the *solvate* plugin in VMD.⁵⁰ The simulation systems were summarized in **Table 1**, with an example computational model shown in **Figure S2**.

Simulation Protocol. The CHARMM36 parameter set⁵¹ was used for the adenosine receptors, G proteins and POPC lipids. Force field parameters of agonists ADO and NECA were obtained from the CHARMM ParamChem web server.^{52,53} For each of the AR-G protein complex systems, initial energy minimization, thermalization, and 20ns cMD equilibration were performed using NAMD2.12⁵⁴. A cutoff distance of 12 Å was used for the Van der Waals and short-range electrostatic interactions and the long-range electrostatic interactions were computed with the particle-mesh Ewald summation method. A 2-fs integration time step was used for all MD

simulations and a multiple-time-stepping algorithm was used with bonded and short-range non-bonded interactions computed every time step and long-range electrostatic interactions every two time steps. The SHAKE algorithm was applied to all hydrogen-containing bonds. The NAMD simulation started with equilibration of the lipid tails. With all other atoms fixed, the lipid tails were energy minimized for 1,000 steps using the conjugate gradient algorithm and melted with a constant number, volume, and temperature (NVT) run for 0.5 ns at 310 K. The four systems were further equilibrated using a constant number, pressure, and temperature (NPT) run at 1 atm and 310 K for 10 ns with 5 kcal/(mol· Å²) harmonic position restraints applied to the protein and ligand atoms. The system volume was found to decrease with a flexible unit cell applied and level off with a 10-ns NPT run, suggesting that solvent and lipid molecules in the system were well equilibrated. Final equilibration of each system was performed using a NPT run at 1 atm pressure and 310 K for 0.5 ns with all atoms unrestrained. After energy minimization and system equilibration, conventional MD simulations were performed on each system for 20 ns at 1 atm pressure and 310 K with a constant ratio constraint applied on the lipid bilayer in the X-Y plane.

With the NAMD output structure, along with the system topology and CHARM36 force field files, the *ParmEd* tool in the AMBER package was used to convert the simulation files into the AMBER format.⁵⁵ The GaMD module implemented in the GPU version of AMBER18^{26,55} was then applied to perform the GaMD simulation, which included a 8-ns short cMD simulation used to collect the potential statistics for calculating GaMD acceleration parameters, a 64-ns equilibration after adding the boost potential, and finally three independent 300-ns GaMD production simulations with randomized initial atomic velocities. All GaMD simulations were run at the “dual-boost” level by setting the reference energy to the lower bound. One boost potential is applied to the dihedral energetic term and the other to the total potential energetic term. The average and SD of the system potential energies were calculated every 800,000 steps (1.6 ns) for

all simulation systems. The upper limit of the boost potential SD, σ_0 was set to 6.0 kcal/mol for both the dihedral and the total potential energetic terms. Similar temperature and pressure parameters were used as in the NAMD simulations. A list of GaMD production simulations on the different ARs-G proteins complex systems is listed in **Table 1**.

Simulation Analysis. CPPTRAJ⁵⁶ and VMD⁵⁰ were used to analyze the GaMD simulation trajectories. Particularly, distances were calculated between the C α atoms of residues Arg^{3,50} and Glu^{6,30}, the center-of-mass (COM) distance between the receptor NPxxY motif and the last 5 residues of the G α α 5 helix, and the COM distance between the G α (excluding residues in the α N helix) and G β (excluding residues 2-45 in the N-terminus) subunits. Root-mean-square fluctuations (RMSFs) were calculated for the protein residues and ligands, averaged over three independent GaMD simulations and color coded for schematic representation of each complex system (**Figure 1**). The representative low-energy conformations of AR-G protein were used to analyze the residue interaction contacts. The residue contact network between the AR and G protein was computed using van der Waals contacts between atoms, as described in Reference 37.⁵⁷ For two-dimensional visualization, software Cytoscape⁵⁸ was utilized to plot the residue contact network.

The PyReweight⁴⁷ toolkit was used to reweight distances, root-mean-square deviations (RMSDs) and the G α α 5 orientation angle to compute the potential of mean force (PMF) profiles. A bin size of 1.0 Å was used for the distances and RMSDs values, and 5.0 degree for the G α α 5 orientation angle. The cutoff was set to 500 frames for 2D PMF calculations. The 2D PMF profiles were obtained for each simulation system regarding agonist RMSD relative to the cryo-EM conformation and the AR:NPxxY-G: α 5 distance (**Figure 2**), RMSD of the helix region in ECL2 relative to the cryo-EM structure and the AR:NPxxY-G: α 5 distance (**Figures 4A and S5**), the distance between atom NE1 of W156 and atom O of G163 and RMSD of the helix region in

ECL2 relative to the cryo-EM structure (**Figure S6**), the Arg^{3.50}-Glu^{6.30} and the AR:NPxxY-G:α5 distances (**Figures 4B and S7**), the G_α α5 orientation angle and the Arg^{3.50}-Glu^{6.30} distance (**Figures 4C and S9**), and increase of the G_α-G_β distance and the Arg^{3.50}-Glu^{6.30} distance (**Figures 4D and S10**). Time courses of these reaction coordinates obtained from the GaMD simulation were plotted in **Figures S14-S17**.

ASSOCIATED CONTENT

Supporting Information

Supplementary file containing figures (Figure S1-S17).

AUTHOR INFORMATION

Corresponding Author

miao@ku.edu

Author Contributions

Y.M. designed research; J.W. performed research; J.W. and Y.M. analyzed data; and J.W. and Y.M. wrote the paper.

Notes

The authors declare no competing financial interest.

Acknowledgements

We thank Waseem Ahmad for proofreading the manuscript. Computing time was provided on the Comet and Stanford EXtream supercomputers through the Extreme Science and Engineering

Discovery Environment award TG-MCB180049 and the Edison and Cori supercomputers through the National Energy Research Scientific Computing Center project M2874. This work was supported in part by the American Heart Association (Award 17SDG33370094) and the startup funding in the College of Liberal Arts and Sciences at the University of Kansas.

Reference

- (1) Finan, C.; Gaulton, A.; Kruger, F. A.; Lumbers, R. T.; Shah, T.; Engmann, J.; Galver, L.; Kelley, R.; Karlsson, A.; Santos, R. et al. The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **2017**, *9* (383).
- (2) Jacobson, K. A.; Gao, Z.-G. Adenosine receptors as therapeutic targets. *Nat. Rev. Drug Discov.* **2006**, *5*, 247.
- (3) Gao, Z.-G.; Inoue, A.; Jacobson, K. A. On the G protein-coupling selectivity of the native A2B adenosine receptor. *Biochem. Pharmacol.* **2018**, *151*, 201.
- (4) Cordeaux, Y.; IJzerman, A. P.; Hill, S. J. Coupling of the human A1 adenosine receptor to different heterotrimeric G proteins: evidence for agonist-specific G protein activation. *Br. J. Pharmacol.* **2004**, *143* (6), 705.
- (5) Cordeaux, Y.; Bridson, S. J.; Megson, A. E.; McDonnell, J.; Dickenson, J. M.; Hill, S. J. Influence of Receptor Number on Functional Responses Elicited by Agonists Acting at the Human Adenosine A1 Receptor: Evidence for Signaling Pathway-Dependent Changes in Agonist Potency and Relative Intrinsic Activity. *Mol. Pharmacol.* **2000**, *58* (5), 1075.
- (6) Stewart, G. D.; Valant, C.; Dowell, S. J.; Mijaljica, D.; Devenish, R. J.; Scammells, P. J.; Sexton, P. M.; Christopoulos, A. Determination of Adenosine A1 Receptor Agonist and Antagonist Pharmacology Using *Saccharomyces cerevisiae*: Implications for Ligand Screening and Functional Selectivity. *J. Pharmacol. Exp. Ther.* **2009**, *331* (1), 277.
- (7) Isberg, V.; Mordalski, S.; Munk, C.; Rataj, K.; Harpsøe, K.; Hauser, A. S.; Vroiling, B.; Bojarski, A. J.; Vriend, G.; Gloriam, D. E. GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res* **2016**, *44* (D1), D356.
- (8) Draper-Joyce, C. J.; Khoshouei, M.; Thal, D. M.; Liang, Y. L.; Nguyen, A. T. N.; Furness, S. G. B.; Venugopal, H.; Baltos, J. A.; Plitzko, J. M.; Danev, R. et al. Structure of the adenosine-bound human adenosine A1 receptor-Gi complex. *Nature* **2018**, *558* (7711), 559.
- (9) García-Nafria, J.; Lee, Y.; Bai, X.; Carpenter, B.; Tate, C. G. Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein. *Elife* **2018**, *7*, e35946.
- (10) Moreira, I. S. Structural features of the G-protein/GPCR interactions. *Biochim. Biophys. Acta. Gen. Subj.* **2014**, *1840* (1), 16.
- (11) Mahoney, J. P.; Sunahara, R. K. Mechanistic insights into GPCR-G protein interactions. *Curr. Opin. Struct. Biol.* **2016**, *41*, 247.
- (12) Preininger, A. M.; Meiler, J.; Hamm, H. E. Conformational Flexibility and Structural Dynamics in GPCR-Mediated G Protein Activation: A Perspective. *J. Mol. Biol.* **2013**, *425* (13), 2288.
- (13) Duc, N. M.; Kim, H. R.; Chung, K. Y. Structural mechanism of G protein activation by G protein-coupled receptor. *Eur. J. Pharmacol.* **2015**, *763* (Pt B), 214.

- (14) Flock, T.; Hauser, A. S.; Lund, N.; Gloriam, D. E.; Balaji, S.; Babu, M. M. Selectivity determinants of GPCR–G-protein binding. *Nature* **2017**, *545*, 317.
- (15) Sgourakis, N. G.; Garcia, A. E. The Membrane Complex between Transducin and Dark-State Rhodopsin Exhibits Large-Amplitude Interface Dynamics on the Sub-Microsecond Timescale: Insights from All-Atom MD Simulations. *J. Mol. Biol.* **2010**, *398* (1), 161.
- (16) Van Eps, N.; Altenbach, C.; Caro, L. N.; Latorraca, N. R.; Hollingsworth, S. A.; Dror, R. O.; Ernst, O. P.; Hubbell, W. L. Gi- and Gs-coupled GPCRs show different modes of G-protein binding. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (10), 2383.
- (17) Rose, A. S.; Elgeti, M.; Zachariae, U.; Grubmüller, H.; Hofmann, K. P.; Scheerer, P.; Hildebrand, P. W. Position of Transmembrane Helix 6 Determines Receptor G Protein Coupling Specificity. *J. Am. Chem. Soc.* **2014**, *136* (32), 11244.
- (18) Goetz, A.; Lanig, H.; Gmeiner, P.; Clark, T. Molecular Dynamics Simulations of the Effect of the G-Protein and Diffusible Ligands on the β 2-Adrenergic Receptor. *J. Mol. Biol.* **2011**, *414* (4), 611.
- (19) Bai, Q.; Zhang, Y.; Ban, Y.; Liu, H.; Yao, X. Computational Study on the Different Ligands Induced Conformation Change of β 2 Adrenergic Receptor-Gs Protein Complex. *Plos One* **2013**, *8* (7), e68138.
- (20) Pachov, D. V.; Fonseca, R.; Arnol, D.; Bernauer, J.; van den Bedem, H. Coupled motions in β 2AR: Gas conformational ensembles. *J. Chem. Theory Comput.* **2016**, *12* (3), 946.
- (21) Dror, R. O.; Mildorf, T. J.; Hilger, D.; Manglik, A.; Borhani, D. W.; Arlow, D. H.; Philippsen, A.; Villanueva, N.; Yang, Z.; Lerch, M. T. Structural basis for nucleotide exchange in heterotrimeric G proteins. *Science* **2015**, *348* (6241), 1361.
- (22) Louet, M.; Perahia, D.; Martinez, J.; Floquet, N. A Concerted Mechanism for Opening the GDP Binding Pocket and Release of the Nucleotide in Hetero-Trimeric G-Proteins. *J. Mol. Biol.* **2011**, *411* (1), 298.
- (23) Louet, M.; Martinez, J.; Floquet, N. GDP Release Preferentially Occurs on the Phosphate Side in Heterotrimeric G-proteins. *PLOS Comput. Biol.* **2012**, *8* (7), e1002595.
- (24) Saleh, N.; Saladino, G.; Gervasio, F. L.; Clark, T. Investigating allosteric effects on the functional dynamics of β 2-adrenergic ternary complexes with enhanced-sampling simulations. *Chem. Sci.* **2017**, *8* (5), 4019.
- (25) Saleh, N.; Ibrahim, P.; Clark, T. Differences between G-Protein-Stabilized Agonist-GPCR Complexes and their Nanobody-Stabilized Equivalents. *Angew. Chem. Int. Ed.* **2017**, *56* (31), 9008.
- (26) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11* (8), 3584.
- (27) Pang, Y. T.; Miao, Y.; Wang, Y.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics in NAMD. *J Chem Theory Comput* **2017**, *13* (1), 9.
- (28) Miao, Y.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Theory, Implementation and Applications. *Annu Rep Comp Chem* **2017**, *13*, 231.
- (29) Miao, Y.; McCammon, J. A. Graded activation and free energy landscapes of a muscarinic G-protein–coupled receptor. *Proc Natl Acad Sci* **2016**, *113* (43), 12162.
- (30) Palermo, G.; Miao, Y.; Walker, R. C.; Jinek, M.; McCammon, J. A. CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations. *Proc Natl Acad Sci* **2017**, *114* (28), 7260.
- (31) Sibener, L. V.; Fernandes, R. A.; Kolawole, E. M.; Carbone, C. B.; Liu, F.; McAfee, D.; Birnbaum, M. E.; Yang, X.; Su, L. F.; Yu, W. et al. Isolation of a Structural Mechanism for

- Uncoupling T Cell Receptor Signaling from Peptide-MHC Binding. *Cell* **2018**, *174* (3), 672.
- (32) Salawu, E. O. The Impairment of TorsinA's Binding to and Interactions With Its Activator: An Atomistic Molecular Dynamics Study of Primary Dystonia. *Front Mol Biosci* **2018**, *5*, 64.
 - (33) Miao, Y.; McCammon, J. A. Mechanism of the G-protein mimetic nanobody binding to a muscarinic G-protein-coupled receptor. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (12), 3036.
 - (34) Nguyen, A. T. N.; Baltos, J.-A.; Thomas, T.; Nguyen, T. D.; Muñoz, L. L.; Gregory, K. J.; White, P. J.; Sexton, P. M.; Christopoulos, A.; May, L. T. Extracellular Loop 2 of the Adenosine A1 Receptor Has a Key Role in Orthosteric Ligand Affinity and Agonist Efficacy. *Mol. Pharmacol.* **2016**, *90* (6), 703.
 - (35) Oldham, W. M.; Hamm, H. E. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat. Rev. Mol. Cell Biol.* **2008**, *9* (1), 60.
 - (36) Glukhova, A.; Thal, D. M.; Nguyen, A. T.; Vecchio, E. A.; Jörg, M.; Scammells, P. J.; May, L. T.; Sexton, P. M.; Christopoulos, A. Structure of the Adenosine A1 Receptor Reveals the Basis for Subtype Selectivity. *Cell* **2017**, *168* (5), 867.
 - (37) Ballesteros, J. A.; Weinstein, H. In *Methods in Neurosciences*; Sealfon, S. C., Ed.; Academic Press, 1995; Vol. 25.
 - (38) Cheng, R. K. Y.; Segala, E.; Robertson, N.; Deflorian, F.; Doré, A. S.; Errey, J. C.; Fiez-Vandal, C.; Marshall, F. H.; Cooke, R. M. Structures of Human A1 and A2A Adenosine Receptors with Xanthines Reveal Determinants of Selectivity. *Structure* **2017**, *25* (8), 1275.
 - (39) Peeters, M. C.; Wisse, L. E.; Dinaj, A.; Vrolijk, B.; Vriend, G.; Ijzerman, A. P. The role of the second and third extracellular loops of the adenosine A1 receptor in activation and allosteric modulation. *Biochem. Pharmacol.* **2012**, *84* (1), 76.
 - (40) Flock, T.; Ravarani, C. N. J.; Sun, D.; Venkatakrishnan, A. J.; Kayikci, M.; Tate, C. G.; Veprintsev, D. B.; Babu, M. M. Universal allosteric mechanism for Gα activation by GPCRs. *Nature* **2015**, *524*, 173.
 - (41) Nair, V.; Fesbender, A. J.; Miller, L. P.; Bruce, J. L. High selectivity of novel isoguanosine analogues for the adenosine A1 receptor. *Bioorganic Med. Chem. Lett.* **1991**, *1* (9), 481.
 - (42) de Lera Ruiz, M.; Lim, Y.-H.; Zheng, J. Adenosine A2A Receptor as a Drug Discovery Target. *J. Med. Chem.* **2014**, *57* (9), 3623.
 - (43) Sun, B.; Bachhawat, P.; Chu, M. L.-H.; Wood, M.; Ceska, T.; Sands, Z. A.; Mercier, J.; Lebon, F.; Kobilka, T. S.; Kobilka, B. K. Crystal structure of the adenosine A2A receptor bound to an antagonist reveals a potential allosteric pocket. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (8), 2066.
 - (44) Nguyen, A. T. N.; Vecchio, E. A.; Thomas, T.; Nguyen, T. D.; Aurelio, L.; Scammells, P. J.; White, P. J.; Sexton, P. M.; Gregory, K. J.; May, L. T. et al. Role of the Second Extracellular Loop of the Adenosine A1 Receptor on Allosteric Modulator Binding, Signaling, and Cooperativity. *Mol. Pharmacol.* **2016**, *90* (6), 715.
 - (45) Miao, Y.; Bhattarai, A.; Nguyen, A. T. N.; Christopoulos, A.; May, L. T. Structural Basis for Binding of Allosteric Drug Leads in the Adenosine A1 Receptor. *Sci. Rep.* **2018**, *8* (1), 16836.
 - (46) Wang, J.; Miao, Y. Recent advances in computational studies of GPCR-G protein interactions. *Advances in Protein Chemistry & Structural Biology* **2019**, *In Press*, DOI: 10.1016/bs.apcsb.2018.11.011.
 - (47) Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J. Chem. Theory Comput.* **2014**, *10* (7), 2677.

- (48) Carpenter, B.; Nehmé, R.; Warne, T.; Leslie, A. G. W.; Tate, C. G. Structure of the adenosine A2A receptor bound to an engineered G protein. *Nature* **2016**, 536, 104.
- (49) Dror, R. O.; Arlow, D. H.; Maragakis, P.; Mildorf, T. J.; Pan, A. C.; Xu, H.; Borhani, D. W.; Shaw, D. E. Activation mechanism of the β 2-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, 108 (46), 18684.
- (50) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, 14 (1), 33.
- (51) Vanommeslaeghe, K.; MacKerell, A. D. CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochim. Biophys. Acta. Gen. Subj.* **2015**, 1850 (5), 861.
- (52) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, 52 (12), 3155.
- (53) Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, 52 (12), 3144.
- (54) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, 26 (16), 1781.
- (55) Case D, et al. . Amber 18 (University of California, San Francisco). **2018**.
- (56) Roe, D. R.; Cheatham, T. E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, 9 (7), 3084.
- (57) Venkatakrishnan, A. J.; Deupi, X.; Lebon, G.; Tate, C. G.; Schertler, G. F.; Babu, M. M. Molecular signatures of G-protein-coupled receptors. *Nature* **2013**, 494 (7436), 185.
- (58) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, 13 (11), 2498.

Table 1 Summary of GaMD simulations performed on the agonist-bound AR-G protein complexes.

^aThe AR:NPxxY-G:α5 distance is the center-of-mass (COM) distance between the receptor NPxxY motif and the last 5 residues of the G_α α5 helix. ^bThe R3.50-E6.30 distance is the distance between the Cα atoms of conserved residues Arg^{3.50} and Glu^{6.30} in the receptors. ^cThe α5 orientation angle is the angle between COMs of the receptor orthosteric pocket, the last 5 and first 5 residues of the G_α α5 helix, illustrated in **Figure S8**. ^dThe increase of G_α-G_β distance is the increase in the distance between the COMs of G_α (excluding the N-terminal helix) and G_β (excluding the C-terminal of β sheet) subunits compared to the cryo-EM structure.

System	ADO-A ₁ AR-G _i	NECA-A _{2A} AR-G _s	NECA-A _{2A} AR-G _i	ADO-A ₁ AR-G _s
Dimension (Å ³)	112x124x150	110x124x148	110x124x148	108x124x153
<i>N</i> _{atoms}	180,394	175,140	176,028	176,940
Simulation Length (ns)	300 x 3	300 x 3	300 x 3	300 x 3
Boost Potential (kcal/mol)	21.43 ± 6.50	21.12 ± 6.49	20.96 ± 6.41	21.88 ± 6.58
<i>Cryo-EM experimental structures</i>				
AR:NPxxY-G:α5 Distance ^a (Å)	12.7	13.6	-	-
R3.50-E6.30 Distance ^b (Å)	16.3	18.4	-	-
α5 Orientation Angle ^c (°)	143.9	123.5	-	-
<i>GaMD simulation low-energy conformations</i>				
AR:NPxxY-G:α5 Distance (Å)	12.2	13.2	11.2	11.8, 13.5
R3.50-E6.30 Distance (Å)	17.4	18.2	17.6	17.5, 22.5
α5 Orientation Angle (°)	135.2	127.5	140.2	123.0, 130.0
Increase of G _α -G _β Distance ^d (Å)	-1.3	1.5	-0.8	1.2, 3.2

Figure 1. Comparison of structural flexibility of active adenosine receptor-G protein complexes obtained from GaMD simulations: **(A)** Change in the root-mean-square fluctuations (RMSFs) of the A₁AR when the G_i protein was changed to the G_s protein. **(B)** Change in the RMSFs of the A_{2A}AR when the G_s protein was changed to the G_i protein. **(C)** Change in the RMSFs of the G_s protein when the receptor was changed from the A_{2A}AR to the A₁AR. **(D)** Change in the RMSFs of the G_i protein when the receptor was changed from the A₁AR to the A_{2A}AR. A color scale of -0.5 Å (blue) to 1.5 Å (red) is used.

Figure 2. 2D potential of mean force (PMF) profiles of the (A) ADO-A₁AR-G_i, (B) NECA-A_{2A}AR-G_s, (C) NECA-A_{2A}AR-G_i and (D) ADO-A₁AR-G_s complex systems regarding the agonist RMSD relative to the cryo-EM conformation and AR:NPxxY-G:α5 distance.

Figure 3. Distinct active conformations of the A₁AR induced by binding of the G_s protein: (A) Representative conformations of two low-energy binding poses of ADO (L1 and L2 in green and red, respectively) in the A₁AR-G_s complex. The cryo-EM structure of the A₁AR-G_i complex (PDB: 6D9H, blue) is shown for comparison. (B) Representative conformations of open (PDB: 6D9H, blue), semi-open (green) and closed (red) states of ECL2 in the A₁AR. (C) Representative conformations of the A₁AR in the "Active" (blue) and "Over-active" (red) states in the A₁AR-G_s system.

Figure 4. Distinct low-energy conformational states of the ADO-A₁AR-G_s system were identified from GaMD simulations. **(A)** 2D PMF of RMSD of the helix region in ECL2 relative to the cryo-EM structure and the A₁AR:NPxxY-G:α5 distance. **(B)** 2D PMF of the R3.50-E6.30 and A₁AR:NPxxY-G:α5 distances. **(C)** 2D PMF of the orientation angle of the Gα α5 helix and the R3.50-E6.30 distance. **(D)** 2D PMF of the increase of the Gα-Gβ distance and the R3.50-E6.30 distance.

Figure 5. Complementary residue interactions at the protein interface in the (A) ADO-A₁AR-G_i, (B) NECA-A_{2A}AR-G_s, (C) NECA-A_{2A}AR-G_i and (D) ADO-A₁AR-G_s (active state) systems. Receptor and G-protein residues are numbered according to the BW and CGN schemes, respectively. Hydrogen bond, van der Waals and salt-bridge interactions are colored in red, black and blue, respectively. The line thickness is proportional to the number of residue interactions. Hydrophobic, polar, acidic and basic residues are colored in green, yellow, red and blue, respectively.

Figure 6. Summary of specific AR-G protein interactions: **(A)** the ADO-bound A₁AR prefers to bind the G_i protein to the G_s. The latter could not stabilize agonist ADO binding in the A₁AR and tended to dissociate from the receptor. **(B)** The A_{2A}AR could bind both the G_s and G_i proteins, which adopted distinct conformations in the complexes.